Xác suất thống kê - Khoảng tin cậy

Dẫn nhập

Khoảng tin cậy

Khoảng tin cậy cho kì vọng

Trường hợp phương sai đã biết

Trường hợp phương sai chưa biết

Khoảng tin cậy cho tỉ lệ

Xác suất thống kê - Khoảng tin cậy

Ngày 22 tháng 11 năm 2024

Xác suất thống kê - Khoảng tin cậy

Dẫn nhập

Khoảng tin cậy

Khoảng tin cậy cho kì vọng

Trường hợp phương sai đã biết

Trường hợp phương sai chưa biết

Khoảng tin cậy cho tỉ lệ

Ví dụ

Trong một xưởng sản xuất bóng đèn, người ta muốn tính tỉ lệ bóng đèn bị lỗi trong lô hàng để gửi đi. Người ta lấy ngẫu nhiên 80 bóng đèn, và người ta thấy có 15 bóng đèn bị lỗi.

Xác suất thống kê - Khoảng tin cậy

Dẫn nhập

Khoảng tin cậy

Khoảng tin cậy cho kì vọng

Trường hợp phương sai đã biết

Trường hợp phương sai chưa biết

Khoảng tin cậy cho tỉ lệ

Ví dụ

Trong một xưởng sản xuất bóng đèn, người ta muốn tính tỉ lệ bóng đèn bị lỗi trong lô hàng để gửi đi. Người ta lấy ngẫu nhiên 80 bóng đèn, và người ta thấy có 15 bóng đèn bị lỗi.

- Như vậy, tỉ lệ thực bóng đèn bị lỗi trong lô hàng này là bao nhiêu ?

Dẫn nhập

Ví dụ

Trong một xưởng sản xuất bóng đèn, người ta muốn tính tỉ lệ bóng đèn bị lỗi trong lô hàng để gửi đi. Người ta lấy ngẫu nhiên 80 bóng đèn, và người ta thấy có 15 bóng đèn bị lỗi.

• Như vậy, tỉ lệ thực bóng đèn bị lỗi trong lô hàng này là bao nhiêu ?

• Nếu bên nhận yêu cầu lô hàng ***không được quá*** 10% tỉ lệ bóng đèn bị lỗi, thì lô hàng có đạt tiêu chuẩn hay không ?

Xác suất thống kê - Khoảng tin cậy

Dẫn nhập

Khoảng tin cậy

Khoảng tin cậy cho kì vọng

Trường hợp phương sai đã biết

Trường hợp phương sai chưa biết

Khoảng tin cậy cho tỉ lệ

Thống kê

Cho mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$ được lấy từ cùng một phân phối f có tham số $\theta \in \Theta$.
Ta nói biến ngẫu nhiên T là một thống kê nếu T chỉ phụ thuộc vào mẫu $(X_1, \ldots, X_n)$

Ví dụ

Cho mẫu ngẫu nhiên $(X_1, \ldots, X_n)$ lấy từ phân phối $\mathcal{B}(10, p)$.

- $T_1 = X_1$, $T_2 = \dfrac{X_1 + X_2}{2}$, $T_3 = X_1 \cdot X_2$, $T_4 = \overline{X}$, $T_5 = S^2$ đều là các thống kê.
- $T_6 = X_1 + p$, $T_7 = \mathbb{P}(X_1 = 0)$, $T_8 = X_1^p$, đều không phải là thống kê.

Xác suất thống kê - Khoảng tin cậy

Dẫn nhập

Khoảng tin cậy

Khoảng tin cậy cho kì vọng

Trường hợp phương sai đã biết

Trường hợp phương sai chưa biết

Khoảng tin cậy cho tỉ lệ

Khoảng tin cậy

Cho mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$ được lấy từ phân phối f có tham số θ. Khi đó khoảng tin cậy là hai thống kê $L = L(X_1, \ldots, X_n), U = U(X_1, \ldots, X_n)$ sao cho $L \leq U$. Ta kí hiệu khoảng tin cậy này là $[L, U]$.

Các loại khoảng tin cậy

Cho mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$ được lấy từ phân phối f có tham số $\theta \in \Theta$, và cho khoảng tin cậy $[L, U]$. Ta nói:

- ***Khoảng tin cậy có độ tin cậy*** $(1 - \alpha) \cdot 100\%$ nếu
$$\mathbb{P}(\theta \in [L, U]) = 1 - \alpha \ \forall \theta \in \Theta$$
- ***Khoảng tin cậy có độ tin cậy xấp xỉ*** $(1 - \alpha) \cdot 100\%$ nếu
$$\lim_{n \to \infty} \mathbb{P}(\theta \in [L, U]) = 1 - \alpha \ \forall \theta \in \Theta$$

Xác suất thống kê - Khoảng tin cậy

Dẫn nhập

Khoảng tin cậy

Khoảng tin cậy cho kì vọng

Trường hợp phương sai đã biết

Trường hợp phương sai chưa biết

Khoảng tin cậy cho tỉ lệ

Khoảng tin cậy cho kì vọng - Phương sai biết

Cho mẫu ngẫu nhiên $(X_1, X_2, \dots, X_n)$ được lấy từ tổng thể có phân phối f. Kí hiệu $\mu = \mathbb{E}(X_1), \sigma^2 = \mathrm{Var}(X_1)$. Giả sử μ chưa biết và σ^2 đã biết.

Khi đó, khoảng tin cậy (xấp xỉ) $(1-\alpha)\cdot 100\%$ cho μ được cho bởi:

$$\left[\overline{X} - \frac{z_{1-\alpha/2}}{\sqrt{n}}\cdot\sigma,\ \overline{X} + \frac{z_{1-\alpha/2}}{\sqrt{n}}\cdot\sigma\right]$$

Nếu f không là phân phối chuẩn thì cỡ mẫu n phải thỏa $n \geq 30$.

Xác suất thống kê - Khoảng tin cậy

Dẫn nhập

Khoảng tin cậy

Khoảng tin cậy cho kì vọng

Trường hợp phương sai đã biết

Trường hợp phương sai chưa biết

Khoảng tin cậy cho tỉ lệ

Khoảng tin cậy cho kì vọng - Phương sai chưa biết

Cho mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$ được lấy từ tổng thể có phân phối f. Kí hiệu $\mu = \mathbb{E}(X_1), \sigma^2 = \mathrm{Var}(X_1)$. Giả sử μ và σ^2 chưa biết.

- Nếu f là phân phối chuẩn, KTC $(1-\alpha)\cdot 100\%$ cho μ được cho bởi:

$$\left[\overline{X} - \frac{t_{1-\alpha/2}^{n-1}}{\sqrt{n}} S, \overline{X} + \frac{t_{1-\alpha/2}^{n-1}}{\sqrt{n}} S\right]$$

- Nếu f không phải là phân phối chuẩn, thì cỡ mẫu n phải thỏa $n \geq 30$, và KTC $(1-\alpha)\cdot 100\%$ cho μ được cho bởi:

$$\left[\overline{X} - \frac{z_{1-\alpha/2}}{\sqrt{n}} S, \overline{X} + \frac{z_{1-\alpha/2}}{\sqrt{n}} S\right]$$

Nhận xét.

Nếu cỡ mẫu $n \geq 30$, hai khoảng tin cậy trên xấp xỉ nhau.

Xác suất thống kê - Khoảng tin cậy

Dẫn nhập

Khoảng tin cậy

Khoảng tin cậy cho kì vọng

Trường hợp phương sai đã biết

Trường hợp phương sai chưa biết

Khoảng tin cậy cho tỉ lệ

Các bước tìm khoảng tin cậy của kì vọng

(B1) Tìm cỡ mẫu n, và giá trị α.

(B2) Tính trung bình mẫu $\bar{x}$, phương sai mẫu s^2.

(B3) Tìm dung sai ϵ:

- Nếu $n \geq 30$, và σ^2 đã biết,
$$\epsilon = \frac{z_{1-\alpha/2}}{\sqrt{n}}\sigma$$
- Nếu $n \geq 30$, và σ^2 chưa biết,
$$\epsilon = \frac{z_{1-\alpha/2}}{\sqrt{n}}s$$
- Nếu $n < 30$, mẫu phải có phân phối chuẩn, và σ^2 đã biết,
$$\epsilon = \frac{z_{1-\alpha/2}}{\sqrt{n}}\sigma$$

Xác suất thống kê - Khoảng tin cậy

Dẫn nhập

Khoảng tin cậy

Khoảng tin cậy cho kì vọng

Trường hợp phương sai đã biết

Trường hợp phương sai chưa biết

Khoảng tin cậy cho tỉ lệ

Các bước tìm khoảng tin cậy của kì vọng

(B3) Tìm dung sai ϵ: (tt)

- Nếu $n < 30$, mẫu phải có phân phối chuẩn, và σ^2 chưa biết,

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

(B4) **Kết luận:** Khoảng tin cậy với độ tin cậy $(1-\alpha)\cdot 100\%$ cho μ là: $[\bar{x}-\epsilon, \bar{x}+\epsilon]$

Ví dụ minh họa

Ví dụ 1: Bài tập 4.16

Kinh nghiệm thực tế đã chỉ ra rằng sức bền của sợi được sử dụng trong sản xuất vật liệu màn treo có phân phối chuẩn với độ lệch chuẩn là 2 psi. Một mẫu ngẫu nhiên gồm 9 sợi được thử nghiệm và sức bền trung bình được tìm thấy là 98 psi. Tìm khoảng tin cậy 95% cho sức bền trung bình.

Ví dụ minh họa

Ví dụ 1: Bài tập 4.16

Kinh nghiệm thực tế đã chỉ ra rằng sức bền của sợi được sử dụng trong sản xuất vật liệu màn treo có phân phối chuẩn với độ lệch chuẩn là 2 psi. Một mẫu ngẫu nhiên gồm 9 sợi được thử nghiệm và sức bền trung bình được tìm thấy là 98 psi. Tìm khoảng tin cậy 95% cho sức bền trung bình.

Ví dụ 2: Bài tập 4.19

Một mẫu ngẫu nhiên gồm chiều cao của 50 sinh viên của một trường cao đẳng cho thấy trung bình là 174.5 cm và độ lệch chuẩn 6.9 cm. Xây dựng khoảng tin cậy 98% cho chiều cao trung bình μ của tất cả các sinh viên của trường cao đẳng này.

Ví dụ 3: Bài tập 4.45

Đo đường kính của một số chi tiết máy do một máy tiện tự động sản xuất, ta ghi nhận được số liệu như sau:

X	N
12,00	2
12,05	3
12,10	7
12,15	9
12,20	10
12,25	8
12,30	6
12,35	5
12,40	3

với N chỉ số trường hợp tính theo từng giá trị của X (mm).

Ước lượng đường kính trung bình μ với độ tin cậy 0.95.

Xác suất thống kê - Khoảng tin cậy

Dẫn nhập

Khoảng tin cậy

Khoảng tin cậy cho kì vọng

Trường hợp phương sai đã biết

Trường hợp phương sai chưa biết

Khoảng tin cậy cho tỉ lệ

Trung bình mẫu, phương sai mẫu dạng bảng tần số

Cho bảng tần số sau:

X	x_1	x_2	$\ldots$	x_{n-1}	x_n
N	N_1	N_2	$\ldots$	N_{n-1}	N_n

trong đó N_k chỉ số lần giá trị x_k xuất hiện trong mẫu quan sát được. Khi đó, trung bình mẫu và phương sai mẫu được cho bởi

$$\bar{x} = \frac{1}{\sum_{k=1}^{n} N_k} \sum_{k=1}^{n} N_k x_k$$

và

$$s^2 = \frac{1}{\left(\sum_{k=1}^{n} N_k\right) - 1} \sum_{k=1}^{n} N_k (x_k - \bar{x})^2$$

Xác suất thống kê - Khoảng tin cậy
Dẫn nhập
Khoảng tin cậy
Khoảng tin cậy cho kì vọng
Trường hợp phương sai đã biết
Trường hợp phương sai chưa biết
Khoảng tin cậy cho tỉ lệ

Sai số của khoảng tin cậy cho kì vọng

Cho khoảng tin cậy của kì vọng: $[\bar{x} - \epsilon, \bar{x} + \epsilon]$. Khi đó, sai số của khoảng tin cậy này là dung sai ϵ.

Nhận xét

Dung sai ϵ của khoảng tin cậy cho kì vọng thường có dạng

$$\epsilon = \frac{\text{Hằng số}}{\sqrt{n}}$$

Do đó, để sai số của khoảng tin cậy không quá E thì

$$\epsilon \leq E \iff n \geq \frac{(\text{Hằng số})^2}{E^2}$$

Xác suất thống kê - Khoảng tin cậy

Dẫn nhập

Khoảng tin cậy

Khoảng tin cậy cho kì vọng

Trường hợp phương sai đã biết

Trường hợp phương sai chưa biết

Khoảng tin cậy cho tỉ lệ

Ví dụ 3: Bài tập 4.45 (tt)

Nếu muốn sai số ước lượng không quá $E = 0.02$ mm ở độ tin cậy 0.95 thì phải quan sát ít nhất mấy trường hợp ?

Khoảng tin cậy cho tỉ lệ

Khoảng tin cậy cho tỉ lệ

Cho mẫu ngẫu nhiên $(X_1, X_2, \ldots, X_n)$ và đặt

$$Y_k = \begin{cases} 1, & \text{nếu } X_k \text{ có tính chất } \mathcal{A} \\ 0, & \text{ngược lại} \end{cases}$$

Khi đó, nếu p là tỉ lệ của tổng thể có tính chất $\mathcal{A}$ và cỡ mẫu $n \geq 30$, thì khoảng tin cậy $(1-\alpha)\cdot 100\%$ cho p là $[\bar{p} - \epsilon, \bar{p} + \epsilon]$ với dung sai

$$\epsilon = z_{1-\alpha/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

trong đó, $\bar{p} = \frac{1}{n}\sum_{k=1}^{n} Y_k$.

Xác suất thống kê - Khoảng tin cậy

Dẫn nhập

Khoảng tin cậy

Khoảng tin cậy cho kì vọng

Trường hợp phương sai đã biết

Trường hợp phương sai chưa biết

Khoảng tin cậy cho tỉ lệ

Các bước tìm khoảng tin cậy của tỉ lệ

(B1) Tìm cỡ mẫu n, giá trị α và $z_{1-\alpha/2}$.

(B2) Tính $\bar{p}$.

(B3) Tìm dung sai $\epsilon = z_{1-\alpha/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$.

(B4) **Kết luận:** Khoảng tin cậy với độ tin cậy $(1-\alpha)\cdot 100\%$ cho p là $[\bar{p}-\epsilon, \bar{p}+\epsilon]$.

Sai số của khoảng tin cậy cho tỉ lệ

Cho khoảng tin cậy của tỉ lệ: $[\bar{p} - \epsilon, \bar{p} + \epsilon]$. Khi đó, sai số của khoảng tin cậy này là dung sai ϵ.

Do đó, để sai số của khoảng tin cậy không quá E thì

$$n \geq (z_{1-\alpha/2})^2 \frac{\bar{p}(1-\bar{p})}{E^2}$$

Trong trường hợp $\bar{p}$ không biết, thì

$$n \geq \frac{(z_{1-\alpha/2})^2}{4E^2}$$

Dẫn nhập
Khoảng tin cậy
Khoảng tin cậy cho kì vọng
Trường hợp phương sai đã biết
Trường hợp phương sai chưa biết
Khoảng tin cậy cho tỉ lệ

Các ví dụ minh họa

Ví dụ 1: Bài tập 4.53

Trong số 1000 trường hợp ung thư phổi được chọn ngẫu nhiên thì có 823 trường hợp tử vong trong vòng 10 năm.

(a) Tìm khoảng tin cậy 95% cho tỉ lệ tử vong do ung thư phổi.

(b) Hỏi cỡ mẫu tối thiểu để sai số khi ước lượng giá trị thực của p nhỏ hơn 0.03 với độ tin cậy 0.95 là bao nhiêu ?

Các ví dụ minh họa

Xác suất thống kê - Khoảng tin cậy

Dẫn nhập

Khoảng tin cậy

Khoảng tin cậy cho kì vọng

Trường hợp phương sai đã biết

Trường hợp phương sai chưa biết

Khoảng tin cậy cho tỉ lệ

Ví dụ 1: Bài tập 4.53

Trong số 1000 trường hợp ung thư phổi được chọn ngẫu nhiên thì có 823 trường hợp tử vong trong vòng 10 năm.

(a) Tìm khoảng tin cậy 95% cho tỉ lệ tử vong do ung thư phổi.

(b) Hỏi cỡ mẫu tối thiểu để sai số khi ước lượng giá trị thực của p nhỏ hơn 0.03 với độ tin cậy 0.95 là bao nhiêu ?

Ví dụ 2: Bài tập 4.56

Một nghiên cứu sẽ được thực hiện về tỉ lệ gia đình có ít nhất hai ti vi. Một mẫu được yêu cầu phải lớn đến mức nào nếu muốn tự tin rằng sai số khi ước tính số lượng này nhỏ hơn 0.017 với độ tin cậy 99% ?