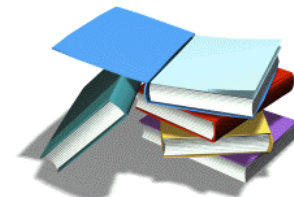




TRƯỜNG ĐẠI HỌC TÀI CHÍNH-MARKETING
KHOA CƠ BẢN
BỘ MÔN TOÁN – THỐNG KÊ

BÀI GIẢNG
LÝ THUYẾT XÁC SUẤT VÀ THỐNG KÊ TOÁN

Giảng viên
ThS. Lê Trường Giang





LÝ THUYẾT XÁC SUẤT & THỐNG KÊ TOÁN

Chương 5

KIỂM ĐỊNH GIẢ THIẾT THỐNG KÊ

Thuật toán kiểm định và bài tập tổng hợp



1. Bài toán kiểm định trung bình.

Giả sử biến ngẫu nhiên X có tham số trung bình $E(X) = \mu$ chưa biết. Ta đặt giả thiết $H_0 : \mu = \mu_0$ và một trong các đối thiết

$$\left[\begin{array}{l} H_1 : \mu > \mu_0 \\ H_1 : \mu < \mu_0 \\ H_1 : \mu \neq \mu_0 \end{array} \right. \text{ . Với một độ tin cậy cho trước và một mẫu cụ thể chọn}$$

được, ta cần khẳng định giả thiết đúng hay đối thiết đúng.

Các bước thực hành: kiểm định trung bình

+ **Bước 1:** Chọn giả thiết $H_0 : \mu = \mu_0; H_1 : \begin{cases} \mu > \mu_0 \\ \mu < \mu_0 \\ \mu \neq \mu_0 \end{cases}$

+ **Bước 2:** Tìm miền bác bỏ

i) Miền bác bỏ dạng Z

Nếu $\mu > \mu_0$ thì $W_\alpha = (z_{0.5-\alpha}; +\infty)$;

Nếu $\mu < \mu_0$ thì $W_\alpha = (-\infty; -z_{0.5-\alpha})$;

Nếu $\mu \neq \mu_0$ thì $W_\alpha = \left(-\infty, -z_{\frac{1-\alpha}{2}}\right) \cup \left(z_{\frac{1-\alpha}{2}}, +\infty\right)$.

ii) Miền bác bỏ dạng T

Nếu $\mu > \mu_0$ thì $W_\alpha = (t_\alpha(n-1); +\infty)$;

Nếu $\mu < \mu_0$ thì $W_\alpha = (-\infty; -t_\alpha(n-1))$;

Nếu $\mu \neq \mu_0$ thì $W_\alpha = \left(-\infty, -t_{\frac{\alpha}{2}}(n-1)\right) \cup \left(t_{\frac{\alpha}{2}}(n-1), +\infty\right)$.

Các bước thực hành: kiểm định trung bình

+ **Bước 3:** Tính giá trị qua sát

$$z_{qs} = \frac{(\bar{x} - \mu_0)\sqrt{n}}{\sigma}$$

Lưu ý nếu không có σ ta thay bằng s .

+ **Bước 4:**

i) Nếu $z_{qs} \in W_\alpha$ thì ta bác bỏ H_0 .

ii) Nếu $z_{qs} \notin W_\alpha$ thì ta chưa có cơ sở bác bỏ H_0 .

2. Bài toán so sánh hai trung bình

Giả sử X_1 và X_2 là hai biến ngẫu nhiên độc có hai trung bình

$E(X_1) = \mu_1$ và $E(X_2) = \mu_2$ chưa biết. Ta đặt giả thiết $H_0 : \mu_1 = \mu_2$ và

một trong các đối thiết $\left[\begin{array}{l} H_1 : \mu_1 > \mu_2 \\ H_2 : \mu_1 < \mu_2 \\ H_3 : \mu_1 \neq \mu_2 \end{array} \right.$. Với độ tin cậy cho trước và hai

mẫu độc lập của X_1 và X_2 ; ta cần kiểm định giả thiết đúng hay đối thiết đúng.

Các bước thực hành: so sánh hai trung bình

+ **Bước 1:** Chọn giả thiết $H_0 : \mu_1 = \mu_2; H_1 : \begin{cases} \mu_1 > \mu_2 \\ \mu_1 < \mu_2 \\ \mu_1 \neq \mu_2 \end{cases}$

+ **Bước 2:** Tìm miền bác bỏ

i) Miền bác bỏ dạng Z

Nếu $\mu_1 > \mu_2$ thì $W_\alpha = (z_{0.5-\alpha}; +\infty)$;

Nếu $\mu_1 < \mu_2$ thì $W_\alpha = (-\infty; -z_{0.5-\alpha})$;

Nếu $\mu_1 \neq \mu_2$ thì $W_\alpha = \left(-\infty, -z_{\frac{1-\alpha}{2}}\right) \cup \left(z_{\frac{1-\alpha}{2}}, +\infty\right)$.

ii) Miền bác bỏ dạng T

Nếu $\mu_1 > \mu_2$ thì $W_\alpha = (t_\alpha (n_1 + n_2 - 2); +\infty)$;

Nếu $\mu_1 < \mu_2$ thì $W_\alpha = (-\infty; -t_\alpha (n_1 + n_2 - 2))$;

Nếu $\mu_1 \neq \mu_2$ thì

$$W_\alpha = \left(-\infty, -t_{\frac{\alpha}{2}} (n_1 + n_2 - 2)\right) \cup \left(t_{\frac{\alpha}{2}} (n_1 + n_2 - 2), +\infty\right).$$

Các bước thực hành: so sánh hai trung bình

+ **Bước 3:** Tính giá trị qua sát

i) Dạng Z:
$$z_{qs} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Lưu ý nếu không có σ_1^2, σ_2^2 ta thay bằng s_1^2, s_2^2 .

ii) Dạng T:
$$z_{qs} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

Trong đó
$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

+ **Bước 4:**

i) Nếu $z_{qs} \in W_\alpha$ thì ta bác bỏ H_0 .

ii) Nếu $z_{qs} \notin W_\alpha$ thì ta chưa có cơ sở bác bỏ H_0 .

3. Bài toán kiểm định tỷ lệ.

Trong tổng thể X ta quan tâm những phần tử có tính chất A với tỷ lệ p chưa biết. Giả sử chúng ta có một giả thiết ban đầu về tỷ lệ phần tử có tính chất A này là $H_0 : p = p_0$ và một trong các đối thiết

$\left[\begin{array}{l} H_1 : p > p_0 \\ H_1 : p < p_0 \\ H_1 : p \neq p_0 \end{array} \right.$. Chọn một mẫu có kích thước n , bài toán kiểm định tỷ lệ

là việc khẳng định giả thiết H_0 đúng hay đối thiết H_1 đúng với độ tin cậy cho trước.

Kiểm định tỷ lệ (điều kiện $\begin{cases} np_0 > 5 \\ n(1-p_0) > 5 \end{cases}$)

+ **Bước 1:** Chọn giả thiết $H_0 : p = p_0; H_1 : \begin{cases} p > p_0 \\ p < p_0 \\ p \neq p_0 \end{cases}$

+ **Bước 2:** Tìm miền bác bỏ

Nếu $p > p_0$ thì $W_\alpha = (z_{0.5-\alpha}; +\infty)$;

Nếu $p < p_0$ thì $W_\alpha = (-\infty; -z_{0.5-\alpha})$;

Nếu $p \neq p_0$ thì $W_\alpha = \left(-\infty, -z_{\frac{1-\alpha}{2}}\right) \cup \left(z_{\frac{1-\alpha}{2}}, +\infty\right)$.

+ **Bước 3:** Tính giá trị qua sát $z_{qs} = \frac{(f - p_0)\sqrt{n}}{\sqrt{p_0(1-p_0)}}$.

+ **Bước 4:** Nếu $z_{qs} \notin W_\alpha$ thì ta bác bỏ H_0 .

4. Bài toán so sánh hai tỷ lệ

Xét hai tổng thể X_1 và X_2 có những phần tử có tính chất A. Gọi p_1, p_2 lần lượt là tỷ lệ phần tử có tính chất A của tổng thể trên. Ta đặt

giả thiết $H_0 : p_1 = p_2$ và một trong các đôi thiết $\left[\begin{array}{l} H_1 : p_1 > p_2 \\ H_2 : p_1 < p_2 \\ H_3 : p_1 \neq p_2 \end{array} \right.$. Với độ

tin cậy cho trước và hai mẫu độc lập của X_1 và X_2 ; ta cần kiểm định giả thiết đúng hay đôi thiết đúng.

So sánh hai tỷ lệ

+ **Bước 1:** Chọn giả thiết $H_0 : p_1 = p_2; H_1 : \begin{cases} p_1 > p_2 \\ p_1 < p_2 \\ p_1 \neq p_2 \end{cases}$

+ **Bước 2:** Tìm miền bác bỏ

Nếu $p_1 > p_2$ thì $W_\alpha = (z_{0.5-\alpha}; +\infty)$;

Nếu $p_1 < p_2$ thì $W_\alpha = (-\infty; -z_{0.5-\alpha})$;

Nếu $p_1 \neq p_2$ thì $W_\alpha = \left(-\infty, -z_{\frac{1-\alpha}{2}}\right) \cup \left(z_{\frac{1-\alpha}{2}}, +\infty\right)$.

+ **Bước 3:** Tính giá trị qua sát $z_{qs} = \frac{f_1 - f_2}{\sqrt{f(1-f)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$.

Trong đó $f = \frac{n_1 f_1 + n_2 f_2}{n_1 + n_2}$.

+ **Bước 4:** Nếu $z_{qs} \notin W_\alpha$ thì ta bác bỏ H_0 .

BÀI TẬP NHÓM

Bài 1. Điều tra thu nhập của 100 hộ gia đình ở tỉnh A thấy có 13 hộ thuộc diện nghèo.

- Ước lượng số hộ nghèo ở tỉnh A với độ tin cậy 95%, biết rằng tỉnh A có 15.000 hộ.
- Tỷ lệ hộ nghèo của tỉnh B là 10%, với mức ý nghĩa 5% có thể cho rằng tỷ lệ hộ nghèo của tỉnh A cao hơn tỉnh B hay không?

Bài 2. Để kiểm tra chất lượng của một lô lớn các màn hình máy tính xuất khẩu người ta lấy ngẫu nhiên 100 màn hình để kiểm tra và thấy 4 màn hình có khuyết tật.

- Với độ tin cậy 95% hãy ước lượng số màn hình có khuyết tật tối đa nếu lô hàng đó có 10.000 màn hình.
- Nhà nhập khẩu chỉ chấp nhận lô màn hình đó nếu tỷ lệ các màn hình có khuyết tật tối đa là 5%. Hỏi lô hàng đó có thể chấp nhận được không?



Bài 3. Khảo sát về thu nhập X (triệu đồng/tháng) của một số công nhân tại một công ty may mặc người ta có bảng số liệu sau:

Thu nhập	2-3	3-4	4-5	5-6	6-7	7-9
Số người	5	9	30	25	10	6

- Tính trung bình mẫu, độ lệch chuẩn mẫu.
- Hãy ước lượng thu nhập trung bình của một người trong một tháng với độ tin cậy 95%?
- Giả sử công ty báo cáo rằng "mức thu nhập trung bình của một người là 5000000 đồng/tháng", với mức ý nghĩa 5% có thể chấp nhận được báo cáo trên hay không?
- Những người có thu nhập không quá 4000000 đồng/tháng là những người có mức thu nhập thấp. Hãy ước lượng những người có mức thu nhập thấp với độ tin cậy 96%?
- Giả sử công ty báo cáo rằng "Tỷ lệ những người có mức thu nhập thấp của công ty là 10%", với mức ý nghĩa 5%, báo cáo này có chấp nhận được không?

Bài 4. Khảo sát về thời gian tự học trong một tuần của một số sinh viên ở một trường đại học trong thời gian gần đây, người ta thu được bảng số liệu sau:

Thời gian tự học (giờ/tuần)	2-4	4-6	6-8	8-10	10-12	12-14	14-16
Số sinh viên	18	25	30	22	15	12	8

- a) Ước lượng giờ tự học trung bình của sinh viên trường này với độ tin cậy 95%.
- b) Trước đây giờ tự học của sinh viên trường này là 10 giờ/tuần. Hãy cho nhận xét về tình hình tự học của sinh viên hệ chính quy trường này trong thời gian gần đây với mức ý nghĩa 5%?
- c) Những sinh viên có giờ tự học từ 10 giờ/tuần trở lên là những sinh viên chăm học. Hãy ước lượng số sinh viên chăm học của trường này với độ tin cậy 98% (trường có 10000 sv)?

Bài 5. Số liệu thống kê về doanh số (DS) bán (triệu đồng/ngày) của một siêu thị như sau:

DS	20-40	40-50	50-60	60-70	70-80	80-90	90-100	100-110	110-130
Số ngày	5	10	20	25	25	15	10	8	3

- Những ngày có doanh số bán hàng trên 90 triệu đồng là những ngày bán đắt hàng. Hãy ước lượng tỷ lệ những ngày bán đắt hàng ở siêu thị này với độ tin cậy 95%.
- Ước lượng doanh số bán trung bình của một ngày ở siêu thị với độ tin cậy 90%, giả sử doanh số bán hàng của những ngày bán là đại lượng ngẫu nhiên có phân phối chuẩn.
- Nếu muốn sai số ước lượng trung bình của một ngày bán hàng ở siêu thị không vượt quá 3 triệu đồng/ngày, ở độ tin cậy 99% thì cần quan sát thêm ít nhất bao nhiêu ngày nữa.
- Trước đây doanh số bán hàng trung bình là 65 triệu đồng/ngày. Số liệu ở trên được thu thập sau khi siêu thị áp dụng phương pháp bán hàng mới. Hãy cho nhận xét về phương pháp bán hàng này với mức ý nghĩa 5%.

Bài 6. Khảo sát chiều cao (m) của 100 sinh viên ở một Trường Đại học (chọn mẫu ngẫu nhiên) ta được bảng số liệu sau

Chiều cao	1,54- 1,58	1,58- 1,62	1,62- 1,66	1,66- 1,70	1,70- 1,74	1,74- 1,78	1,78- 1,82
Số sv	25	15	30	14	10	4	2

- Hãy ước lượng chiều cao trung bình của sinh viên với độ tin cậy 95%.
- Với độ tin cậy 90%, hãy ước lượng tỷ lệ sinh viên có chiều cao từ 1,7m trở đi.
- Với số liệu thống kê trên, nếu muốn ước lượng chiều cao trung bình của sinh viên đạt độ tin cậy 99% và độ chính xác 0,01m thì cần điều tra thêm bao nhiêu sinh viên nữa?
- Một người khẳng định rằng chiều cao trung bình của sinh viên trường này là 1,67m. Hãy kết luận về lời khẳng định đó với mức ý nghĩa 5%.

Bài 7. Khảo sát năng suất (tạ/ha) của một giống lúa mới khi thu hoạch ở 41 điểm tại vùng A, ta thu được kết quả sau

Năng suất	37	38	39	40	41
Số điểm	5	8	10	11	7

- a) Ước lượng năng suất trung bình tối thiểu của giống lúa này tại vùng A với độ tin cậy 95%.
- b) Giống lúa mới được coi là đạt yêu cầu nếu đạt năng suất 39,5 tạ/ha. Với mức ý nghĩa 5% có thể cho rằng giống lúa trên đạt yêu cầu hay không?

Bài 8. Theo dõi doanh thu (triệu đồng) của một đại lý bán xăng dầu qua một số ngày thu được kết quả:

Doanh thu	11	12	13	14	15
Số ngày	3	7	10	7	14

- a) Ước lượng doanh thu trung bình tối thiểu của đại lý trên với độ tin cậy 95%.
- b) Năm trước theo dõi doanh thu qua 36 ngày tính được doanh thu trung bình hằng ngày là 12,5 triệu đồng và độ lệch chuẩn là 500 ngàn đồng, với mức ý nghĩa 5% có thể cho rằng doanh thu hằng ngày đã thay đổi? (biết rằng doanh thu là biến ngẫu nhiên có phân phối chuẩn)



XIN CHÂN THÀNH CẢM ƠN!

