XÁC SUẤT THỐNG KÊ VÀ ỨNG DỤNG

KIỂM ĐỊNH GIẢ THIẾT THỐNG KÊ

Buổi học 8

KIỂM ĐỊNH GIẢ THIẾT

1. KHÁI NIỆM VỀ KIỂM ĐỊNH GIẢ THIẾT
4. KIỂM ĐỊNH ĐẶC BIỆT VỚI MẪU LỚN
5. KIỂM ĐỊNH ĐẶC BIỆT VỚI MẪU NHỎ
6. QUAN HỆ ƯỚC LƯỢNG VÀ KIỂM ĐỊNH
   ĐƯỜNG CONG ĐẶC TRƯNG–LỰC KIỂM ĐỊNH
7. KIỂM ĐỊNH CHI BÌNH PHƯƠNG
8. PHƯƠNG PHÁP ĐIỀU CHỈNH YATE
9. HỆ SỐ NGẪU NHIÊN

KHÁI NIỆM VỀ KIỂM ĐỊNH GIẢ THIẾT

1. CÁC QUYẾT ĐỊNH THỐNG KÊ
2. GIẢ THIẾT THỐNG KÊ–GIẢ THIẾT KHÔNG
4. KIỂM ĐỊNH GIẢ THIẾT VÀ Ý NGHĨA
5. SAI LẦM LOẠI I VÀ SAI LẦM LOẠI II
6. MỨC Ý NGHĨA
7. CÁC KIỂM ĐỊNH CÓ LIÊN QUAN ĐẾN PHÂN PHỐI CHUẨN
8. KIỂM ĐỊNH MỘT PHÍA VÀ HAI PHÍA
9. GIÁ TRỊ P

CÁC QUYẾT ĐỊNH THỐNG KÊ

Trong thực tế ta thường đưa ra một quyết định về tổng thể dựa trên cơ sở thông tin trên mẫu. Các quyết định như thế được gọi là *các quyết định thống kê*.

Ví dụ, trên cơ sở dữ liệu mẫu, ta quyết định

- liệu một loại huyết thanh mới thực sự có hiệu quả trong việc chữa khỏi một căn bệnh nào đó hay không?
- liệu một phương án giáo dục này có tốt hơn phương án khác hay không?
- liệu một đồng xu có công bằng không?

GIẢ THIẾT THỐNG KÊ
- GIẢI THIẾT KHÔNG

Nhằm đưa ra một quyết định, ta lập các giả thiết. Các giả thiết có thể đúng hoặc sai, được gọi là các *giả thiết thống kê*.

- Ví dụ, muốn quyết định xem đồng xu có công bằng hay không, ta lập giả thiết đồng xu công bằng, nghĩa là p = 0,5. Giả thiết này được gọi là *giả thiết không*, ký hiệu là H_0.
- Bất kỳ giả thiết nào khác với giả thiết không thì được gọi là *giả thiết phòng hờ* hay *giả thiết thay thế*. Ví dụ, nếu giả thiết không là p = 0,5 thì giả thiết phòng hờ có thể là p = 0,7, p ≠ 0,5 hay p > 0,5, ký hiệu là H_1.

KIỂM ĐỊNH GIẢ THIẾT THỐNG KÊ VÀ Ý NGHĨA

- Nếu dựa trên giả định cho rằng giả thiết không (H_0) đúng, nhưng các kết quả được quan sát trên một mẫu khác với điều mà chúng ta mong đợi của giả thiết H_0 thì ta có khuynh hướng bác bỏ giả thiết này (hay ít nhất không chấp nhận giả thiết trên).
- Ví dụ, trong 20 lần tung một đồng xu có 16 lần xuất hiện mặt ngửa thì ta có khuynh hướng bác bỏ giả thiết cho rằng đồng công bằng, mặc dù chúng ta có thể mắc sai lầm.
- Các thủ tục giúp chúng ta quyết định chấp nhận hay bác bỏ một giả thiết thì được gọi là *kiểm định giả thiết, kiểm định ý nghĩa hay luật kiểm định.*

SAI LẦM LOẠI I VÀ LOẠI II

- *Sai lầm loại I*: khi chúng ta bác bỏ một giả thiết mà giả thiết đó là đúng.
- *Sai lầm loại II*: ta chấp nhận một giả thiết mà giả thiết đó lẽ ra phải được bác bỏ.

Hướng giải quyết sai lầm trong quyết định

- Thiết kế một kiểm định sao cho khi đưa ra quyết định thì sai lầm xảy ra nhỏ nhất. Điều này khó thực hiện vì giảm sai lầm loại này thì tăng sai lầm loại khác nên có sự thỏa hiệp: hạn chế loại sai lầm nghiêm trọng hơn.
- Một phương pháp duy nhất hạn chế cả hai loại sai lầm là tăng kích thước mẫu, mà điều này có thể hoặc không thể được.

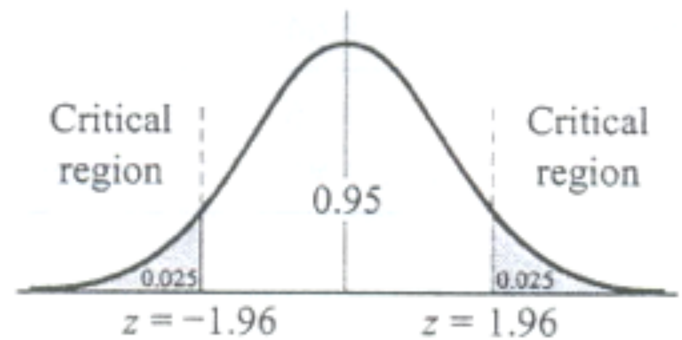
MỨC Ý NGHĨA

- Khi kiểm định một giả thiết, rủi ro mắc sai lầm loại I với xác suất lớn nhất được gọi là *mức ý nghĩa* của kiểm định (thường được xác định trước khi tiến hành chọn mẫu).
- Trong thực hành, mức ý nghĩa thường là 0,05 hay 0,01, tuy nhiên các giá trị khác cũng được sử dụng. Chẳng hạn, mức ý nghĩa được chọn để kiểm định một giả thiết là 0,05 hay 5%. ie, nếu giả thiết H_0 đúng thì chúng ta *tin chắc* 95% quyết định là đúng. Trong trường hợp này, ta nói giả thiết đã bị *bác bỏ ở 0,05 mức ý nghĩa*, hay chúng ta quyết định sai lầm với xác suất 0,05.

KIỂM ĐỊNH CÓ LIÊN QUAN ĐẾN PHÂN PHỐI CHUẨN

Minh họa, giả sử thống kê mẫu S ~ $N(\mu_S, \sigma_S^2)$.

- Nếu giả thiết đúng thì ta tin chắc 95% giá trị z của thống kê mẫu S sẽ nằm giữa -1,96 đến 1,96.

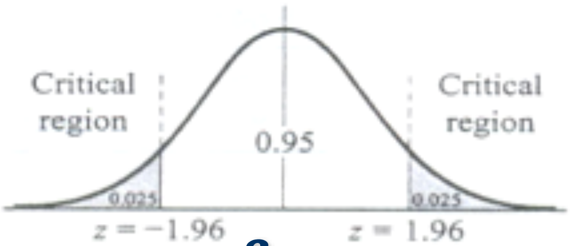


- Chúng ta có phát biểu:

(a) Bác bỏ giả thiết với mức ý nghĩa 0,05 nếu giá trị z của thống kê S nằm ngoài khoảng –1,96 đến 1,96 (ie, z >1,96 hay z <–1,96).

(b) Chấp nhận giả thiết (hay không có quyết định nào cả) nếu ngược lại.

KIỂM ĐỊNH 1 PHÍA VÀ 2 PHÍA



- *Kiểm định hai phía*: giá trị biên tương ứng của z ở hai phía của giá trị trung bình.
- *Kiểm định phía phải (trái)*: gồm các giá trị nằm ở phía phải (trái) của giá trị z.
- z_α: $1-\alpha = 2\Phi(z_\alpha)$ (2 phía); $1-2\alpha = 2\Phi(z_{2\alpha})$ (1 phía)

Mức ý nghĩa α	0,10	0,05	0,01	0,005	0,002
Giá trị z dùng để kiểm định một phía	±1,28	±1,645	±2,33	±2,58	±2,88
Giá trị z dùng để kiểm định hai phía	±1,645	±1,96	±2,58	±2,81	±3,08

GIÁ TRỊ P

Giả thiết H_0 là một khẳng định của tham số tổng thể và H_1 là một trong các khẳng định:

(i) Tham số này lớn hơn giá trị được phát biểu (kiểm định phía phải).

(ii) Tham số này nhỏ hơn giá trị được phát biểu (kiểm định phía trái).

(iii) Tham số này lớn hơn hay nhỏ hơn giá trị được phát biểu (kiểm định hai phía).

Sau khi kiểm định được thực hiện và thống kê S được tính thì giá trị P của kiểm định này là xác suất sao cho giá trị của S về hướng H_1 nếu giả thiết H_0 là đúng.

GIÁ TRỊ P

VÍ DỤ. Giả sử một tổng thể có phân phối chuẩn, với σ = 3 và H_0 khẳng định trung bình μ = 12. Chọn ngẫu nhiên từ tổng thể một mẫu có kích thước 36 và tính được trung bình mẫu $\bar{x}$ = 12,95. Thống kê được chọn để kiểm định là

$$Z = \frac{\bar{X} - 12}{\sigma / \sqrt{n}} = \frac{\bar{X} - 12}{0,5} \sim N(0,1)$$

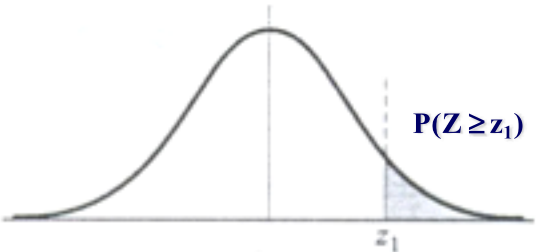
Nếu H_0 đúng thì giá trị kiểm định Z = (12,95 – 12)/0,5=1,9. Giá trị P tùy vào giả thiết H_1:

(i) Với H_1: μ > 12, giá trị P = P(Z ≥ 1,9) = 0,029.

(ii) Với H_1: μ < 12, giá trị P = P(Z ≤ 1,9) = 0,97.

(iii)Với H_1: $\mu \neq$ 12, P = P(Z ≥1,9)+P(Z ≤1,9)=0,057.

Ý NGHĨA CỦA GIÁ TRỊ P



- Nếu giá trị P nhỏ, cho bằng chứng bác bỏ H_0 để tham khảo H_1. Trường hợp (i) giá trị P= 0,029 cho thấy trung bình của tổng thể > 12.
- Nếu giá trị P lớn, bằng chứng không bác bỏ H_0 để tham khảo H_1. Trường hợp (ii), P = 0,97 nên không thể bác bỏ H_0: μ=12 để tham khảo H_1: $\mu < 12$
- Trường hợp (iii), giá trị P = 0,057 là bằng chứng bác bỏ H_0 để tham khảo H_1: $\mu \neq 12$ nhưng không có nhiều bằng chứng bác bỏ giả thiết H_0 để tham khảo giả thiết H_1: $\mu > 12$.

CÁC KIỂM ĐỊNH
ĐẶC BIỆT VỚI MẪU LỚN

1. KIỂM ĐỊNH TRUNG BÌNH TỔNG THỂ
2. KIỂM ĐỊNH TỶ LỆ TỔNG THỂ
3. KIỂM ĐỊNH HIỆU CỦA CÁC TRUNG BÌNH
4. KIỂM ĐỊNH HIỆU CỦA CÁC TỶ LỆ

KIỂM ĐỊNH TRUNG BÌNH TỔNG THỂ

Với mẫu lớn, nhiều thống kê S rất gần phân phối chuẩn, trung bình μ_S, độ lệch chuẩn σ_S.

Trường hợp thống kê là trung bình mẫu S = $\bar{X}$, ta có $\mu_S = \mu_{\bar{X}} = \mu$ và $\sigma_S = \sigma_{\bar{X}} = \sigma/\sqrt{n}$

trong đó σ là độ lệch chuẩn của tổng thể và n là kích thước mẫu. ĐLNN được chuẩn hóa là

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

Khi chưa biết độ lệch chuẩn của tổng thể, ta có thể sử dụng độ lệch chuẩn mẫu s (hoặc s^) để ước lượng cho σ.

KIỂM ĐỊNH TỶ LỆ TỔNG THỂ

Trường hợp thống kê là tỷ lệ mẫu S = f, ta có tỷ lệ thành công trên mẫu $\mu_S = \mu_f = P$, trong đó P là tỷ lệ thành công của tổng thể và n là kích thước mẫu. ĐLNN $\sigma_S = \sigma_f = \sqrt{P(1-P)/n}$ được chuẩn hóa là

$$Z = \frac{f-P}{\sqrt{P(1-P)}}\sqrt{n}$$

Trong trường hợp f = X/n, trong đó X là số lần thành công trên mẫu, biểu thức trên trở thành

$$Z = \frac{X-np}{\sqrt{npq}}$$

KIỂM ĐỊNH HIỆU CỦA CÁC TRUNG BÌNH

Gọi $\overline{X}_1$ và $\overline{X}_2$ là trung bình mẫu thu được từ các mẫu có kích thước n_1 và n_2 được chọn từ hai tổng thể có trung bình lần lượt là μ_1, μ_2 và độ lệch chuẩn là σ_1, σ_2. Xét giả thiết H_0: $\mu_1 = \mu_2$ là không có sự khác biệt nào giữa hai trung bình tổng thể.

$$\mu_{\overline{X}_1-\overline{X}_2} = 0 \qquad \sigma_{\overline{X}_1-\overline{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

ĐLNN được chuẩn hóa là

$$Z = \frac{\overline{X}_1 - \overline{X}_2 - 0}{\sigma_{\overline{X}_1-\overline{X}_2}} = \frac{\overline{X}_1 - \overline{X}_2}{\sigma_{\overline{X}_1-\overline{X}_2}}$$

Chúng ta có thể kiểm định giả thiết H_0 đối với giả thiết H_1 ở mức ý nghĩa thích hợp.

KIỂM ĐỊNH HIỆU CỦA CÁC TỶ LỆ

Gọi f_1 và f_2 là tỷ lệ mẫu thu được từ các mẫu có kích thước n_1 và n_2 được chọn từ hai tổng thể có tỷ lệ lần lượt là P_1, P_2. Xét giả thiết H_0: $P_1 = P_2$ là không có sự khác biệt giữa hai tỷ lệ tổng thể.

$$\mu_{f_1-f_2} = 0 \qquad \sigma_{f_1-f_2} = \sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

trong đó $\bar{p} = \dfrac{n_1 f_1 + n_2 f_2}{n_1 + n_2}$ dùng để ước lượng p.

ĐLNN được chuẩn hóa là

$$Z = \frac{f_1 - f_2 - 0}{\sigma_{f_1-f_2}} = \frac{f_1 - f_2}{\sigma_{f_1-f_2}}$$

Chúng ta có thể kiểm định giả thiết H_0 đối với giả thiết H_1 ở mức ý nghĩa thích hợp.

CÁC KIỂM ĐỊNH
ĐẶC BIỆT VỚI MẪU NHỎ

1. KIỂM ĐỊNH TRUNG BÌNH TỔNG THỂ
2. KIỂM ĐỊNH HIỆU CỦA CÁC TRUNG BÌNH
3. KIỂM ĐỊNH PHƯƠNG SAI TỔNG THỂ
4. KIỂM ĐỊNH TỶ LỆ CỦA CÁC PHƯƠNG SAI

KIỂM ĐỊNH TRUNG BÌNH TỔNG THỂ

Trường hợp mẫu nhỏ (n < 30), có thể kiểm định giả thiết bằng cách sử dụng phân phối Student, chi bình phương và phân phối F. Điều này dĩ nhiên vẫn đúng với mẫu có kích thước lớn.

Để kiểm định giả thiết H_0: $\mu = \mu_0$, i.e, tổng thể có phân phối chuẩn với trung bình là μ, ta sử dụng

$$T = \frac{\bar{X} - \mu}{S}\sqrt{n-1} = \frac{\bar{X} - \mu}{\hat{S}}\sqrt{n}$$

Do chưa biết độ lệch chuẩn của tổng thể nên ta sử dụng độ lệch chuẩn mẫu s (hoặc s^) để ước lượng cho σ.

KIỂM ĐỊNH HIỆU CỦA CÁC TRUNG BÌNH

Gọi $\overline{X}_1$, $\overline{X}_2$ và S_1, S_2 lần lượt là trung bình và độ lệch chuẩn mẫu thu được từ hai mẫu có kích thước n_1 và n_2 được chọn từ hai tổng thể có trung bình và độ lệch chuẩn lần lượt là μ_1, μ_2 và σ_1, σ_2. Xét giả thiết H_0: $\mu_1 = \mu_2$ và $\sigma_1 = \sigma_2$, i.e, hai mẫu NN trên có cùng tổng thể.

Ta sử dụng giá trị kiểm định

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

trong đó $$\sigma = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}}$$

T có phân phối Student với ($n_1 + n_2 - 2$) bậc tự do. σ^2 là ước lượng *phương sai chung.*

KIỂM ĐỊNH PHƯƠNG SAI TỔNG THỂ

Kiểm định giả thiết H_0: tổng thể có phân phối chuẩn với phương sai σ^2, ĐLNN

$$\chi^2 = \frac{nS^2}{\sigma^2} = \frac{(n-1)\hat{S}^2}{\sigma^2}$$

có phân phối chi bình phương n−1 bậc tự do.

- Kiểm định hai phía, chúng ta sẽ chấp nhận H_0: $\sigma^2 = \sigma_0^2$ (hay ít ra cũng không bác bỏ H_0) ở mức ý nghĩa α, nếu $\chi^2_{\alpha/2} \le \chi^2 \le \chi^2_{1-\alpha/2}$
- Để kiểm định giả thiết H_1: $\sigma^2 > \sigma_0^2$, vẫn với giả thiết H_0 bên trên, dùng kiểm định một phía, chúng ta sẽ bác bỏ H_0 ở mức ý nghĩa α (kết luận H_1 đúng) nếu $\chi^2 \ge \chi^2_{\alpha}$.

KIỂM ĐỊNH TỶ LỆ CỦA CÁC PHƯƠNG SAI

Trong bài toán, ta muốn quyết định liệu xem 2 mẫu lần lượt có kích thước là m, n và có phương sai tương ứng là s_1^2, s_2^2 được chọn từ 2 tổng thể có phân phối chuẩn thì hai tổng thể đó có cùng phương sai hay không? Chúng ta sẽ sử dụng thống kê

$$F = \frac{\hat{S}_1^2 / \sigma_1^2}{\hat{S}_2^2 / \sigma_2^2}$$

trong đó σ_1^2, σ_2^2 lần lượt là phương sai của hai tổng thể có phân phối chuẩn. Giả sử giả thiết H_0 là không có sự khác biệt nào giữa các phương sai của tổng thể, $\sigma_1^2 = \sigma_2^2$. Với giả thiết này thì biểu (16) trở thành

$$F = \frac{\hat{S}_1^2}{\hat{S}_2^2}$$

QUAN HỆ GIỮA ƯỚC LƯỢNG VÀ KIỂM ĐỊNH

Từ các kết quả trên ta thấy có mối quan hệ giữa lý thuyết ước lượng và lý thuyết kiểm định. Ví dụ, để chấp nhận giả thiết H_0 ở mức ý nghĩa α thì tương đương với kết quả $(1-\alpha)\%$ khoảng tin cậy

$$\bar{x} - z_{\alpha}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha}\frac{\sigma}{\sqrt{n}}$$

Vậy ít nhất trong trường hợp kiểm định hai phía, chúng ta có thể sử dụng khoảng tin cậy của chương 6 để kiểm định giả thiết. Một kết quả tương tự cho kiểm định một phía sẽ đòi hỏi khoảng tin cậy một phía.

ĐƯỜNG CONG ĐẶC TRƯNG – LỰC KIỂM ĐỊNH

- Hạn chế sai lầm loại I tùy thuộc vào chọn mức ý nghĩa của kiểm định.
- Tránh sai lầm loại II bằng cách không thực hiện chúng, i.e, không bao giờ chấp nhận H_0. Tuy nhiên, nhiều trường hợp trong thực tế, điều này không thể thực hiện được.
- Ta thường sử dụng *đường cong đặc trưng* hay *đường cong OC* là biểu đồ chỉ ra xác suất của sai lầm loại II dưới các giả thiết khác nhau để giảm đến mức thấp nhất của sai lầm loại II. i.e, sẽ chỉ ra *lực kiểm định.*

KIỂM ĐỊNH CHI BÌNH PHƯƠNG

- Để xác định liệu phân phối mẫu có kích thước n có cùng phân phối với tổng thể ở mức ý nghĩa α hay không, chúng ta sử dụng tiến trình kiểm định giả thiết được gọi là *kiểm định chi bình phương*

$$\chi^2 = \sum_{j=1}^{k} \frac{\left(X_j - np_j\right)^2}{np_j} \quad \text{hoặc} \quad \chi^2 = \sum_{j=1}^{k} \frac{X_j^2}{np_j} - n$$

- Nếu $\chi^2 > \chi^2_\alpha$: bác bỏ H_0, i.e, phân phối mẫu không cùng phân phối với tổng thể.
- Nếu $\chi^2 \le \chi^2_\alpha$: chấp nhận giả thiết H_0 (ít nhất cũng không bác bỏ), i.e, phân phối mẫu có cùng phân phối với tổng thể.

THE END