

NỘI DUNG THI CUỐI KỲ

1. Phân Phối Xác Suất

- Phân phối Chuẩn

2. Ước lượng (cho trung bình, tỷ lệ)

- Ước lượng điểm
- Ước lượng khoảng

3. Kiểm định (cho trung bình, tỷ lệ)

- Kiểm định một phía
- Kiểm định hai phía

4. Hồi quy tuyến tính 2 biến

Ghi chú:

- sinh viên được sử dụng tài liệu giấy và máy tính khi thi cuối kỳ
- sinh viên nên chuẩn bị bảng tra mà mình quen thuộc

TÓM TẮT BÀI TOÁN ƯỚC LƯỢNG THAM SỐ

I. ƯỚC LƯỢNG ĐIỂM

- Ước lượng điểm cho kỳ vọng của quần thể (μ) là giá trị trung bình của mẫu (\bar{x})
- Ước lượng điểm cho độ lệch chuẩn của quần thể (σ) là độ lệch chuẩn của mẫu (s)
- Ước lượng điểm cho tỷ lệ của quần thể (p) là tỷ lệ mẫu (f)

II. ƯỚC LƯỢNG KHOẢNG

1. Ước lượng khoảng cho kỳ vọng

- Bài toán: cần tìm một khoảng ước lượng cho kỳ vọng của quần thể (μ) với độ tin cậy là $1-\alpha$. Để ước lượng khoảng, ta dựa vào tập mẫu quan sát được tìm khoảng ước lượng sao cho:

$$P(\bar{x} - \varepsilon; \bar{x} + \varepsilon) = 1 - \alpha$$

(ε : còn được gọi là sai số ước lượng)

- Công thức tính sai số ước lượng:

- o **Trường hợp 1:** $n > 30$, biết σ :

$$\varepsilon = z_{\alpha/2} * \frac{\sigma}{\sqrt{n}}$$

- o **Trường hợp 2:** $n > 30$, không biết σ :

$$\varepsilon = z_{\alpha/2} * \frac{s}{\sqrt{n}}$$

- o **Trường hợp 3:** $n \leq 30$, X có phân phối chuẩn và biết σ :

$$\varepsilon = z_{\alpha/2} * \frac{\sigma}{\sqrt{n}}$$

- o **Trường hợp 4:** $n \leq 30$, X có phân phối chuẩn và không biết σ :

$$\varepsilon = t_{\alpha/2}^{n-1} * \frac{s}{\sqrt{n}}$$

- Với:

- o $z_{\alpha/2}$ được tính dựa vào phân phối Z sao cho:

$$\varphi(z_{\alpha/2}) = (1 - \alpha)/2$$

- $t_{\alpha/2}^{n-1}$ được tính theo phân phối T
- Khoảng tin cậy cho kỳ vọng là: $(\bar{x} - \varepsilon; \bar{x} + \varepsilon)$

2. Ước lượng khoảng cho tỷ lệ

- Bài toán: cần tìm một khoảng ước lượng cho tỷ lệ của quần thể (p) với độ tin cậy là $1 - \alpha$. Để ước lượng khoảng, ta dựa vào tập mẫu quan sát được tìm khoảng ước lượng $(f - \varepsilon, f + \varepsilon)$ sao cho:

$$P(f - \varepsilon; f + \varepsilon) = 1 - \alpha$$

(ε : còn được gọi là sai số ước lượng)

- Công thức tính sai số ước lượng:

$$\varepsilon = z_{\alpha/2} * \sqrt{\frac{f(1-f)}{n}}$$

- Với $z_{\alpha/2}$ được tính dựa vào phân phối Z sao cho:

$$\varphi(z_{\alpha/2}) = (1 - \alpha)/2$$

- Khoảng tin cậy cho tỷ lệ là: $(f - \varepsilon, f + \varepsilon)$

III. VÍ DỤ

(Tham khảo đáp án bài TK03)

IV. BÀI TẬP RÈN LUYỆN

(Xem trong file bài tập)

TÓM TẮT KIỂM ĐỊNH GIẢ THUYẾT THỐNG KÊ

I. CÁC BƯỚC THỰC HIỆN BÀI TOÁN KIỂM ĐỊNH

- Bước 1: Phát biểu cặp giả thuyết/ đối thuyết

Khi phát biểu cần xem xét kỹ bài toán là kiểm định một phía hay hai phía để phát biểu cho chính xác.

Lưu ý: giả thuyết không (H_0) luôn đi kèm với dấu bằng

- Bước 2: Xác định mức ý nghĩa α của kiểm định

- Bước 3: Tính giá trị thống kê kiểm định (z_0 hoặc t_0). Việc tính toán trị thống kê (z_0 hoặc t_0) tùy thuộc vào phân phối tương ứng (phân phối Z hoặc phân phối T)

- Bước 4: Xác định miền bác bỏ giả thuyết

Lưu ý: cần để ý bài toán là kiểm định một phía hay hai phía mà xác định cho chính xác miền bác bỏ

- Bước 5: Dựa vào trị thống kê kiểm định và miền bác bỏ, đưa ra kết luận về giả thuyết ban đầu (bác bỏ H_0 hoặc chưa đủ cơ sở để bác bỏ H_0)

II. KIỂM ĐỊNH GIẢ THUYẾT CHO KỲ VỌNG

1. Trị thống kê

- **Trường hợp 1:** $n > 30$, biết σ : (sử dụng phân phối Z)

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

- **Trường hợp 2:** $n > 30$, không biết σ : (sử dụng phân phối Z)

$$z_0 = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

- **Trường hợp 3:** $n \leq 30$, X có phân phối chuẩn và biết σ : (sử dụng phân phối Z)

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

- **Trường hợp 4:** $n \leq 30$, X có phân phối chuẩn và không biết σ : (sử dụng phân phối T)

$$t_0 = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

2. Miền bác bỏ

- Trường hợp z-test (trị thống kê là z_0):

Giả thuyết	Miền bác bỏ
$H_0: \mu = \mu_0; H_1: \mu \neq \mu_0$	$\{z: z > z_{\alpha/2}\}$
$H_0: \mu = \mu_0; H_1: \mu < \mu_0$	$\{z: z < -z_{\alpha}\}$
$H_0: \mu = \mu_0; H_1: \mu > \mu_0$	$\{z: z > z_{\alpha}\}$

- Trường hợp t-test (trị thống kê là t_0):

Giả thuyết	Miền bác bỏ
$H_0: \mu = \mu_0; H_1: \mu \neq \mu_0$	$\{t: t > t_{\alpha/2}^{n-1}\}$
$H_0: \mu = \mu_0; H_1: \mu < \mu_0$	$\{t: t < -t_{\alpha}^{n-1}\}$
$H_0: \mu = \mu_0; H_1: \mu > \mu_0$	$\{t: t > t_{\alpha}^{n-1}\}$

III. TRỊ THỐNG KÊ KHI KIỂM ĐỊNH GIẢ THUYẾT CHO TỶ LỆ

1. Trị thống kê

- Điều kiện kiểm định: $n.p_0 > 5$ và $n(1 - p_0) > 5$
- Trị thống kê kiểm định:

$$z_o = \frac{f - p_o}{\sqrt{\frac{p_o(1 - p_o)}{n}}}$$

2. Miền bác bỏ

Giả thuyết	Miền bác bỏ
$H_0: p = p_0; H_1: p \neq p_0$	$\{z: z > z_{\alpha/2}\}$
$H_0: p = p_0; H_1: p < p_0$	$\{z: z < -z_{\alpha}\}$
$H_0: p = p_0; H_1: p > p_0$	$\{z: z > z_{\alpha}\}$

IV. VÍ DỤ

1. **Ví dụ 1:** Trọng lượng của một gói chè do một máy tự động đóng gói là 500gam/gói. Người ta lấy ngẫu nhiên 30 gói chè và cân được trọng lượng trung bình là 495gam/gói, độ lệch chuẩn là 10gam. Một ý kiến cho rằng máy làm việc không bình thường nên làm cho trọng lượng chè trung bình khi lấy ra kiểm thử giảm sút. Với mức ý nghĩa 5%, hãy cho biết ý kiến này có đáng tin hay không?

GIAI:

- Ta có các thông tin sau từ đề bài:

$$\mu_0 = 500; n = 30; \bar{x} = 495; s = 10, \alpha = 0.05$$

- B1: Vì cần kiểm định xem trọng lượng chè trung bình do máy đóng gói có giảm sút hay không (nghĩa là có nhỏ hơn so với mức làm việc bình thường của máy là 500gam/gói). Ta đặt giả thiết như sau:

$$\begin{cases} H_0: \mu = 500 \\ H_1: \mu < 500 \end{cases}$$

- B2: Mức ý nghĩa của kiểm định: $\alpha = 0.05$.

$$\text{Suy ra: } t_{\alpha}^{n-1} = t_{0.05}^{29} = 1.699 \text{ (tra bảng phân phối T)}$$

- B3: Trị thống kê:

$$t_o = \frac{\bar{x} - \mu_o}{s/\sqrt{n}} = \frac{495 - 500}{10/\sqrt{30}} = -2.739$$

- B4: Miền bác bỏ:

$$\{t: t < -t_{\alpha}^{n-1}\}$$

Hay:

$$\{t: t < -1.699\}$$

- B5: Vì $t_o < -1.699$. Vậy ta bác bỏ giả thuyết H_0 , chấp nhận giả thuyết H_1 .

Kết luận: vậy ý kiến: “máy làm việc không bình thường nên trọng lượng của gói chè giảm sút” là đáng tin.

2. **Ví dụ 2:** Điều tra doanh thu (triệu/ tháng) bán hàng một sản phẩm X của các hộ kinh doanh hàng năm, người ta thu được bảng số liệu như sau:

X (triệu/tháng)	11	11.5	12	12.5	13	13.5
n_i (số hộ)	10	15	20	30	15	10

Những hộ có doanh thu cao là những hộ có doanh thu trên 12.5 triệu đồng/tháng. Theo một báo cáo, tỉ lệ hộ có doanh thu cao là 35%. Với mức ý nghĩa 5%, hãy cho biết số liệu trong bảng báo cáo có đáng tin hay không?

GIẢI:

- Ta có các thông tin sau từ đề bài:

$$p_0 = 0.35; n = 100; f = \frac{15 + 10}{100} = 0.25; \alpha = 0.05$$

- B1: Ta đặt giả thiết như sau:

$$\begin{cases} H_0: p = 0.35 \\ H_1: p \neq 0.35 \end{cases}$$

- B2: Mức ý nghĩa của kiểm định: $\alpha = 0.05$.

$$\text{Suy ra: } Z_{\alpha/2} = Z_{0.475} = 1.96 \text{ (tra bảng phân phối Z)}$$

- B3:

- Kiểm tra điều kiện kiểm định: $n.p_0 > 5$ và $n.(1-p_0) > 5 \rightarrow$ THỎA
- Trị thống kê:

$$z_o = \frac{f - p_o}{\sqrt{\frac{p_o(1 - p_o)}{n}}} = \frac{0.25 - 0.35}{\sqrt{\frac{0.35(1 - 0.35)}{100}}} = -2.097$$

- B4: Miền bác bỏ:

$$\{z: |z| > z_{\alpha/2}\}$$

Hay:

$$\{z: |z| > 1.96\}$$

- B5: Vì $|z_o| > 1.96$. Vậy ta bác bỏ giả thuyết H_o , chấp nhận giả thuyết H_1 .

Kết luận: Vậy với mức ý nghĩa 5%, báo cáo trên là không đáng tin.

V. BÀI TẬP RÈN LUYỆN

Tham khảo file [bai_tap_phanphoi_uocluong_kiemdinh.pdf](#) phần **7.1, 7.2, 8.1, 8.3**

