



HCMUS
Viet Nam National University
Ho Chi Minh City
University of Science



Khoa Toán - Tin học
Fac. of Math. & Computer Science

BỔ SUNG KIẾN THỨC ƯỚC LƯỢNG - KIỂM ĐỊNH

Lưu Giang Nam

Bộ môn Ứng dụng Tin học
Khoa Toán - Tin học
Trường Đại học KHTN, ĐHQG TPHCM

10/2025

Mục lục

1 Ước lượng

- Giới thiệu
- Ước lượng trung bình quần thể
- Ước lượng tỉ lệ quần thể

2 Kiểm định

- Các khái niệm cơ bản
- Kiểm định về giá trị trung bình
- Kiểm định về giá trị tỉ lệ

Phần 1

Ước lượng

Phần 1

Ước lượng

Mục 1: GIỚI THIỆU

Giới thiệu

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

- Phần này trình bày hai vấn đề chính của thống kê suy diễn:
 - 1 Sử dụng dữ liệu mẫu để ước lượng các giá trị của tham số quần thể.
 - 2 Kiểm định giả thuyết hoặc đưa ra các phát biểu về các tham số quần thể.
- Chúng ta giới thiệu các phương pháp để ước lượng các giá trị tham số quan trọng trên quần thể như:
 - **Trung bình**
 - **Tỉ lệ**
- Chúng ta cũng trình bày các phương pháp xác định kích thước mẫu cần thiết để ước tính các thông số đó.



Ước lượng tham số

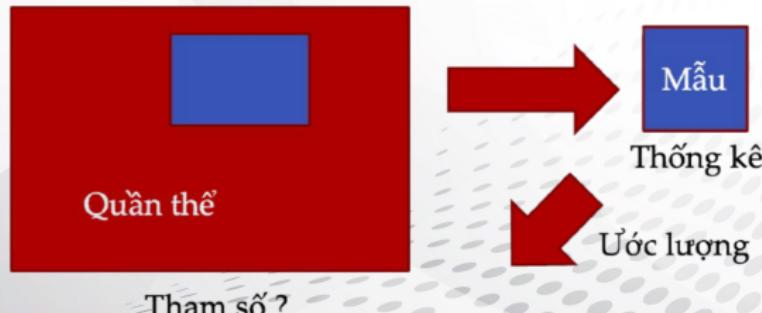
■ **Ước lượng tham số:** Sử dụng thống kê mẫu để ước lượng cho tham số quần thể.

■ **Ước lượng điểm (Point Estimation):**

■ Xác định một giá trị số là giá trị ước lượng cho tham số quần thể.

■ **Ước lượng khoảng (Interval Estimation):**

■ Xác định một khoảng giá trị có nhiều khả năng chứa giá trị tham số quần thể.



Ước lượng điểm và ước lượng khoảng

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Ước lượng điểm (Point Estimation):

- Là quá trình xác định một giá trị duy nhất từ dữ liệu mẫu để làm đại diện cho tham số của quần thể.
- Ví dụ: Sử dụng trung bình mẫu \bar{X} để ước lượng kỳ vọng μ của quần thể.

Đặc điểm:

- Đơn giản và dễ sử dụng.
- Chỉ cung cấp một giá trị duy nhất, không phản ánh được độ bất định.

Công thức:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$



Ước lượng điểm và ước lượng khoảng

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Ước lượng khoảng (Interval Estimation):

- Là quá trình xác định một khoảng giá trị có khả năng cao chứa tham số của quần thể.
- Khoảng tin cậy thường được biểu diễn dưới dạng:

$$CI = (\text{Ước lượng điểm} \pm \text{Sai số chuẩn}) \quad (2)$$

Đặc điểm:

- Phản ánh được độ bất định của ước lượng.
- Độ chính xác phụ thuộc vào kích thước mẫu và mức ý nghĩa.

Phần 1

Ước lượng

Mục 2: ƯỚC LƯỢNG TRUNG BÌNH QUẦN
THỂ

Khái niệm chính về ước lượng tham số

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

- 1 Trung bình mẫu \bar{x} là ước lượng điểm tốt nhất cho trung bình quần thể μ .
- 2 Sử dụng dữ liệu mẫu để xây dựng một khoảng giá trị, gọi là khoảng tin cậy (confidence interval), có khả năng cao chứa trung bình quần thể μ .
- 3 Để đạt được một khoảng tin cậy với độ chính xác mong muốn, kích thước mẫu n phải được tính toán.

Ước lượng trung bình quần thể

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tần số quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

- Trung bình mẫu \bar{X} là ước lượng điểm tốt nhất cho trung bình quần thể μ .

Khoảng ước lượng cho trung bình μ :

- Công thức khoảng ước lượng:

$$\bar{X} \pm E \quad \text{hoặc} \quad \bar{X} - E < \mu < \bar{X} + E \quad (3)$$

- Trong đó:

- \bar{X} : Trung bình mẫu, tính từ một tập mẫu có kích thước n .
- E : Biên độ lỗi (margin of error).

Vấn đề: Biên độ lỗi E được tính như thế nào?

Ước lượng trung bình quần thể

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

- E được tính dựa vào định lý giới hạn trung tâm.
- Định lý giới hạn trung tâm: khi ta lấy mẫu ngẫu nhiên, kích thước tập mẫu càng lớn thì phân phối xác suất của đặc trưng trung bình của tập mẫu càng gần với phân phối chuẩn.
- Biên độ lỗi E được xác định bằng công thức:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \quad (4)$$

- Trong đó:
 - $z_{\alpha/2}$: Giá trị critical value tương ứng với mức ý nghĩa α .
 - σ : Độ lệch chuẩn của quần thể (nếu biết).
 - n : Kích thước mẫu.

Độ tin cậy (Confidence Level)

- Độ tin cậy (confidence level) là xác suất $1 - \alpha$.
- Biểu thị khả năng rằng khoảng tin cậy thực sự chứa tham số quần thể, giả định rằng quá trình ước lượng được lặp lại một số lượng lớn lần.
- Độ tin cậy còn được gọi là:
 - Bậc tin cậy (degree of confidence).
 - Hệ số tin cậy (confidence coefficient).

Ý nghĩa:

- Độ tin cậy càng cao, khoảng tin cậy càng rộng.
- Thể hiện mức độ chắc chắn về kết quả ước lượng tham số quần thể. Chúng ta phải cẩn thận để diễn giải các khoảng tin cậy một cách chính xác.
 - Ví dụ: khoảng tin cậy $0.828 < p < 0.872$.
 - "Chúng ta tin tưởng $(1 - \alpha) \cdot 100\%$ rằng khoảng giá trị từ 0.828 đến 0.872 chứa giá trị thực của tỷ lệ quần thể p ".

Phân phối Student - t-distribution

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ



- Nếu phân phối của quần thể là phân phối chuẩn, thì phân phối của:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \quad (5)$$

là một phân phối t cho tất cả các mẫu có kích thước n .

- Phân phối này thường được gọi là **phân phối t** và được sử dụng để tìm các giá trị tới hạn (critical values), được biểu thị bằng $t_{\alpha/2}$.
- Phân phối t phù hợp cho các trường hợp:
 - Kích thước mẫu nhỏ ($n \leq 30$).
 - Độ lệch chuẩn của quần thể (σ) chưa biết.
 - Quần thể có phân phối chuẩn.

Bậc Tự Do (Degrees of Freedom)

Ước lượng

Giới thiệu

Ước lượng trung bình quản thể

Ước lượng tỉ lệ quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

- **Số bậc tự do** của tập hợp dữ liệu mẫu là số các giá trị mẫu có thể khác nhau sau khi có ràng buộc cụ thể trên tất cả các giá trị dữ liệu.
- Bậc tự do thường được viết tắt là **df**.
- Công thức tính bậc tự do:

$$df = n - 1 \quad (6)$$

- Đối với các phương pháp trong phần này, bậc tự do sẽ được sử dụng để xác định phân phối t và các giá trị tới hạn tương ứng.

Tìm $z_{\alpha/2}$ cho Mức độ Tin cậy 95%

Để tìm giá trị $z_{\alpha/2}$ cho mức độ tin cậy 95%, chúng ta thực hiện các bước sau:

- Mức độ tin cậy 95% có nghĩa là 5% dữ liệu nằm ngoài khoảng tin cậy (2.5% ở mỗi đuôi phân phối chuẩn).
- Tính giá trị α :

$$\alpha = 1 - 0.95 = 0.05$$

- Tính $\alpha/2$:

$$\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$$

- Tìm giá trị $z_{\alpha/2}$ từ bảng phân phối **Student distribution**:

$$z_{\alpha/2} = 1.96, \text{ với } Pr(T > t) = 0.025 \text{ và } df = +\infty$$

Vậy, với mức độ tin cậy 95%, giá trị $z_{\alpha/2}$ là 1.96.

Giải thích chính xác về khoảng tin cậy

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thể

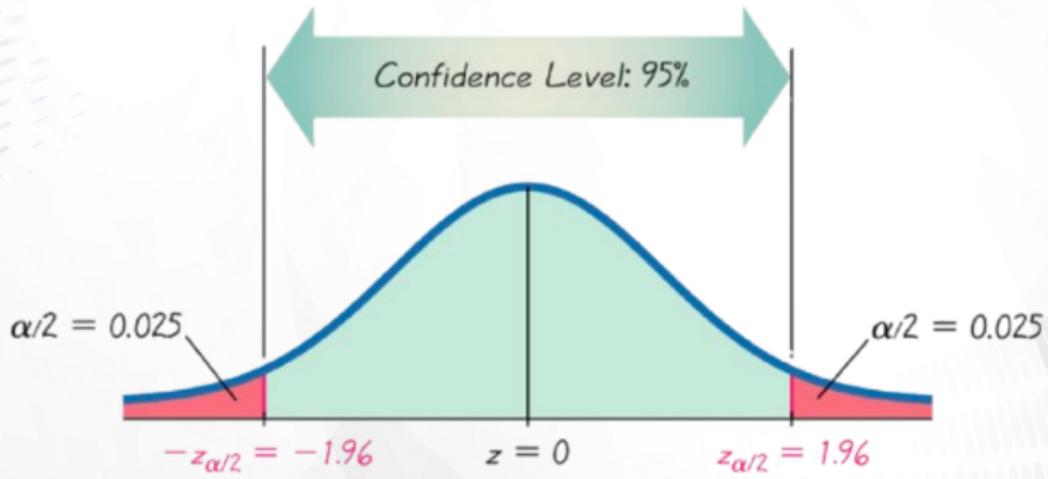
Ước lượng tỉ lệ quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tỉ lệ



Critical Values

Ví dụ 1

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Giả sử quần thể gồm các bao đựng gạo nhỏ, người ta lấy mẫu gồm 35 bao và cân trọng lượng của chúng. Trung bình khối lượng mẫu là 362.3 gram và độ lệch chuẩn quần thể là 5 gram. Cần xây dựng khoảng ước lượng cho trung bình khối lượng của toàn bộ các bao đựng gạo với độ tin cậy $1 - \alpha = 0.95$. Công thức khoảng ước lượng cho trung bình quần thể là:

$$\bar{X} \pm z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}$$

Trong đó:

- $\bar{X} = 362.3$ gram
- $\sigma = 5$ gram
- $n = 35$
- $z_{\alpha/2} = 1.96$ (mức tin cậy 95%)

Ví dụ 1: Tính toán khoảng ước lượng

Bước 1: Tính sai số chuẩn (SE):

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{5}{\sqrt{35}} \approx 0.843$$

Bước 2: Tính độ rộng khoảng ước lượng E :

$$E = z_{\alpha/2} \times SE = 1.96 \times 0.843 \approx 1.653$$

Bước 3: Xây dựng khoảng ước lượng:

$$\bar{X} - E < \mu < \bar{X} + E$$

$$362.3 - 1.653 < \mu < 362.3 + 1.653$$

$$360.65 < \mu < 363.95$$

Kết luận: Với độ tin cậy 95%, khoảng ước lượng trung bình khối lượng các bao gạo là [360.65; 363.95] gram.

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thể

Ước lượng tỉ lệ quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tỉ lệ



Chọn Phân Phối Phù Hợp

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Sử dụng phân phối chuẩn:

- σ đã biết và quần thể có phân phối chuẩn hoặc $n > 30$.
- Công thức:

$$\bar{x} - t_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \quad (7)$$

Sử dụng phân phối t:

- σ chưa biết và quần thể có phân phối chuẩn hoặc $n < 30$.
- Công thức:

$$\bar{x} - t_{\alpha/2}^{n-1} \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2}^{n-1} \cdot \frac{s}{\sqrt{n}} \quad (8)$$



Ví dụ 2:

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

- Một thử nghiệm về hiệu quả của tỏi đối với mức cholesterol LDL (mg/dL).
- Số liệu:
 - Số mẫu: $n = 49$
 - Trung bình mẫu: $\bar{x} = 0.4$
 - Độ lệch chuẩn mẫu: $s = 21.0$
- Yêu cầu: Xây dựng khoảng tin cậy 95% cho thay đổi ròng trung bình trong nồng độ cholesterol LDL.

Giải ví dụ 2

■ Công thức khoảng tin cậy:

$$\bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

- Với $\alpha = 0.05$ và $df = n - 1 = 48$, tra bảng t : $t_{\alpha/2} = 2.01$.
- Tính sai số biên E :

$$E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} = 2.01 \cdot \frac{21.0}{\sqrt{49}} = 2.01 \cdot 3.0 = 6.03$$

■ Khoảng tin cậy:

$$0.4 \pm 6.03 \implies [-5.63, 6.43]$$

- Kết luận: Khoảng tin cậy 95% cho thấy thay đổi ròng trung bình trong nồng độ cholesterol LDL nằm trong khoảng từ -5.63 đến 6.43 mg/dL. Không đủ bằng chứng cho thấy tôi làm giảm cholesterol.

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thể

Ước lượng tÍt lít quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
đinh

Kiểm định về giá trị tÍt lít



Các Thuộc Tính Quan Trọng của Phân Phối t

Ước lượng

Giới thiệu

Ước lượng trung bình quân
thể

Ước lượng tỉ lệ quân thể

Kiểm định

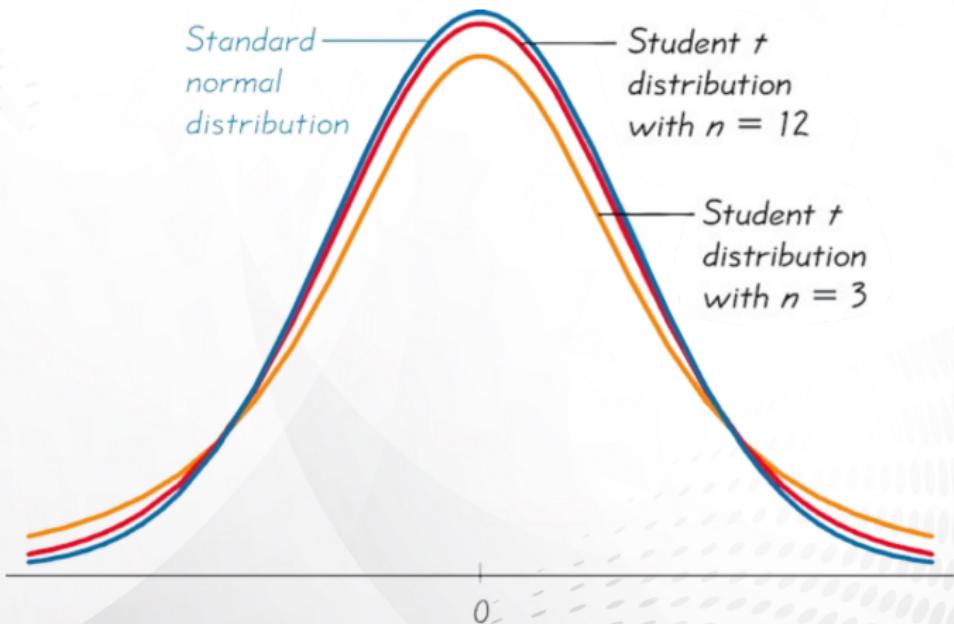
Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tỉ lệ

- 1 Phân phối t khác nhau đối với các cỡ mẫu khác nhau.
- 2 Phân phối t có:
 - Hình dạng chuông đối xứng, tương tự phân bố chuẩn.
 - Phản ánh sự biến thiên lớn hơn (với phân bố rộng hơn) khi cỡ mẫu nhỏ.
- 3 Giá trị trung bình của phân phối t : $t = 0$ (giống như phân phối chuẩn có giá trị trung bình là 0).
- 4 Độ lệch chuẩn của phân phối t :
 - Thay đổi theo cỡ mẫu và lớn hơn 1.
 - Không giống như phân phối chuẩn, có $\sigma = 1$.
- 5 Khi cỡ mẫu n lớn hơn, phân phối t sẽ tiến gần hơn đến phân phối chuẩn.

Ví dụ: Phân phối t với $n = 3$ và $n = 12$



Hình: Phân phối t với $n = 3$ và $n = 12$.

Kích Thước Mẫu: Ước Tính Trung Bình Quần Thể

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

- **Câu hỏi đặt ra:** Phải lấy mẫu kích thước bao nhiêu để ước tính trung bình quần thể chính xác?
- **Công thức xác định kích thước mẫu:**

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E} \right)^2 \quad (9)$$

Nếu cỡ mẫu được tính n không phải là số nguyên, hãy làm tròn giá trị của n đến số nguyên lớn hơn tiếp theo.

- **Ý nghĩa các ký hiệu:**

- μ : trung bình quần thể.
- σ : độ lệch chuẩn của quần thể.
- \bar{x} : trung bình mẫu.
- E : biên độ lỗi mong đợi (sai số tối đa cho phép).
- $z_{\alpha/2}$: giá trị z tương ứng với mức tin cậy $1 - \alpha$.

Tìm Kích Thước Mẫu n Khi σ Không Xác Định

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thể

Ước lượng tỉ lệ quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tỉ lệ

■ Công thức xác định kích thước mẫu:

$$n = \left(\frac{t_{\alpha/2} \cdot s}{E} \right)^2 \quad (10)$$

■ Trong đó:

- $t_{\alpha/2}$: giá trị t tương ứng với mức tin cậy $1 - \alpha$ và bậc tự do $df = n - 1$.
- s : độ lệch chuẩn mẫu.
- E : biên độ lỗi mong đợi.



Ví dụ

Giả sử rằng chúng ta muốn ước tính điểm số IQ trung bình của quần thể sinh viên ngành thống kê. Cần lấy mẫu bao nhiêu sinh viên để làm bài kiểm tra IQ để tin tưởng 95% rằng giá trị ước lượng chứa giá trị trung bình thực sự μ của quần thể với sai số là 3 điểm IQ.

■ Dữ liệu đầu vào:

- Mức độ tin cậy $1 - \alpha = 0.95$, do đó $\alpha = 0.05$ và $\alpha/2 = 0.025$.
- Giá trị $z_{\alpha/2} = 1.96$ từ phân phối chuẩn.
- Biên độ lỗi $E = 3$ điểm IQ.
- Độ lệch chuẩn quần thể $\sigma = 15$ điểm IQ.

■ Tính toán kích thước mẫu:

- Thay các giá trị vào công thức:

$$n = \left(\frac{1.96 \cdot 15}{3} \right)^2 = \left(\frac{29.4}{3} \right)^2 = 9.8^2 = 96.04 \rightarrow n = 97$$

Bài Tập 1:

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

- Một công ty sản xuất gạo muốn ước lượng trọng lượng trung bình của mỗi bao gạo trong kho. Họ chọn ngẫu nhiên một mẫu 50 bao gạo, và trọng lượng trung bình của mẫu là 45 kg với độ lệch chuẩn quần thể là 3 kg. Xây dựng khoảng tin cậy 95% cho trọng lượng trung bình của tất cả các bao gạo trong kho.
- Dữ liệu:
 - $n = 50$
 - $\bar{x} = 45 \text{ kg}$
 - $\sigma = 3 \text{ kg}$
 - Mức độ tin cậy $1 - \alpha = 0.95$
- Yêu cầu: Tính khoảng tin cậy 95% cho trọng lượng trung bình của quần thể.

Bài Tập 2:

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

- Một trường đại học muốn ước lượng điểm số trung bình của các sinh viên trong kỳ thi cuối kỳ. Họ lấy mẫu ngẫu nhiên 30 sinh viên và nhận được điểm trung bình mẫu là 75 điểm, với độ lệch chuẩn quần thể là 10 điểm. Xây dựng khoảng tin cậy 99% cho điểm số trung bình của toàn bộ sinh viên.
- Dữ liệu:
 - $n = 30$
 - $\bar{x} = 75$ điểm
 - $\sigma = 10$ điểm
 - Mức độ tin cậy $1 - \alpha = 0.99$
- Yêu cầu: Tính khoảng tin cậy 99% cho điểm số trung bình của quần thể sinh viên.

Bài Tập 3:

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

- Một nhóm nghiên cứu muốn ước lượng thời gian trung bình để hoàn thành một bài thi nghiệm trong phòng thí nghiệm. Họ thu thập dữ liệu từ 25 thí nghiệm, với thời gian trung bình là 120 phút và độ lệch chuẩn mẫu là 15 phút. Hãy xây dựng khoảng tin cậy 90% cho thời gian trung bình hoàn thành bài thi nghiệm trong quần thể.
- Dữ liệu:
 - $n = 25$
 - $\bar{x} = 120$ phút
 - $s = 15$ phút
 - Mức độ tin cậy $1 - \alpha = 0.90$
- Yêu cầu: Tính khoảng tin cậy 90% cho thời gian trung bình hoàn thành bài thi nghiệm trong quần thể.

Phần 1

Ước lượng

Mục 3: ƯỚC LƯỢNG TỈ LỆ QUẦN THỂ

Ước Lượng Tỷ Lệ Quần Thể

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỷ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỷ lệ

1 Tỉ lệ mẫu: Là ước lượng điểm tốt nhất của tỉ lệ quần thể.

2 Xây dựng khoảng tin cậy:

- Sử dụng tỉ lệ mẫu để xây dựng khoảng tin cậy cho tỉ lệ quần thể.
- Điển giải ý nghĩa của khoảng tin cậy để ước lượng giá trị đúng của tỉ lệ quần thể.

3 Xác định kích thước mẫu:

- Cần phải xác định kích thước mẫu cần thiết để ước lượng tỉ lệ quần thể với độ chính xác mong muốn.



Khoảng Ước Lượng Cho Tỷ Lệ p

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thể

Ước lượng tần số quản
thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tần số

- Phân phối tỷ lệ mẫu của biến X là phân phối nhị thức vì thỏa mãn các điều kiện sau:
 - Số lần thí nghiệm là cố định.
 - Kết quả thí nghiệm có thể phân thành 2 lớp: thành công hay thất bại.
 - Xác suất thành công trong mọi lần thí nghiệm là như nhau.
 - Các lần thí nghiệm là độc lập.
- X là số lần thí nghiệm thành công trong n lần thí nghiệm.
- Trong phương pháp nhị thức, tính xác suất khi số phép thử lớn (ví dụ như 100) là gần như không thể. Phân phối chuẩn có thể được dùng để xấp xỉ phân phối nhị thức khi n lớn.



Điều Kiện Để Xấp Xỉ Phân Phối Nhị Thức Bằng Phân Phối Chuẩn

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thể

Ước lượng tần số quản
thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tần
lệ

- Điều kiện để phân phối chuẩn có thể được dùng để xấp xỉ phân phối nhị thức khi n lớn:

$$np > 5 \quad \text{và} \quad n(1 - p) > 5 \quad (11)$$

- n càng lớn thì xấp xỉ càng chính xác.
- Khi phân phối nhị thức xấp xỉ phân phối chuẩn, phân phối của tỷ lệ mẫu có dạng chuẩn với các đặc điểm:
 - Trung bình: $\mu_{\hat{p}} = p$
 - Độ lệch chuẩn: $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$



Tính Ước Lượng Tỷ Lệ Quần Thể p

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỷ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỷ lệ

- Ước lượng điểm tốt nhất của tỷ lệ quần thể p là tỷ lệ mẫu \hat{p} , được tính bằng $\hat{p} = \frac{x}{n}$ với x là số lần thành công và n là tổng số thử nghiệm.
- Sử dụng tỷ lệ mẫu để xây dựng khoảng tin cậy cho tỷ lệ quần thể p :

$$\hat{p} \pm t_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \quad (12)$$

Trong đó:

- $t_{\alpha/2}$ là giá trị tương ứng với mức độ tin cậy $1 - \alpha$.
- \hat{p} là tỷ lệ mẫu.
- n là kích thước mẫu.



Ví dụ

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tần số quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tần số

- VỚI TỶ LỆ MẪU $\hat{p} = 0.85$, $n = 1007$, MỨC ĐỘ TIN CẬY 95%

$$E = 1.96 \times \sqrt{\frac{0.85(1 - 0.85)}{1007}} \approx 0.030$$

DO ĐÓ, KHOẢNG TIN CẬY CHO TỶ LỆ QUẦN THỂ p LÀ:

$$0.85 \pm 0.030$$

ĐIỀU NÀY CÓ NGHĨA LÀ CHÚNG TA TIN TƯỞNG 95% RẰNG TỶ LỆ QUẦN THỂ p NẰM TRONG KHOẢNG TỪ 0.82 ĐẾN 0.88.

Bài tập luyện tập

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thể

Ước lượng tỉ lệ quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tỉ lệ

Bài 1: Một công ty sản xuất bóng đèn muôn ước lượng độ sáng trung bình (đơn vị lumen) của loại bóng mới. Một mẫu ngẫu nhiên gồm 50 bóng đèn có độ sáng trung bình là 800 lumen. Giả sử độ lệch chuẩn tổng thể là 40 lumen.

- Hãy tính khoảng tin cậy 95% cho độ sáng trung bình của loại bóng này.
- Nếu muôn khoảng tin cậy có độ rộng không quá 5 lumen, kích thước mẫu tối thiểu cần lấy là bao nhiêu?
- Nếu lấy một mẫu có kích thước lớn hơn 50 thì khoảng tin cậy có rộng hơn hay hẹp hơn? Giải thích.
- Nếu độ lệch chuẩn tổng thể không được biết, cần phải sử dụng phương pháp nào để ước lượng?

Bài tập luyện tập

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Bài 2: Một nghiên cứu đo thời gian ngủ trung bình mỗi đêm của sinh viên đại học. Một mẫu ngẫu nhiên gồm 25 sinh viên có thời gian ngủ trung bình là 6,8 giờ với độ lệch chuẩn mẫu là 1,2 giờ.

- Hãy tính khoảng tin cậy 90% cho thời gian ngủ trung bình của sinh viên.
- Nếu độ lệch chuẩn thực tế của tổng thể lớn hơn độ lệch chuẩn mẫu, khoảng tin cậy có thay đổi thế nào?
- Nếu muốn sai số không vượt quá 0.5 giờ với độ tin cậy 95%, cần bao nhiêu sinh viên trong mẫu?
- Tại sao trong trường hợp này ta không thể dùng phân phối chuẩn để ước lượng?



Bài tập luyện tập

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thể

Ước lượng tần số quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tần số

Bài 3: Một nhóm nghiên cứu thu thập thời gian chờ trung bình của khách hàng tại một quầy thanh toán siêu thị (tính theo phút). Dữ liệu từ một mẫu ngẫu nhiên 10 khách hàng như sau:

5.2, 6.1, 4.8, 5.5, 5.9, 6.3, 4.7, 5.4, 6.0, 5.1

- Tính trung bình mẫu và độ lệch chuẩn mẫu.
- Hãy tính khoảng tin cậy 95% cho thời gian chờ trung bình của khách hàng.
- Nếu tăng kích thước mẫu lên 30, khoảng tin cậy có thay đổi thế nào?
- Trong trường hợp nào thì có thể sử dụng phân phối chuẩn thay vì phân phối t để ước lượng?



Bài tập luyện tập

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tần số quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tần số

Bài 4: Một nhà sản xuất đo lường hàm lượng đường (gram) trong một loại nước ép trái cây. Một mẫu ngẫu nhiên 100 chai có hàm lượng đường trung bình là 20,5 gram với độ lệch chuẩn mẫu là 3 gram.

- Hãy tính khoảng tin cậy 99% cho hàm lượng đường trung bình của loại nước ép này.
- Nếu chỉ lấy mẫu 25 chai thay vì 100 chai, có thể tiếp tục sử dụng phân phối chuẩn không? Vì sao?
- Nếu muốn sai số tối đa là 0.5 gram với độ tin cậy 95%, cần bao nhiêu chai nước ép trong mẫu?
- Khoảng tin cậy sẽ thay đổi thế nào nếu độ tin cậy tăng lên 99.9%?

Bài tập luyện tập

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Bài 5: Một khảo sát về mức lương (triệu đồng/tháng) của nhân viên trong một công ty thu được mẫu ngẫu nhiên sau:

12.5, 13.2, 11.8, 14.0, 13.5, 12.9, 11.7, 13.8, 12.6, 14.1

- Hãy tính trung bình mẫu và độ lệch chuẩn mẫu.
- Hãy tính khoảng tin cậy 95% cho mức lương trung bình của nhân viên công ty này.
- Nếu số lượng nhân viên trong mẫu tăng lên 50, có cần sử dụng phân phối t không? Giải thích.
- Nếu độ tin cậy tăng lên 99%, khoảng tin cậy sẽ rộng hơn hay hẹp hơn? Giải thích.



Bài tập luyện tập

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỷ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỷ lệ

- Bài 6:** Một cuộc khảo sát ngẫu nhiên 800 người về sở thích xem phim hành động cho thấy 480 người thích thể loại này.
- a) Hãy ước lượng khoảng tin cậy 95% cho tỷ lệ người thích xem phim hành động. b) Nếu muốn sai số không vượt quá 0.02 với độ tin cậy 95%, cần bao nhiêu người trong mẫu?
 - c) Nếu tỷ lệ thực tế của tổng thể thấp hơn tỷ lệ mẫu, khoảng tin cậy sẽ thay đổi thế nào?
 - d) Nếu lấy mẫu 1000 người thay vì 800 người, khoảng tin cậy sẽ thay đổi như thế nào?

Bài tập luyện tập

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thể

Ước lượng tần số quản
thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tần số

Bài 7: Một nghiên cứu khảo sát 60 sinh viên về việc có sử dụng điện thoại trong lớp học không. Kết quả có 15 sinh viên trả lời "Có".

- Hãy tính khoảng tin cậy 90% cho tỷ lệ sinh viên sử dụng điện thoại trong lớp học.
- Có thể dùng công thức chuẩn để tính khoảng tin cậy trong trường hợp này không? Giải thích.
- Nếu số lượng sinh viên khảo sát tăng lên 200, phương pháp ước lượng có thay đổi không? Vì sao?
- Nếu tăng mức độ tin cậy lên 95%, khoảng tin cậy sẽ thay đổi thế nào?



Bài tập luyện tập

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thể

Ước lượng tỷ lệ quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tỷ lệ

Bài 8: Một giáo viên ghi lại điểm số cuối kỳ của 15 học sinh trong lớp (thang điểm 10) như sau:

4.5, 6.2, 7.0, 5.8, 3.9, 8.1, 5.0, 6.5, 7.3, 4.8, 5.4, 6.0, 3.5, 7.8, 5.2

Biết rằng điểm qua môn là 5.0.

- Hãy xác định tỷ lệ học sinh qua môn trong mẫu này.
- Hãy ước lượng khoảng tin cậy 95% cho tỷ lệ học sinh qua môn.
- Nếu muốn sai số tối đa là 5% với độ tin cậy 95%, cần bao nhiêu học sinh trong mẫu?
- Nếu tăng kích thước mẫu lên 50 học sinh, khoảng tin cậy có thay đổi không? Giải thích.

Phần 2

Kiểm định

Phần 2

Kiểm định

Mục 1: CÁC KHÁI NIỆM CƠ BẢN

Mục tiêu chính

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Mục tiêu chính của phần này:

- Phát triển hiểu biết về phương pháp kiểm định giả thuyết thống kê.
- Xác định và xây dựng giả thuyết null (H_0) và giả thuyết thay thế (H_1).
- Tiến hành kiểm định giả thuyết cho các phát biểu về:
 - Giá trị trung bình của quần thể (μ).
 - Tỷ lệ của quần thể (p).
- Hiểu và áp dụng các phương pháp tính toán giá trị thống kê kiểm định.
- Đưa ra kết luận dựa trên kết quả kiểm định giả thuyết.



Định nghĩa kiểm định giả thuyết thống kê

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Giả thuyết là một phát biểu về một thuộc tính của quần thể.
Kiểm định giả thuyết thống kê là một thủ tục để kiểm định một phát biểu về một tham số của quần thể bằng cách sử dụng dữ liệu mẫu.

Quá trình kiểm định giả thuyết bao gồm:

- 1 Xây dựng giả thuyết null (H_0) và giả thuyết thay thế (H_1).
- 2 Thu thập dữ liệu mẫu và tính toán giá trị thống kê kiểm định.
- 3 So sánh giá trị thống kê kiểm định với giá trị tới hạn hoặc sử dụng giá trị p để đưa ra kết luận.

Giả thuyết Null và Giả thuyết Thay thế

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

- **Giả thuyết null (H_0):** là phát biểu cho rằng tham số quần thể bằng một giá trị cụ thể. Được kiểm định trực tiếp bằng cách giả định là đúng cho đến khi có đủ bằng chứng bác bỏ.
- **Giả thuyết thay thế (H_1 hoặc H_A):** là phát biểu cho rằng tham số quần thể có giá trị khác so với giả thuyết null. Được ký hiệu bằng các dấu: $<$, $>$, hoặc \neq .

Ví dụ: Bạn muốn kiểm tra liệu hơn 50% cử tri có ủng hộ ứng cử viên không.

- $H_0: p = 0.5$
- $H_1: p > 0.5$

Các loại kiểm định giả thuyết: Đuôi trái, đuôi phải, hai đuôi

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Kiểm định đuôi trái: Kiểm tra xem tham số quần thể nhỏ hơn một giá trị cụ thể.

- $H_0: \mu = 10, H_1: \mu < 10$

Kiểm định đuôi phải: Kiểm tra xem tham số quần thể lớn hơn một giá trị cụ thể.

- $H_0: \mu = 10, H_1: \mu > 10$

Kiểm định hai đuôi: Kiểm tra xem tham số quần thể có khác giá trị cụ thể không.

- $H_0: \mu = 10, H_1: \mu \neq 10$

Giả thuyết thay thế (H_1) luôn sử dụng một trong các dấu: $<$, $>$, hoặc \neq .



Phần 2

Kiểm định

Mục 2: KIỂM ĐỊNH VỀ GIÁ TRỊ TRUNG
BÌNH

Kiểm định về giá trị trung bình

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Mục tiêu:

- Kiểm định giả thuyết về giá trị trung bình của quần thể (μ).
- Xác định khi nào sử dụng phân phối z hoặc phân phối t.
- Hiểu cách tính giá trị thông kê kiểm định.
- Sử dụng giá trị p hoặc khoảng tin cậy để đưa ra kết luận.

Các trường hợp kiểm định:

- Khi phương sai của quần thể đã biết hoặc cỡ mẫu lớn ($n > 30$) \Rightarrow dùng phân phối z.
- Khi phương sai của quần thể chưa biết và cỡ mẫu nhỏ ($n \leq 30$) \Rightarrow dùng phân phối t.



Công thức kiểm định giá trị trung bình

Giá trị thống kê kiểm định phụ thuộc vào phân phối sử dụng:
Trường hợp phương sai đã biết hoặc mẫu lớn:

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \quad (13)$$

với \bar{x} là giá trị trung bình mẫu, σ là độ lệch chuẩn quần thể, n là kích thước mẫu.

Trường hợp phương sai chưa biết và mẫu nhỏ:

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \quad (14)$$

với s là độ lệch chuẩn mẫu. Trong trường hợp này, sử dụng phân phối Student- t với $df = n - 1$ bậc tự do.

Sau đó dùng **Bảng phân phối chuẩn**.

Ví dụ 1: Kiểm định đuôi phải ($H_1 : \mu > \mu_0$)

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Bài toán: Một công ty sản xuất pin khẳng định thời lượng trung bình của pin là 300 phút. Một nhóm kỹ sư ngờ rằng thời lượng pin có thể cao hơn. Để kiểm tra điều này, họ lấy mẫu $n = 36$ viên pin và đo được trung bình $\bar{x} = 310$ phút, với độ lệch chuẩn quần thể $\sigma = 20$ phút. Kiểm định với mức ý nghĩa $\alpha = 0.05$.

Bước 1: Đặt giả thuyết kiểm định

- $H_0 : \mu = 300$ (Thời lượng trung bình không thay đổi)
- $H_1 : \mu > 300$ (Thời lượng trung bình tăng lên)

Bước 2: Tính giá trị thống kê kiểm định

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{310 - 300}{20/\sqrt{36}} = \frac{10}{3.33} = 3.00$$

Ví dụ 1: Kiểm định đuôi phải ($H_1 : \mu > \mu_0$)

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Bước 3: Xác định giá trị p

$$p = P(Z > 3.00) = 1 - P(Z \leq 3.00)$$

Tra bảng phân phối chuẩn, $P(Z \leq 3.00) = 0.9987$.

$$p = 1 - 0.9987 = 0.0013$$

Bước 4: So sánh với mức ý nghĩa

- Vì $p = 0.0013 < \alpha = 0.05$, bác bỏ H_0 .
- Kết luận: Có đủ bằng chứng để khẳng định thời lượng trung bình pin lớn hơn 300 phút.



Ví dụ 2: Kiểm định đuôi trái ($H_1 : \mu < \mu_0$)

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Bài toán: Một nhà máy sản xuất dây điện có đường kính trung bình là 2.5 mm. Gần đây, có nghi ngờ rằng đường kính trung bình đã giảm. Một mẫu gồm $n = 40$ dây điện được kiểm tra, cho thấy trung bình $\bar{x} = 2.45$ mm, với độ lệch chuẩn quần thể $\sigma = 0.1$ mm. Kiểm định với mức ý nghĩa $\alpha = 0.01$.

Bước 1: Đặt giả thuyết kiểm định

- $H_0 : \mu = 2.5$ (Đường kính trung bình không đổi)
- $H_1 : \mu < 2.5$ (Đường kính trung bình giảm)

Bước 2: Tính giá trị thống kê kiểm định

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{2.45 - 2.5}{0.1 / \sqrt{40}} = \frac{-0.05}{0.0158} = -3.16$$

Ví dụ 2: Kiểm định đuôi trái ($H_1 : \mu < \mu_0$)

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Bước 3: Xác định giá trị p

$$p = P(Z < -3.16)$$

Tra bảng phân phối chuẩn, $P(Z < -3.16) = 0.0008$.

Bước 4: So sánh với mức ý nghĩa

- Vì $p = 0.0008 < \alpha = 0.01$, bác bỏ H_0 .
- Kết luận: Có đủ bằng chứng để kết luận rằng đường kính trung bình của dây điện đã giảm.

Ví dụ 3: Kiểm định giá trị trung bình 2 đuôi

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thái

Ước lượng tỉ lệ quản
thái

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tỉ lệ

Bài toán: Một nhà sản xuất bóng đèn cho biết tuổi thọ trung bình của bóng đèn là 350 giờ với độ lệch chuẩn $\sigma = 100$ giờ. Một khách hàng nghi ngờ tuyên bố này và lấy mẫu ngẫu nhiên 64 bóng đèn, tính được $\bar{x} = 340$ giờ. Kiểm định giả thuyết với mức ý nghĩa $\alpha = 0.05$.

Các bước thực hiện:

Bước 1: Xác định giả thuyết

- $H_0 : \mu = 350$ (Tuổi thọ trung bình đúng như công bố)
- $H_1 : \mu \neq 350$ (Tuổi thọ trung bình khác 350)

Bước 2: Tính giá trị thống kê kiểm định

$$z = \frac{340 - 350}{100/\sqrt{64}} = \frac{-10}{12.5} = -0.8$$

Ví dụ 3: Kiểm định giá trị trung bình 2 đuôi

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Bước 3: Xác định giá trị p

$$p = 2 \times P(Z < -0.8) = 2 \times 0.2119 = 0.4238$$

Bước 4: So sánh với mức ý nghĩa

- Vì $p = 0.4238 > \alpha = 0.05$, không bác bỏ H_0 .
- Kết luận: Không có đủ bằng chứng để khẳng định tuổi thọ trung bình khác 350 giờ.



Ví dụ 4: Kiểm định giá trị trung bình với phân phối t

Ước lượng

Giới thiệu

Ước lượng trung bình quản thể

Ước lượng tỉ lệ quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Bài toán: Một nhà hàng thức ăn nhanh muốn kiểm tra xem thời gian chờ trung bình có khác 4.5 phút hay không. Một mẫu gồm 25 khách hàng có thời gian chờ trung bình $\bar{x} = 5.1$ phút với độ lệch chuẩn mẫu $s = 1.2$ phút. Kiểm định giả thuyết với $\alpha = 0.05$.

Bước 1: Xác định giả thuyết

- $H_0 : \mu = 4.5$ (Thời gian chờ trung bình không đổi)
- $H_1 : \mu \neq 4.5$ (Thời gian chờ trung bình đã thay đổi)

Bước 2: Tính giá trị thống kê kiểm định

$$t = \frac{5.1 - 4.5}{1.2/\sqrt{25}} = \frac{0.6}{0.24} = 2.5$$

Ví dụ 4: Kiểm định giá trị trung bình với phân phối t

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Bước 3: Tính t_0 Với $df = 25 - 1 = 24$ và $\alpha = 0.05 \rightarrow \alpha/2 = 0.025$ tra **bảng Student-t**, ta có:

$$t_0 = 2.064$$

Bước 4: So sánh với mức ý nghĩa

- Vì $t = 2.5 > t_0 = 2.064$ nên bác bỏ H_0 .
- Kết luận: Thời gian chờ trung bình đã thay đổi đáng kể.



Phần 2

Kiểm định

Mục 3: KIỂM ĐỊNH VỀ GIÁ TRỊ TỈ LỆ

Kiểm định về giá trị tỉ lệ

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Mục tiêu:

- Kiểm định giả thuyết về tỷ lệ của quần thể (p).
- Sử dụng phân phối chuẩn để xấp xỉ phân phối nhị thức.
- Xác định giá trị thống kê kiểm định và sử dụng giá trị p để ra quyết định.

Điều kiện áp dụng:

- Mẫu ngẫu nhiên đơn giản.
- Các điều kiện để xấp xỉ phân phối chuẩn phải thỏa mãn:

$$np \geq 5, \quad nq \geq 5 \quad (15)$$

với $q = 1 - p$.



Công thức kiểm định giá trị tỉ lệ

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Giả thuyết kiểm định:

- $H_0 : p = p_0$ (Tỷ lệ quần thể không đổi)
- $H_1 : p \neq p_0$ hoặc $p > p_0$ hoặc $p < p_0$

Giá trị thống kê kiểm định:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} \quad (16)$$

với:

- $\hat{p} = \frac{x}{n}$: Tỷ lệ mẫu.
- p_0 : Tỷ lệ giả định theo giả thuyết null.
- $q_0 = 1 - p_0$.
- n : Kích thước mẫu.

Ví dụ 1: Kiểm định đuôi phải ($H_1 : p > p_0$)

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thế

Ước lượng tỉ lệ quản thế

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tỉ lệ

Bài toán: Một nhà sản xuất điện thoại tuyên bố rằng ít nhất 75% khách hàng hài lòng với sản phẩm của họ. Một cuộc khảo sát ngẫu nhiên với 500 khách hàng cho thấy 395 người hài lòng. Có đủ bằng chứng để kết luận rằng tỷ lệ khách hàng hài lòng thực sự cao hơn 75% không? Kiểm định với mức ý nghĩa $\alpha = 0.01$.

Bước 1: Đặt giả thuyết kiểm định

- $H_0 : p = 0.75$ (Tỷ lệ khách hàng hài lòng là 75%)
- $H_1 : p > 0.75$ (Tỷ lệ khách hàng hài lòng cao hơn 75%)

Bước 2: Tính giá trị thống kê kiểm định

$$\hat{p} = \frac{395}{500} = 0.79, \quad z = \frac{0.79 - 0.75}{\sqrt{\frac{0.75 \times 0.25}{500}}} = \frac{0.04}{0.0194} = 2.06$$

Ví dụ 1: Kiểm định đuôi phải ($H_1 : p > p_0$)

Ước lượng

Giới thiệu

Ước lượng trung bình quản thể

Ước lượng tỉ lệ quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Bước 3: Xác định giá trị p

$$p = P(Z > 2.06)$$

Tra bảng chuẩn, $P(Z < 2.06) = 0.9803$.

$$p = 1 - 0.9803 = 0.0197$$

Bước 4: So sánh với mức ý nghĩa

- Vì $p = 0.0197 > \alpha = 0.01$, không bác bỏ H_0 .
- Kết luận: Không có đủ bằng chứng để khẳng định rằng tỷ lệ khách hàng hài lòng thực sự cao hơn 75%.



Ví dụ 2: Kiểm định đuôi trái ($H_1 : p < p_0$)

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thái

Ước lượng tỷ lệ quản
thái

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tỉ lệ

Bài toán: Một trường đại học tuyên bố rằng tỷ lệ sinh viên tốt nghiệp trong vòng 4 năm là ít nhất 85%. Một nghiên cứu ngẫu nhiên với 400 sinh viên tốt nghiệp cho thấy chỉ có 320 sinh viên hoàn thành trong 4 năm. Có đủ bằng chứng để kết luận rằng tỷ lệ sinh viên tốt nghiệp trong 4 năm thấp hơn 85% không? Kiểm định với mức ý nghĩa $\alpha = 0.05$.

Bước 1: Đặt giả thuyết kiểm định

- $H_0 : p = 0.85$ (Tỷ lệ sinh viên tốt nghiệp đúng hạn là 85%)
- $H_1 : p < 0.85$ (Tỷ lệ sinh viên tốt nghiệp đúng hạn thấp hơn 85%)

Bước 2: Tính giá trị thống kê kiểm định

$$\hat{p} = \frac{320}{400} = 0.80, \quad z = \frac{0.80 - 0.85}{\sqrt{\frac{0.85 \times 0.15}{400}}} = \frac{-0.05}{0.0165} = -3.03$$

Ví dụ 2: Kiểm định đuôi trái ($H_1 : p < p_0$)

Ước lượng

Giới thiệu

Ước lượng trung bình quản thể

Ước lượng tỉ lệ quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Bước 3: Xác định giá trị p

$$p = P(Z < -3.03)$$

Tra bảng chuẩn, $P(Z < -3.03) = 0.0012$.

Bước 4: So sánh với mức ý nghĩa

- Vì $p = 0.0012 < \alpha = 0.05$, bác bỏ H_0 .
- Kết luận: Có đủ bằng chứng để kết luận rằng tỷ lệ sinh viên tốt nghiệp đúng hạn thực sự thấp hơn 85%.

Ví dụ 3: Kiểm định hai đuôi ($H_1 : p \neq p_0$)

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thái

Ước lượng tỉ lệ quản
thái

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tỉ lệ

Bài toán: Theo thống kê, 93% chủ sở hữu máy tính tin rằng họ có phần mềm chống vi-rút. Một cuộc khảo sát với 400 người cho thấy chỉ có 380 người thực sự có phần mềm chống vi-rút. Có đủ bằng chứng để bác bỏ tỷ lệ 93% không? Kiểm định với mức ý nghĩa $\alpha = 0.05$.

Bước 1: Đặt giả thuyết kiểm định

- $H_0 : p = 0.93$ (Tỷ lệ thực sự là 93%)
- $H_1 : p \neq 0.93$ (Tỷ lệ thực sự khác 93%)

Bước 2: Tính giá trị thống kê kiểm định

$$\hat{p} = \frac{380}{400} = 0.95, \quad z = \frac{0.95 - 0.93}{\sqrt{\frac{0.93 \times 0.07}{400}}} = \frac{0.02}{0.013} = 1.54$$

Ví dụ 4: Kiểm định hai đuôi ($H_1 : p \neq p_0$)

Ước lượng

Giới thiệu

Ước lượng trung bình quần thể

Ước lượng tỉ lệ quần thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Bước 3: Xác định giá trị p

$$p = 2 \times P(Z > 1.54)$$

Tra bảng chuẩn, $P(Z < 1.54) = 0.9380$.

$$p = 2 \times (1 - 0.9380) = 2 \times 0.0620 = 0.124$$

Bước 4: So sánh với mức ý nghĩa

- Vì $p = 0.124 > \alpha = 0.05$, không bác bỏ H_0 .
- Kết luận: Không có đủ bằng chứng để bác bỏ tỷ lệ 93%.



Tóm tắt kiểm định giá trị tỉ lệ

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thế

Ước lượng tỉ lệ quản thế

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tỉ lệ

Các bước thực hiện kiểm định giả thuyết:

- Xác định giả thuyết null (H_0) và giả thuyết thay thế (H_1).
- Kiểm tra điều kiện áp dụng phân phối chuẩn:

$$np \geq 5, \quad nq \geq 5$$

- Tính giá trị thống kê kiểm định:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

- Xác định giá trị p hoặc so sánh với giá trị tới hạn.

- Đưa ra kết luận:

- Nếu $p < \alpha$, bác bỏ H_0 .
- Nếu $p > \alpha$, không bác bỏ H_0 .



Bài tập: Kiểm định về giá trị trung bình

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thể

Ước lượng tỉ lệ quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tỉ lệ

Bài 1: Một công ty sản xuất nước giải khát cho rằng lượng nước trong mỗi chai trung bình là 500ml. Một khách hàng nghi ngờ và đo ngẫu nhiên 30 chai, thu được trung bình 495ml, độ lệch chuẩn 8ml. Kiểm định với $\alpha = 0.05$ xem lượng nước có khác 500ml không.

Bài 2: Một hãng xe tuyên bố mức tiêu thụ nhiên liệu trung bình của xe họ là 6.5L/100km. Một kiểm tra với 25 chiếc xe cho thấy mức tiêu thụ trung bình 6.8L/100km, độ lệch chuẩn 0.4L/100km. Kiểm định với $\alpha = 0.01$ xem xe có tiêu hao nhiên liệu hơn không.

Bài 3: Một nhà sản xuất thép cho rằng độ bền trung bình của thép là 1200MPa. Một nhóm kỹ sư kiểm tra 40 mẫu và thu được trung bình 1185MPa, độ lệch chuẩn 50MPa. Kiểm định với $\alpha = 0.05$ xem độ bền có giảm không.

Bài tập: Kiểm định về giá trị trung bình (tt)

Ước lượng

Giới thiệu

Ước lượng trung bình quản thể

Ước lượng tỉ lệ quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỉ lệ

Bài 4: Một giáo viên nhận xét rằng điểm trung bình của học sinh lớp A là 75. Một khảo sát với 36 học sinh thu được trung bình 78, độ lệch chuẩn 5 điểm. Kiểm định với $\alpha = 0.05$ xem điểm trung bình có cao hơn 75 không.

Bài 5: Một công ty giao hàng cam kết thời gian giao hàng trung bình là 2 ngày. Một khảo sát với 50 đơn hàng cho thấy trung bình 2.3 ngày, độ lệch chuẩn 0.5 ngày. Kiểm định với $\alpha = 0.05$ xem thời gian giao hàng có chậm hơn không.

Bài tập: Kiểm định về giá trị tỷ lệ

Ước lượng

Giới thiệu

Ước lượng trung bình quản
thế

Ước lượng tỉ lệ quản thế

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung
bình

Kiểm định về giá trị tỉ lệ

Bài 1: Một cuộc khảo sát trước bầu cử cho rằng 55% cử tri ủng hộ ứng cử viên A. Một cuộc thăm dò với 500 cử tri cho thấy 270 người ủng hộ A. Kiểm định với $\alpha = 0.05$ xem tỷ lệ thực tế có khác 55% không.

Bài 2: Một công ty bảo hiểm cho rằng 90% khách hàng hài lòng với dịch vụ của họ. Một khảo sát với 400 khách hàng cho thấy 350 người hài lòng. Kiểm định với $\alpha = 0.05$ xem tỷ lệ hài lòng có thấp hơn 90% không.

Bài 3: Một trường đại học tuyên bố 80% sinh viên tốt nghiệp có việc làm trong vòng 6 tháng. Một khảo sát với 300 sinh viên tốt nghiệp cho thấy 235 người có việc làm. Kiểm định với $\alpha = 0.05$ xem tỷ lệ có thấp hơn 80% không.

Bài tập: Kiểm định về giá trị tỷ lệ (tt)

Ước lượng

Giới thiệu

Ước lượng trung bình quản thể

Ước lượng tỷ lệ quản thể

Kiểm định

Các khái niệm cơ bản

Kiểm định về giá trị trung bình

Kiểm định về giá trị tỷ lệ

Bài 4: Một công ty viễn thông cho rằng 60% khách hàng sử dụng gói cước cao cấp. Một khảo sát với 600 khách hàng cho thấy 390 người sử dụng gói cao cấp. Kiểm định với $\alpha = 0.05$ xem tỷ lệ thực tế có cao hơn 60% không.

Bài 5: Một bệnh viện báo cáo rằng tỷ lệ bệnh nhân phục hồi sau phẫu thuật là 85%. Một nghiên cứu với 250 bệnh nhân cho thấy 200 người phục hồi. Kiểm định với $\alpha = 0.01$ xem tỷ lệ phục hồi có giảm không.

Good luck!