

CHƯƠNG 5: KIỂM ĐỊNH GIẢ THUYẾT

Bài giảng môn Xác suất Thống kê

Giảng viên: TS. Nguyễn Kiều Linh

Hà Nội, năm 2021

1 BÀI 1: MỘT SỐ KHÁI NIỆM

- 1. Giả thuyết thống kê
- 2. Tiêu chuẩn kiểm định giả thuyết thống kê
- 3. Miền bác bỏ giả thuyết
- 4. Giá trị quan sát của tiêu chuẩn kiểm định
- 5. Quy tắc kiểm định giả thuyết thống kê
- 6. Các bước kiểm định giả thuyết thống kê

2 BÀI 2: KIỂM ĐỊNH Giả thuyết MỘT MẪU

- 1. Kiểm định giả thuyết về giá trị trung bình
- 2. Kiểm định giả thuyết về tỉ lệ

Giới thiệu Chương 5

Trong chương 4 ta đã giải quyết bài toán ước lượng các tham số đặc trưng của biến ngẫu nhiên gốc của tổng thể. Trong chương này ta sẽ sử dụng phương pháp kiểm định giả thuyết thống kê để kiểm định các tham số đặc trưng của tổng thể.

- Một dạng khác của quy nạp thống kê là kiểm định giả thuyết thống kê. Đây là một phương pháp quan trọng cho phép giải quyết nhiều bài toán trong thực tế.
- Nội dung của kiểm định giả thuyết thống kê là dựa vào mẫu cụ thể và các quy tắc hay thủ tục quyết định dẫn đến bác bỏ hay chấp nhận giả thuyết của tổng thể.

1 BÀI 1: MỘT SỐ KHÁI NIỆM

- 1. Giả thuyết thống kê
- 2. Tiêu chuẩn kiểm định giả thuyết thống kê
- 3. Miền bác bỏ giả thuyết
- 4. Giá trị quan sát của tiêu chuẩn kiểm định
- 5. Quy tắc kiểm định giả thuyết thống kê
- 6. Các bước kiểm định giả thuyết thống kê

2 BÀI 2: KIỂM ĐỊNH Giả thuyết MỘT MẪU

- 1. Kiểm định giả thuyết về giá trị trung bình
- 2. Kiểm định giả thuyết về tỉ lệ

1. Giả thuyết thống kê

Giả thuyết thống kê

- Giả thuyết thống kê là giả thuyết về biến ngẫu nhiên gốc của tổng thể, bao gồm: dạng phân bố xác suất, các đặc trưng tham số của biến ngẫu nhiên gốc hoặc giả thuyết về sự độc lập của các biến ngẫu nhiên gốc.
- Giả thuyết đưa ra kiểm nghiệm được ký hiệu là H_0 , gọi là “giả thuyết không” (Null hypothesis). Đó là giả thuyết mà ta nghi ngờ muốn bác bỏ hoặc giả thuyết ta muốn bảo vệ. Ngoài giả thuyết H_0 ra, ta còn phải định ra một giả thuyết cạnh tranh với H_0 gọi là giả thuyết đối (Alternative hypothesis), ký hiệu H_1 . Giả thuyết đối H_1 sẽ được chấp nhận khi H_0 bị bác bỏ.

1. Giả thuyết thống kê

Chú ý

Giả thuyết đối H_1 không nhất thiết là phủ định của giả thuyết H_0 . Chẳng hạn giả thuyết H_0 : nhu cầu thị trường về loại hàng hóa này là $\mu = 1000$ đơn vị/tháng. Nếu ta nghi ngờ rằng nhu cầu này không đúng thì giả thuyết đối H_1 là $\mu \neq 1000$, nhưng nếu do tiếp thị tốt, do chính sách hậu mãi tốt người ta nghĩ rằng nhu cầu về mặt hàng này tăng lên thì đối thiết H_1 là $\mu > 1000$.

1 BÀI 1: MỘT SỐ KHÁI NIỆM

- 1. Giả thuyết thống kê
- 2. Tiêu chuẩn kiểm định giả thuyết thống kê
- 3. Miền bác bỏ giả thuyết
- 4. Giá trị quan sát của tiêu chuẩn kiểm định
- 5. Quy tắc kiểm định giả thuyết thống kê
- 6. Các bước kiểm định giả thuyết thống kê

2 BÀI 2: KIỂM ĐỊNH Giả thuyết MỘT MẪU

- 1. Kiểm định giả thuyết về giá trị trung bình
- 2. Kiểm định giả thuyết về tỉ lệ

2. Tiêu chuẩn kiểm định giả thuyết thống kê

Quy tắc kiểm định

Quy tắc kiểm định dựa trên hai nguyên lý sau:

- Nguyên lý xác suất nhỏ: "Nếu một biến cố có xác suất nhỏ thì trong một phép thử thì biến cố đó coi như không xảy ra".
- Phương pháp phản chứng: "Để bác bỏ A ta giả sử A đúng; nếu A đúng dẫn đến một điều vô lý thì ta bác bỏ A".

2. Tiêu chuẩn kiểm định giả thuyết thống kê

Phương pháp chung của kiểm định

Dựa vào hai nguyên lý trên ta đưa ra phương pháp chung để kiểm định một giả thuyết thống kê như sau:

- Để kiểm định H_0 trước hết giả sử H_0 đúng, từ đó ta tìm được biến cố A mà xác suất xuất hiện biến cố A là rất bé và ta có thể xem A không thể xảy ra trong một phép thử về biến cố này.
- Khi đó nếu trên một mẫu cụ thể quan sát được mà biến cố A xuất hiện thì điều này trái với nguyên lý xác suất nhỏ. Vậy H_0 sai và bác bỏ nó. Còn nếu A không xảy ra thì ta chưa có cơ sở để bác bỏ H_0 .

2. Tiêu chuẩn kiểm định giả thuyết thống kê

Phương pháp cụ thể để kiểm định

Ta thực hiện phương pháp trên bằng các bước cụ thể sau:

- Từ biến ngẫu nhiên gốc X của tổng thể lập mẫu ngẫu nhiên $W = (X_1, X_2, \dots, X_n)$.
- Chọn thống kê $T = T(X_1, X_2, \dots, X_n, \theta)$, trong đó θ là tham số liên quan đến giả thuyết cần kiểm định.
- Nếu H_0 đúng thì thống kê T có quy luật phân bố xác suất xác định. Thống kê được gọi là *tiêu chuẩn kiểm định*.

1 BÀI 1: MỘT SỐ KHÁI NIỆM

- 1. Giả thuyết thống kê
- 2. Tiêu chuẩn kiểm định giả thuyết thống kê
- 3. Miền bác bỏ giả thuyết
- 4. Giá trị quan sát của tiêu chuẩn kiểm định
- 5. Quy tắc kiểm định giả thuyết thống kê
- 6. Các bước kiểm định giả thuyết thống kê

2 BÀI 2: KIỂM ĐỊNH Giả thuyết MỘT MẪU

- 1. Kiểm định giả thuyết về giá trị trung bình
- 2. Kiểm định giả thuyết về tỉ lệ

3. Miền bác bỏ giả thuyết

Miền bác bỏ giả thuyết

- Sau khi đã chọn tiêu chuẩn kiểm định T , với α bé cho trước (thường α được lấy bằng 0,05 hoặc 0,01) và với giả thuyết H_0 đúng ta có thể tìm được miền W_α sao cho T nhận giá trị trong miền W_α với xác suất bằng α :

$$P(T \in W_\alpha | H_0) = \alpha.$$

- Giá trị α được gọi là *mức ý nghĩa* của tiêu chuẩn kiểm định và miền W_α gọi là *miền bác bỏ* giả thuyết H_0 với mức ý nghĩa α .

1 BÀI 1: MỘT SỐ KHÁI NIỆM

- 1. Giả thuyết thống kê
- 2. Tiêu chuẩn kiểm định giả thuyết thống kê
- 3. Miền bác bỏ giả thuyết
- 4. Giá trị quan sát của tiêu chuẩn kiểm định
- 5. Quy tắc kiểm định giả thuyết thống kê
- 6. Các bước kiểm định giả thuyết thống kê

2 BÀI 2: KIỂM ĐỊNH Giả thuyết MỘT MẪU

- 1. Kiểm định giả thuyết về giá trị trung bình
- 2. Kiểm định giả thuyết về tỉ lệ

4. Giá trị quan sát của tiêu chuẩn kiểm định

Giá trị quan sát của tiêu chuẩn kiểm định

Thực hiện phép thử với mẫu ngẫu nhiên $W = (X_1, X_2, \dots, X_n)$ thu được mẫu cụ thể $w = (x_1, x_2, \dots, x_n)$, thay giá trị này vào thống kê T ta được giá trị quan sát của tiêu chuẩn kiểm định:

$$T_{qs} = (x_1, x_2, \dots, x_n, \theta).$$

1 BÀI 1: MỘT SỐ KHÁI NIỆM

- 1. Giả thuyết thống kê
- 2. Tiêu chuẩn kiểm định giả thuyết thống kê
- 3. Miền bác bỏ giả thuyết
- 4. Giá trị quan sát của tiêu chuẩn kiểm định
- 5. Quy tắc kiểm định giả thuyết thống kê
- 6. Các bước kiểm định giả thuyết thống kê

2 BÀI 2: KIỂM ĐỊNH Giả thuyết MỘT MẪU

- 1. Kiểm định giả thuyết về giá trị trung bình
- 2. Kiểm định giả thuyết về tỉ lệ

5. Quy tắc kiểm định giả thuyết thống kê

So sánh giá trị quan sát của tiêu chuẩn kiểm định với miền bác bỏ và kết luận theo quy tắc sau:

Quy tắc kiểm định giả thuyết thống kê

- Nếu $T_{qs} \in W_\alpha$, theo nguyên tắc kiểm định thì H_0 sai, do đó ta bác bỏ H_0 thừa nhận H_1 .
- Nếu $T_{qs} \notin W_\alpha$ thì điều này chưa khẳng định rằng H_0 đúng mà chỉ có nghĩa là qua mẫu cụ thể này chưa khẳng định được là H_0 sai. Do đó ta chỉ có thể nói rằng qua mẫu cụ thể này chưa có cơ sở để bác bỏ H_0 (trên thực tế là vẫn thừa nhận H_0).

1 BÀI 1: MỘT SỐ KHÁI NIỆM

- 1. Giả thuyết thống kê
- 2. Tiêu chuẩn kiểm định giả thuyết thống kê
- 3. Miền bác bỏ giả thuyết
- 4. Giá trị quan sát của tiêu chuẩn kiểm định
- 5. Quy tắc kiểm định giả thuyết thống kê
- 6. Các bước kiểm định giả thuyết thống kê

2 BÀI 2: KIỂM ĐỊNH Giả thuyết MỘT MẪU

- 1. Kiểm định giả thuyết về giá trị trung bình
- 2. Kiểm định giả thuyết về tỉ lệ

6. Các bước kiểm định giả thuyết thống kê

Qua nội dung trình bày ở trên ta có thể xây dựng các bước kiểm định giả thuyết thống kê như sau:

Các bước kiểm định giả thuyết thống kê

- Phát biểu giả thuyết H_0 và đối thiết H_1 .
- Từ tổng thể nghiên cứu lập mẫu ngẫu nhiên kích thước n .
- Chọn tiêu chuẩn kiểm định T và xác định quy luật phân bố xác suất của T với điều kiện giả thuyết H_0 đúng.
- Với mức ý nghĩa α , xác định miền bác bỏ W_α tốt nhất tùy thuộc vào giả thuyết đối H_1 .
- Từ mẫu cụ thể tính giá trị quan sát của tiêu chuẩn kiểm định T_{qs} .
- So sánh giá trị quan sát của tiêu chuẩn kiểm định T_{qs} với miền bác bỏ W_α và kết luận.

1 BÀI 1: MỘT SỐ KHÁI NIỆM

- 1. Giả thuyết thống kê
- 2. Tiêu chuẩn kiểm định giả thuyết thống kê
- 3. Miền bác bỏ giả thuyết
- 4. Giá trị quan sát của tiêu chuẩn kiểm định
- 5. Quy tắc kiểm định giả thuyết thống kê
- 6. Các bước kiểm định giả thuyết thống kê

2 BÀI 2: KIỂM ĐỊNH Giả thuyết MỘT MẪU

- 1. Kiểm định giả thuyết về giá trị trung bình
- 2. Kiểm định giả thuyết về tỉ lệ

1. Kiểm định giả thuyết về giá trị trung bình

Bài toán

- Giả sử biến ngẫu nhiên X có phân bố chuẩn $N(\mu, \delta^2)$ nhưng ta chưa biết kỳ vọng $E(X) = \mu$ của X . Nếu có cơ sở để giả thuyết rằng giá trị của nó bằng μ_0 ta đưa ra giả thuyết thống kê:

$$H_0 : \mu = \mu_0.$$

- Cần kiểm tra giả thuyết này với các giả thuyết đối

$$H_1 : \mu \neq \mu_0, \quad H_1 : \mu > \mu_0, \quad H_1 : \mu < \mu_0.$$

- Xét bài toán trong ba trường hợp:

- Trường hợp 1: Đã biết phương sai $DX = \delta^2$;
- Trường hợp 2: Chưa biết phương sai $DX = \delta^2$, $n \geq 30$;
- Trường hợp 3: Chưa biết phương sai $DX = \delta^2$, $n < 30$;

1.1. Trường hợp 1: Đã biết phương sai $DX = \sigma^2$

Trường hợp 1: Đã biết phương sai $DX = \sigma^2$

Với mức ý nghĩa α cho trước ta xây dựng miền bác bỏ phụ thuộc vào giả thuyết đối H_1 như sau:

- Bài toán 1: $H_0 : \mu = \mu_0; H_1 : \mu \neq \mu_0$.

Miền bác bỏ

$$W_\alpha = (-\infty, -u_{\frac{\alpha}{2}}] \cup [u_{\frac{\alpha}{2}}, +\infty).$$

- Bài toán 2: $H_0 : \mu = \mu_0; H_1 : \mu > \mu_0$.

Miền bác bỏ

$$W_\alpha = [u_\alpha, +\infty).$$

- Bài toán 3: $H_0 : \mu = \mu_0; H_1 : \mu < \mu_0$.

Miền bác bỏ

$$W_\alpha = (-\infty, -u_\alpha].$$

1.1. Trường hợp 1: Đã biết phương sai $DX = \delta^2$

Trường hợp 1: Đã biết phương sai $DX = \delta^2$

Giá trị quan sát

$$u_{qs} = \frac{(\bar{x} - \mu_0)\sqrt{n}}{\delta}.$$

Ta kiểm tra u_{qs} có thuộc miền bác bỏ W_α không để kết luận

- Nếu $u_{qs} \in W_\alpha$ thì ta bác bỏ H_0 , thừa nhận H_1 .
- Nếu $u_{qs} \notin W_\alpha$ thì ta chưa có cơ sở bác bỏ giả thuyết H_0 , tức là chưa có cơ sở để thừa nhận giả thuyết H_1 .

1.1. Trường hợp 1: Đã biết phương sai $DX = \delta^2$

Lời giải

Ta có $\bar{x} = \frac{49.10 + 50.60 + 51.20 + 52.5 + 53.5}{100} = 50,35$.

Ta kiểm định giả thuyết: $H_0 : \mu = 50$; $H_1 \neq 50$.

Ta có $\frac{\alpha}{2} = \frac{0,05}{2} = 0,025$. Tra bảng ta được $u_{\frac{\alpha}{2}} = 1,96$. Do đó ta xây dựng được miền bác bỏ như sau:

$$\begin{aligned}W_{\alpha} &= (-\infty, -u_{\frac{\alpha}{2}}] \cup [u_{\frac{\alpha}{2}}, +\infty). \\&= (-\infty; -1,96] \cup [1,96; +\infty)\end{aligned}$$

Giá trị quan sát $u_{qs} = \frac{(\bar{x} - \mu_0)\sqrt{n}}{\delta} = \frac{(50,35 - 50)\sqrt{100}}{2} = 1,75$.

Ta thấy $u_{qs} \notin W_{\alpha}$, ta chưa có cơ sở bác bỏ giả thuyết H_0 , tức là chưa có cơ sở để thừa nhận giả thuyết H_1 . Vậy điều nghi ngờ là sai.

1.2. Trường hợp 2: Chưa biết $DX = \delta^2$, $n \geq 30$

Trường hợp 2: Chưa biết phương sai $DX = \delta^2$, $n \geq 30$

Trong trường hợp này thì miền bác bỏ W_α và quy tắc kiểm định y hệt như Trường hợp 1, chỉ khác ở chỗ giá trị quan sát được tính theo công thức

$$u_{qs} = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s}.$$

1.2. Trường hợp 2: Chưa biết $DX = \delta^2$, $n \geq 30$

Ví dụ

Lượng nước sạch của một gia đình 4 người ở Hà Nội sử dụng trong 6 tháng năm ngoái là 17 m^3 . Thao dõi lượng nước sạch sử dụng trong 6 tháng năm nay của 60 gia đình 4 người thu được số liệu sau:

Lượng nước sạch (m^3)	15-16	16-17	17-18	18-19	19-20
Số gia đình tương ứng	7	15	21	12	5

Giả sử lượng nước sạch tiêu thụ của các hộ gia đình là một biến ngẫu nhiên có phân bố chuẩn.

- a) Hãy ước lượng bằng khoảng tin cậy lượng nước sạch trung bình của các hộ sử dụng trong 6 tháng năm nay với độ tin cậy 95%.
- b) Có ý kiến cho rằng lượng nước tiêu thụ năm nay tăng lên. Sử dụng bảng số liệu trên hãy kiểm định ý kiến đó với mức ý nghĩa 2,5%.

1.2. Trường hợp 2: Chưa biết $DX = \delta^2$, $n \geq 30$

Lời giải

Thực hiện phép đổi biến $u_i = \frac{x_i - 17,5}{1}$ với $x_0 = 17,5$ và $h = 1$. Ta có bảng tính sau:

x_i	u_i	n_i	$n_i u_i$	$n_i u_i^2$
15,5	-2	7	-14	28
16,5	-1	15	-15	15
17,5	0	21	0	0
18,5	1	12	12	12
19,5	2	5	10	20
\sum		$n = 60$	-7	75

Khi đó
$$\bar{x} = x_0 + h\bar{u} = 17,5 + 1 \frac{-7}{60} = 17,38 \quad (1)$$

$$s^2 = h^2 s_u^2 = 1^2 \frac{1}{59} \left(75 - \frac{(-7)^2}{60} \right) = \frac{4451}{3540}. \quad (2)$$

$$s = \sqrt{s^2} \simeq 1,12. \quad (3)$$

1.2. Trường hợp 2: Chưa biết $DX = \delta^2$, $n \geq 30$

Lời giải

a) Độ tin cậy 95%, suy ra $1 - \alpha = 0,95$ hay $\alpha = 0,05$. Khi đó $\frac{\alpha}{2} = 0,025$. Tra bảng ta được $u_{\frac{\alpha}{2}} = 1,96$.

Độ chính xác của ước lượng là

$$\epsilon = u_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} = 1,96 \frac{1,12}{\sqrt{60}} = 0,28.$$

Khoảng tin cậy của lượng nước sạch trung bình của các hộ sử dụng trong 6 tháng năm nay

$$(\bar{x} - \epsilon, \bar{x} + \epsilon) = (17,10; 17,66).$$

1.2. Trường hợp 2: Chưa biết $DX = \sigma^2$, $n \geq 30$

Lời giải

b) Ta kiểm định giả thuyết: $H_0 : \mu = 17$; $H_1 > 17$.

Ta có $\alpha = 0,025$. Tra bảng ta được $u_\alpha = 1,96$. Do đó ta xây dựng được miền bác bỏ như sau:

$$W_\alpha = [u_\alpha, +\infty) = [1,96; +\infty).$$

Giá trị quan sát $u_{qs} = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s} = \frac{(17,38 - 17)\sqrt{60}}{1,12} = 2,63$.

Ta thấy $u_{qs} \in W_\alpha$. Vậy ta bác bỏ H_0 , thừa nhận H_1 . Do đó lượng nước tiêu thụ năm nay tăng lên.

1.3. Trường hợp 3: Chưa biết $DX = \delta^2$, $n < 30$

Trường hợp 3: Chưa biết phương sai $DX = \delta^2$, $n < 30$

Với mức ý nghĩa α cho trước ta xây dựng miền bác bỏ phụ thuộc vào giả thuyết đối H_1 như sau:

- Bài toán 1: $H_0 : \mu = \mu_0; H_1 : \mu \neq \mu_0$.

Miền bác bỏ

$$W_\alpha = (-\infty, -t_{\frac{\alpha}{2}}(n-1)] \cup [t_{\frac{\alpha}{2}}(n-1), +\infty).$$

- Bài toán 2: $H_0 : \mu = \mu_0; H_1 : \mu > \mu_0$.

Miền bác bỏ

$$W_\alpha = [t_\alpha(n-1), +\infty).$$

- Bài toán 3: $H_0 : \mu = \mu_0; H_1 : \mu < \mu_0$.

Miền bác bỏ

$$W_\alpha = (-\infty, -t_\alpha(n-1)].$$

1.3. Trường hợp 3: Chưa biết $DX = \delta^2$, $n < 30$

Trường hợp 3: Chưa biết phương sai $DX = \delta^2$, $n < 30$

Giá trị quan sát

$$t_{qs} = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s}.$$

Ta kiểm tra t_{qs} có thuộc miền bác bỏ W_α không để kết luận

- Nếu $t_{qs} \in W_\alpha$ thì ta bác bỏ H_0 , thừa nhận H_1 .
- Nếu $t_{qs} \notin W_\alpha$ thì ta chưa có cơ sở bác bỏ giả thuyết H_0 , tức là chưa có cơ sở để thừa nhận giả thuyết H_1 .

1.3. Trường hợp 3: Chưa biết $DX = \delta^2$, $n < 30$

Lời giải

Lập bảng tính toán

x_i	n_i	$n_i x_i$	$n_i x_i^2$
8,5	5	42,5	361,25
9	8	72	648
11	10	110	1210
12,5	2	25	312,5
\sum	$n = 25$	249,5	2531,75

Khi đó
$$\bar{x} = \frac{249,5}{25} = 9,98 \quad (4)$$

$$s^2 = \frac{1}{24} \left(2531,75 - \frac{(249,5)^2}{25} \right) \simeq 1,7392. \quad (5)$$

$$s = \sqrt{s^2} \simeq 1,32. \quad (6)$$

1.3. Trường hợp 3: Chưa biết $DX = \delta^2$, $n < 30$

Lời giải

Ta kiểm định giả thuyết: $H_0 : \mu = 11$; $H_1 < 11$.

Ta có $\alpha = 0,05$. Tra bảng ta được $t_\alpha(n-1) = t_{0,05}(24) = -1,711$. Do đó ta xây dựng được miền bác bỏ như sau:

$$W_\alpha = [-\infty; u_\alpha) = [-\infty; -1,711).$$

Giá trị quan sát $t_{qs} = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s} = \frac{(9,98 - 11)\sqrt{25}}{1,32} = -3,86$.

Ta thấy $u_{qs} \in W_\alpha$. Vậy ta bác bỏ H_0 , thừa nhận H_1 . Vậy có thể kết luận mức xăng tiêu hao trung bình thấp hơn 11 lít

1 BÀI 1: MỘT SỐ KHÁI NIỆM

- 1. Giả thuyết thống kê
- 2. Tiêu chuẩn kiểm định giả thuyết thống kê
- 3. Miền bác bỏ giả thuyết
- 4. Giá trị quan sát của tiêu chuẩn kiểm định
- 5. Quy tắc kiểm định giả thuyết thống kê
- 6. Các bước kiểm định giả thuyết thống kê

2 BÀI 2: KIỂM ĐỊNH Giả thuyết MỘT MẪU

- 1. Kiểm định giả thuyết về giá trị trung bình
- 2. Kiểm định giả thuyết về tỉ lệ

2. Kiểm định giả thuyết về tỉ lệ

Kiểm định giả thuyết về tỉ lệ

- Giả sử tổng thể gồm hai loại phần tử, có tính chất A và không có tính chất A , trong đó tỉ lệ có phần tử có tính chất A là p . Thông thường p chưa biết, nhưng có cơ sở để giả thuyết rằng giá trị của nó bằng p_0 , ta đưa ra giả thuyết thống kê $H_0 : p = p_0$.
- Cần kiểm tra giả thuyết này với các giả thuyết đối

$$H_1 : p \neq p_0, \quad H_1 : p > p_0, \quad H_1 : p < p_0.$$

2. Kiểm định giả thuyết về tỉ lệ

Kiểm định giả thuyết về tỉ lệ

Mỗi mức ý nghĩa α cho trước, ta xây dựng miền bác bỏ phụ thuộc vào giả thuyết đối H_1 như sau:

- Bài toán 1: $H_0 : p = p_0; H_1 : p \neq p_0$.

Miền bác bỏ

$$W_\alpha = (-\infty, -u_{\frac{\alpha}{2}}] \cup [u_{\frac{\alpha}{2}}, +\infty).$$

- Bài toán 2: $H_0 : p = p_0; H_1 : p > p_0$.

Miền bác bỏ

$$W_\alpha = [u_\alpha, +\infty).$$

- Bài toán 3: $H_0 : p = p_0; H_1 : p < p_0$.

Miền bác bỏ

$$W_\alpha = (-\infty, -u_\alpha].$$

2. Kiểm định giả thuyết về tỉ lệ

Kiểm định giả thuyết về tỉ lệ

Giá trị quan sát

$$u_{qs} = \frac{(f - p_0)\sqrt{n}}{\sqrt{p_0(1 - p_0)}}.$$

Ta kiểm tra u_{qs} có thuộc miền bác bỏ W_α không để kết luận:

- Nếu $u_{qs} \in W_\alpha$ thì ta bác bỏ H_0 , thừa nhận H_1 .
- Nếu $u_{qs} \notin W_\alpha$ thì ta chưa có cơ sở bác bỏ giả thuyết H_0 , tức là chưa có cơ sở để thừa nhận giả thuyết H_1 .

2. Kiểm định giả thuyết về tỉ lệ

Ví dụ

Tỷ lệ khách hàng trở lại sử dụng dịch vụ của công ty là 60%. Có ý kiến cho rằng tỷ lệ này giảm do chính sách hậu mãi của công ty không tốt. Theo dõi ngẫu nhiên 300 khách hàng thấy có 162 khách hàng trở lại sử dụng dịch vụ của công ty. Hãy kết luận ý kiến trên với mức ý nghĩa $\alpha = 0,025$.

2. Kiểm định giả thuyết về tỉ lệ

Lời giải

Ta kiểm định giả thuyết $H_0 : p = 0,6$; $H_1 : p < 0,6$.

Ta có $\alpha = 0,025$. Tra bảng ta có $u_\alpha = 1,96$.

Miền bác bỏ

$$W_\alpha = (-\infty; -u_\alpha] = (-\infty; -1,96).$$

Ta có $f = \frac{162}{300} = 0,54$.

Giá trị quan sát

$$u_{qs} = \frac{(f - p_0)\sqrt{n}}{\sqrt{p_0(1 - p_0)}} = \frac{(0,54 - 0,6)\sqrt{300}}{\sqrt{0,6 \cdot (1 - 0,6)}} \simeq -2,12.$$

Ta thấy $u_{qs} \in W_\alpha$. Do đó ta bác bỏ H_0 , thừa nhận H_1 . Vậy số khách hàng quay trở lại công ty có giảm.