

# Lý thuyết Xác suất và Thống kê Toán học

PGS.TS. Trần Lộc Hùng

Tp. Hồ Chí Minh, 10/ 2013

Ngày 12 tháng 10 năm 2013

-----

# LÝ THUYẾT XÁC SUẤT VÀ THỐNG KÊ TOÁN HỌC

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ↺ 🔍 ↻

# Lý thuyết Xác suất và Thống kê Toán học

PGS.TS. Trần Lộc Hùng

Tp. Hồ Chí Minh, 10/ 2013

Ngày 12 tháng 10 năm 2013

## Chương 6. Ước lượng tham số tổng thể

# Từ khóa (Key Words)

- Ước lượng tham số

# Từ khóa (Key Words)

- Ước lượng tham số
- Độ tin cậy

# Từ khóa (Key Words)

- Ước lượng tham số
- Độ tin cậy
- Khoảng ước lượng

# Chương 6. Ước lượng tham số tổng thể

## 1 Đặt vấn đề

# Chương 6. Ước lượng tham số tổng thể

- 1 Đặt vấn đề
- 2 Ước lượng điểm



# Chương 6. Ước lượng tham số tổng thể

- 1 Đặt vấn đề
- 2 Ước lượng điểm
- 3 Ước lượng khoảng

# Chương 6. Ước lượng tham số tổng thể

- ① Đặt vấn đề
- ② Ước lượng điểm
- ③ Ước lượng khoảng
- ④ Ước lượng và cỡ mẫu

# Chương 6. Ước lượng tham số tổng thể

- ① Đặt vấn đề
- ② Ước lượng điểm
- ③ Ước lượng khoảng
- ④ Ước lượng và cỡ mẫu
- ⑤ Bài tập

# Đặt vấn đề

Giả sử  $\omega_n = \{X_1, X_2, \dots, X_n\}$  là một mẫu ngẫu nhiên sinh bởi biến ngẫu nhiên  $X$  có quy luật xác suất  $Q(x, \theta)$ . Vấn đề đặt ra là:

- 1 Ước lượng tham số chưa biết  $\theta$  của tổng thể  $\Omega$  bao gồm  $\mu, \sigma^2, p$ .

# Đặt vấn đề

Giả sử  $\omega_n = \{X_1, X_2, \dots, X_n\}$  là một mẫu ngẫu nhiên sinh bởi biến ngẫu nhiên  $X$  có quy luật xác suất  $Q(x, \theta)$ . Vấn đề đặt ra là:

- 1 Ước lượng tham số chưa biết  $\theta$  của tổng thể  $\Omega$  bao gồm  $\mu, \sigma^2, p$ .
- 2 Hàm ước lượng là các thống kê của mẫu  
$$\hat{\theta} = f(\omega_n) = f(X_1, X_2, \dots, X_n)$$

# Đặt vấn đề

Giả sử  $\omega_n = \{X_1, X_2, \dots, X_n\}$  là một mẫu ngẫu nhiên sinh bởi biến ngẫu nhiên  $X$  có quy luật xác suất  $Q(x, \theta)$ . Vấn đề đặt ra là:

- 1 Ước lượng tham số chưa biết  $\theta$  của tổng thể  $\Omega$  bao gồm  $\mu, \sigma^2, p$ .
- 2 Hàm ước lượng là các thống kê của mẫu  
$$\hat{\theta} = f(\omega_n) = f(X_1, X_2, \dots, X_n)$$
- 3 Ước lượng điểm

## 6.1 Ước lượng điểm

- 1 Ước lượng không chệch
- 2 Ước lượng vững
- 3 Ước lượng hiệu quả
- 4 Ước lượng hợp lý cực đại
- 5 Ước lượng theo mô men

# Ước lượng không chệch

## Định nghĩa

Thông kê  $\hat{\theta}$  được gọi là một ước lượng không chệch của tham số  $\theta$ , nếu

$$E(\hat{\theta}) = \theta$$

- 1 Bản chất là đẳng thức

$$E(\hat{\theta} - \theta) = 0$$



# Ước lượng không chệch

## Định nghĩa

Thống kê  $\hat{\theta}$  được gọi là một ước lượng không chệch của tham số  $\theta$ , nếu

$$E(\hat{\theta}) = \theta$$

- 1 Bản chất là đẳng thức

$$E(\hat{\theta} - \theta) = 0$$

- 2 Nếu  $E(\hat{\theta}) \neq \theta$ , thì thống kê  $\hat{\theta}$  là ước lượng chệch so với tham số  $\theta$

# Ước lượng không chệch

## Ví dụ 1

Trung bình mẫu  $\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$  là một ước lượng không chệch của tham số tổng thể  $\mu$

- ① Dùng định nghĩa và tính chất của kỳ vọng, ta có

$$E(\bar{X}) = E\left(\frac{1}{n} \sum_{j=1}^n X_j\right) = \frac{1}{n} \sum_{j=1}^n E(X_j) = \mu$$

# Ước lượng không chệch

## Ví dụ 1

Trung bình mẫu  $\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$  là một ước lượng không chệch của tham số tổng thể  $\mu$

- ① Dùng định nghĩa và tính chất của kỳ vọng, ta có

$$E(\bar{X}) = E\left(\frac{1}{n} \sum_{j=1}^n X_j\right) = \frac{1}{n} \sum_{j=1}^n E(X_j) = \mu$$

- ② Điều này không phụ thuộc vào cỡ mẫu và đúng cho mọi quy luật xác suất

# Ước lượng không chệch

## Ví dụ 2

Phương sai mẫu  $S_n^2 = \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})^2$  là một ước lượng chệch của phương sai tổng thể  $\sigma^2$

- 1 Dùng định nghĩa và tính chất của kỳ vọng, ta có

$$E(S_n^2) = E\left(\frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})^2\right) = \frac{n-1}{n} \sigma^2 \neq \sigma^2$$

# Ước lượng không chệch

## Ví dụ 2

Phương sai mẫu  $S_n^2 = \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})^2$  là một ước lượng chệch của phương sai tổng thể  $\sigma^2$

- ❶ Dùng định nghĩa và tính chất của kỳ vọng, ta có

$$E(S_n^2) = E\left(\frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})^2\right) = \frac{n-1}{n} \sigma^2 \neq \sigma^2$$

- ❷ Điều này là lý do cần phải có phương sai mẫu điều chỉnh

$$\hat{S}_n^2 = \frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X})^2 = \frac{n}{n-1} S_n^2$$

# Ước lượng không chệch

## Ví dụ 3

Phương sai mẫu điều chỉnh  $\hat{S}_n^2 = \frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X})^2$  là một ước lượng không chệch của phương sai tổng thể  $\sigma^2$

- 1 Dùng định nghĩa và tính chất của kỳ vọng, ta có

$$E(\hat{S}_n^2) = E\left(\frac{n-1}{n} \cdot \frac{n}{n-1} \sigma^2\right) = \sigma^2$$

# Ước lượng không chệch

## Ví dụ 3

Phương sai mẫu điều chỉnh  $\hat{S}_n^2 = \frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X})^2$  là một ước lượng không chệch của phương sai tổng thể  $\sigma^2$

- ① Dùng định nghĩa và tính chất của kỳ vọng, ta có

$$E(\hat{S}_n^2) = E\left(\frac{n-1}{n} \cdot \frac{n}{n-1} \sigma^2\right) = \sigma^2$$

- ② Khi  $n$  lớn thì phương sai mẫu là ước lượng không chệch tiệm cận của tham số  $\sigma^2$ , vì

$$\lim_{n \rightarrow \infty} E(\hat{S}_n^2) = \lim_{n \rightarrow \infty} \frac{n-1}{n} \sigma^2 = \sigma^2$$

# Ước lượng không chệch

Giả sử  $X \sim B_n(p)$ ,  $k$  là số phép thử thành công của  $n$  phép thử độc lập Bernoulli. Khi đó,

## Ví dụ 4

Tần suất mẫu  $f_n = \frac{k}{n}$  là một ước lượng không chệch của tham số  $p$

① Dùng định nghĩa và tính chất của kỳ vọng, ta có

$$E\left(\frac{k}{n}\right) = \frac{E(k)}{n} = \frac{np}{n} = p$$



# Ước lượng không chệch

Giả sử  $X \sim B_n(p)$ ,  $k$  là số phép thử thành công của  $n$  phép thử độc lập Bernoulli. Khi đó,

## Ví dụ 4

Tần suất mẫu  $f_n = \frac{k}{n}$  là một ước lượng không chệch của tham số  $p$

- ① Dùng định nghĩa và tính chất của kỳ vọng, ta có

$$E\left(\frac{k}{n}\right) = \frac{E(k)}{n} = \frac{np}{n} = p$$

- ② Khi  $n$  lớn thì tần suất mẫu  $f_n \approx p$ . Đây là bản chất của thống kê (Luật yếu các số lớn)

# Ước lượng vững

## Định nghĩa

Thống kê  $\hat{\theta}$  được gọi là một ước lượng vững của tham số  $\theta$ , nếu

$$\forall \epsilon > 0, \lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| > \epsilon) = 0$$

- 1 Là kết quả của Luật yếu số lớn

# Ước lượng vững

## Định nghĩa

Thống kê  $\hat{\theta}$  được gọi là một ước lượng vững của tham số  $\theta$ , nếu

$$\forall \epsilon > 0, \lim_{n \rightarrow \infty} P\left(|\hat{\theta} - \theta| > \epsilon\right) = 0$$

- 1 Là kết quả của Luật yếu số lớn
- 2 Bản chất: khi  $n$  lớn, thì  $\hat{\theta} \approx \theta$

# Ước lượng vững

## Định lý 1

Thống kê  $\hat{\theta}$  là một ước lượng vững của tham số  $\theta$ , nếu

- 1 Thống kê  $\hat{\theta}$  là ước lượng không chệch của tham số  $\theta$

# Ước lượng vững

## Định lý 1

Thống kê  $\hat{\theta}$  là một ước lượng vững của tham số  $\theta$ , nếu

- 1 Thống kê  $\hat{\theta}$  là ước lượng không chệch của tham số  $\theta$
- 2  $\lim_{n \rightarrow \infty} D(\hat{\theta}) = 0$

# Chứng minh

- ① Xét  $\forall \epsilon > 0$ , do  $\hat{\theta}$  là một ước lượng không chệch của tham số  $\theta$ ,

$$\begin{aligned} 0 < P\left( |\hat{\theta} - \theta| \geq \epsilon \right) &= P\left( |\hat{\theta} - E(\hat{\theta})| \geq \epsilon \right) \leq \\ &\leq \frac{1}{\epsilon^2} D(\hat{\theta}) \end{aligned}$$

# Chứng minh

- ① Xét  $\forall \epsilon > 0$ , do  $\hat{\theta}$  là một ước lượng không chệch của tham số  $\theta$ ,

$$\begin{aligned} 0 < P\left(|\hat{\theta} - \theta|\right) &= P\left(|\hat{\theta} - E(\hat{\theta})| \geq \epsilon\right) \leq \\ &\leq \frac{1}{\epsilon^2} D(\hat{\theta}) \end{aligned}$$

- ② Nếu  $D(\hat{\theta}) \rightarrow 0$  khi  $n \rightarrow \infty$ , ta có điều phải chứng minh.

# Ước lượng vững

## Ví dụ 1

Trung bình mẫu  $\bar{X}$  là một ước lượng vững của tham số  $\mu$ , vì

- 1 Trung bình mẫu  $\bar{X}$  là ước lượng không chệch của tham số  $\mu$



# Ước lượng vững

## Ví dụ 1

Trung bình mẫu  $\bar{X}$  là một ước lượng vững của tham số  $\mu$ , vì

- 1 Trung bình mẫu  $\bar{X}$  là ước lượng không chệch của tham số  $\mu$
- 2  $\lim_{n \rightarrow \infty} D(\bar{X}) = \frac{\sigma^2}{n} = 0$

# Ước lượng vững

## Ví dụ 2

Phương sai mẫu  $S_n^2$  không là một ước lượng vững của tham số  $\sigma^2$ , vì phương sai mẫu  $\overline{X}$  không phải là ước lượng không chệch của tham số  $\sigma^2$

# Ước lượng hiệu quả

## Định nghĩa

Thống kê  $\hat{\theta}$  được gọi là một ước lượng hiệu quả của tham số  $\theta$ , nếu

- 1 Thống kê  $\hat{\theta}$  là một ước lượng không chệch của tham số  $\theta$

# Ước lượng hiệu quả

## Định nghĩa

Thống kê  $\hat{\theta}$  được gọi là một ước lượng hiệu quả của tham số  $\theta$ , nếu

- 1 Thống kê  $\hat{\theta}$  là một ước lượng không chệch của tham số  $\theta$
- 2 Phương sai  $D(\hat{\theta}) \leq D(\bar{\theta})$ , với  $\bar{\theta}$  là một ước lượng của tham số  $\theta$

# Ước lượng hiệu quả

## Định nghĩa

Thống kê  $\hat{\theta}$  còn được gọi là một ước lượng không chệch với phương sai bé nhất của tham số  $\theta$

# Chú ý

- ① Điều kiện để thống kê  $\hat{\theta}$  là một ước lượng không chệch của tham số  $\theta$  có thể dễ dàng kiểm tra

# Chú ý

- ① Điều kiện để thống kê  $\hat{\theta}$  là một ước lượng không chệch của tham số  $\theta$  có thể dễ dàng kiểm tra
- ② Khó khăn trong việc xác định sự bé nhất của phương sai  $D(\hat{\theta})$  so với các phương sai của các ước lượng không chệch khác của tham số  $\theta$

# Chú ý

- ① Điều kiện để thống kê  $\hat{\theta}$  là một ước lượng không chệch của tham số  $\theta$  có thể dễ dàng kiểm tra
- ② Khó khăn trong việc xác định sự bé nhất của phương sai  $D(\hat{\theta})$  so với các phương sai của các ước lượng không chệch khác của tham số  $\theta$
- ③ Đó là nguyên nhân phải sử dụng bất đẳng thức thông tin của Crame-Rao



# Bất đẳng thức thông tin Crame-Rao

## Bất đẳng thức

Với mọi ước lượng không chệch  $\hat{\theta}$  của tham số  $\theta$ , có

$$D(\hat{\theta}) \geq \frac{1}{nI_n(\theta)}$$

- ① Đại lượng  $I_n(\theta)$  là đơn vị thông tin Fisher, xác định bởi

$$I_n(\theta) = E\left(\frac{d}{d\theta} \ln p(x, \theta)\right)^2$$

# Bất đẳng thức thông tin Crame-Rao

## Bất đẳng thức

Với mọi ước lượng không chệch  $\hat{\theta}$  của tham số  $\theta$ , có

$$D(\hat{\theta}) \geq \frac{1}{nI_n(\theta)}$$

- ① Đại lượng  $I_n(\theta)$  là đơn vị thông tin Fisher, xác định bởi

$$I_n(\theta) = E\left(\frac{d}{d\theta} \ln p(x, \theta)\right)^2$$

- ② Bất đẳng thức đúng với mọi ước lượng không chệch của tham số  $\theta$

# Ước lượng hiệu quả

## Kết luận

Thống kê  $\hat{\theta}$  là một ước lượng không chệch với phương sai bé nhất (hiệu quả) của tham số  $\theta$ , nếu

$$D(\hat{\theta}) = \frac{1}{nI_n(\theta)}$$

# Ước lượng hiệu quả

## Ví dụ 1

Giả sử  $X \sim \mathcal{N}(\mu, \sigma^2)$ . Khi đó, trung bình mẫu  $\bar{X}$  là một ước lượng hiệu quả của tham số  $\mu$

# Lời giải

- 1 Dễ dàng thấy rằng trung bình mẫu  $\bar{X}$  là một ước lượng không chệch của tham số  $\mu$

# Lời giải

- 1 Dễ dàng thấy rằng trung bình mẫu  $\bar{X}$  là một ước lượng không chệch của tham số  $\mu$
- 2 Mặt khác, ta có  $D(\bar{X}) = \frac{\sigma^2}{n}$

# Lời giải

- 1 Dễ dàng thấy rằng trung bình mẫu  $\bar{X}$  là một ước lượng không chệch của tham số  $\mu$
- 2 Mặt khác, ta có  $D(\bar{X}) = \frac{\sigma^2}{n}$
- 3 Lời khẳng định là đúng, nếu

$$I_n(\mu) = \frac{1}{\sigma^2}$$

# Lời giải

- 1 Dễ dàng thấy rằng trung bình mẫu  $\bar{X}$  là một ước lượng không chệch của tham số  $\mu$
- 2 Mặt khác, ta có  $D(\bar{X}) = \frac{\sigma^2}{n}$
- 3 Lời khẳng định là đúng, nếu

$$I_n(\mu) = \frac{1}{\sigma^2}$$

- 4 Xét đơn vị thông tin  $I_n(\mu)$  của biến ngẫu nhiên  $X \sim \mathcal{N}(\mu, \sigma^2)$



# Lời giải

- ❶ Biến  $X$  có hàm mật độ

$$p(x, \mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

# Lời giải

- 1 Biến  $X$  có hàm mật độ

$$p(x, \mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

- 2 Khi đó,  $\ln p(x, \mu) = \ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$

# Lời giải

- ❶ Biến  $X$  có hàm mật độ

$$p(x, \mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

- ❷ Khi đó,  $\ln p(x, \mu) = \ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$

- ❸ Lấy đạo hàm riêng theo  $\mu$ , ta có

$$\frac{d}{d\mu} \ln p(x, \mu) = \frac{x - \mu}{\sigma^2}$$

# Lời giải

- ❶ Biến  $X$  có hàm mật độ

$$p(x, \mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

- ❷ Khi đó,  $\ln p(x, \mu) = \ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$

- ❸ Lấy đạo hàm riêng theo  $\mu$ , ta có

$$\frac{d}{d\mu} \ln p(x, \mu) = \frac{x - \mu}{\sigma^2}$$

- ❹ Suy ra,

$$I_n(\mu) = E\left(\frac{d}{dx} \ln p(x, \mu)\right)^2 = E\left(\frac{x - \mu}{\sigma^2}\right)^2 = \frac{1}{\sigma^2}$$

# Ước lượng hợp lý cực đại

## Định nghĩa 1

Hàm  $L(X_1, X_2, \dots, X_n, \theta) = \prod_{j=1}^n p_{X_j}(x, \theta)$  được gọi là hàm hợp lý của tham số  $\theta$

## Định nghĩa 2

Thông kê  $\hat{\theta}$  được gọi là một ước lượng hợp lý cực đại của tham số  $\theta$ , nếu hàm hợp lý  $L(X_1, X_2, \dots, X_n, \theta)$  đạt giá trị cực đại (địa phương) tại điểm  $\theta$

# Thuật toán Fisher tìm ước lượng hợp lý cực đại

## Thuật toán

- 1 Xác định hàm hợp lý  $L(X_1, X_2, \dots, X_n, \theta)$  của tham số  $\theta$
- 2 Giải phương trình

$$\frac{d}{d\theta} \ln L(X_1, X_2, \dots, X_n, \theta) = 0$$

tìm điểm có khả năng đạt cực trị  $\hat{\theta}$

- 3 Kiểm tra điều kiện

$$\frac{d^2}{d\theta^2} \ln L(X_1, X_2, \dots, X_n, \hat{\theta}) < 0$$

- 4 Nếu điều kiện thỏa mãn, thì thống kê  $\hat{\theta}$  là ước lượng hợp lý cực đại của tham số  $\theta$

# Giải thích

- 1 Dựa vào định lý Fermat: nếu hàm số  $y = f(x)$  có cực trị tại điểm  $x_0$ , thì  $\frac{d}{dx}f(x_0) = 0$ . (Điều ngược lại không đúng)
- 2 Dựa vào định lý về cực trị: nếu hàm số  $y = f(x)$  có cực trị tại điểm  $x_0$  và  $f(x_0) < 0$ , thì điểm  $x_0$  là điểm cực đại của hàm số  $f$
- 3 Do hàm số  $f > 0$  nên hàm số  $f$  và  $\ln(f)$  có cùng cực trị.
- 4 Vì vậy mà Fisher đã xét điểm cực đại của hàm  $\ln L(X_1, X_2, \dots, X_n, \theta)$  thay cho hàm  $L(X_1, X_2, \dots, X_n, \theta) > 0$ .

# Ước lượng hợp lý cực đại

## Ví dụ 1

Giả sử  $X \sim \mathcal{N}(\mu, \sigma^2)$ . Khi đó, trung bình mẫu  $\bar{X}$  là một ước lượng hợp lý cực đại của tham số  $\mu$



# Lời giải

- ① Xác định hàm hợp lý của tham số  $\mu$  là

$$L(X_1, X_2, \dots, X_n, \mu) = \left( \frac{1}{\sigma\sqrt{2\pi}} \right)^n e^{-\frac{1}{2\sigma^2} \sum_{j=1}^n (x_j - \mu)^2}$$

# Lời giải

- ① Xác định hàm hợp lý của tham số  $\mu$  là

$$L(X_1, X_2, \dots, X_n, \mu) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2} \sum_{j=1}^n (x_j - \mu)^2}$$

- ② Giải phương trình

$$\frac{d}{d\mu} \ln L(X_1, X_2, \dots, X_n, \mu) = 0$$

cho nghiệm  $\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$

# Lời giải

- ① Xác định hàm hợp lý của tham số  $\mu$  là

$$L(X_1, X_2, \dots, X_n, \mu) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2} \sum_{j=1}^n (x_j - \mu)^2}$$

- ② Giải phương trình

$$\frac{d}{d\mu} \ln L(X_1, X_2, \dots, X_n, \mu) = 0$$

cho nghiệm  $\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$

- ③ Kiểm tra điều kiện

$$\frac{d^2}{d\mu^2} \ln L(X_1, X_2, \dots, X_n, \hat{\mu}) = -\frac{n}{\sigma^2} < 0$$

# Lời giải

- ① Xác định hàm hợp lý của tham số  $\mu$  là

$$L(X_1, X_2, \dots, X_n, \mu) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2} \sum_{j=1}^n (x_j - \mu)^2}$$

- ② Giải phương trình

$$\frac{d}{d\mu} \ln L(X_1, X_2, \dots, X_n, \mu) = 0$$

cho nghiệm  $\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$

- ③ Kiểm tra điều kiện

$$\frac{d^2}{d\mu^2} \ln L(X_1, X_2, \dots, X_n, \hat{\mu}) = -\frac{n}{\sigma^2} < 0$$

- ④ Kết luận, trung bình mẫu  $\bar{X}$  là ước lượng hợp lý cực đại của tham số  $\mu$

# Ước lượng hợp lý cực đại

## Ví dụ 2

Giả sử  $X \sim \mathcal{N}(\mu, \sigma^2)$ . Khi đó, phương sai mẫu  $S_n^2$  là một ước lượng hợp lý cực đại của tham số  $\sigma^2$

Xin mời các bạn giải ví dụ 2!

# Chú ý

- Trung bình mẫu  $\bar{X}$  là các ước lượng tốt cho tham số  $\mu$  của tổng thể sinh bởi biến ngẫu nhiên  $X$  phân phối chuẩn  $\mathcal{N}(\mu, \sigma^2)$

# Chú ý

- Trung bình mẫu  $\bar{X}$  là các ước lượng tốt cho tham số  $\mu$  của tổng thể sinh bởi biến ngẫu nhiên  $X$  phân phối chuẩn  $\mathcal{N}(\mu, \sigma^2)$ 
  - ① là ước lượng không chệch của tham số  $\mu$

# Chú ý

- Trung bình mẫu  $\bar{X}$  là các ước lượng tốt cho tham số  $\mu$  của tổng thể sinh bởi biến ngẫu nhiên  $X$  phân phối chuẩn  $\mathcal{N}(\mu, \sigma^2)$ 
  - ① là ước lượng không chệch của tham số  $\mu$
  - ② là ước lượng vững của tham số  $\mu$



# Chú ý

- Trung bình mẫu  $\bar{X}$  là các ước lượng tốt cho tham số  $\mu$  của tổng thể sinh bởi biến ngẫu nhiên  $X$  phân phối chuẩn  $\mathcal{N}(\mu, \sigma^2)$ 
  - ① là ước lượng không chệch của tham số  $\mu$
  - ② là ước lượng vững của tham số  $\mu$
  - ③ là ước lượng hợp lý cực đại của tham số  $\mu$

# Chú ý

- Trung bình mẫu  $\bar{X}$  là các ước lượng tốt cho tham số  $\mu$  của tổng thể sinh bởi biến ngẫu nhiên  $X$  phân phối chuẩn  $\mathcal{N}(\mu, \sigma^2)$ 
  - ① là ước lượng không chệch của tham số  $\mu$
  - ② là ước lượng vững của tham số  $\mu$
  - ③ là ước lượng hợp lý cực đại của tham số  $\mu$
  - ④ là ước lượng hiệu quả của tham số  $\mu$

# Chú ý

- Trung bình mẫu  $\bar{X}$  là các ước lượng tốt cho tham số  $\mu$  của tổng thể sinh bởi biến ngẫu nhiên  $X$  phân phối chuẩn  $\mathcal{N}(\mu, \sigma^2)$ 
  - ① là ước lượng không chệch của tham số  $\mu$
  - ② là ước lượng vững của tham số  $\mu$
  - ③ là ước lượng hợp lý cực đại của tham số  $\mu$
  - ④ là ước lượng hiệu quả của tham số  $\mu$
- Vì vậy, trong nhiều bài toán thống kê mà biến ngẫu nhiên  $X$  có phân phối chuẩn với tham số  $\mu$  và  $\sigma^2$ , trung bình mẫu  $\bar{X}$  có thể thay thế cho trung bình tổng thể  $\mu$

# Ước lượng mô men

## Định nghĩa

Thống kê  $\hat{\theta}$  được gọi là ước lượng mô men của tham số  $\theta$ , nếu nó là nghiệm của hệ phương trình

$$\hat{\mu}_j = \mu_j$$

- ở đây,  $\hat{\mu}_j = E_n(X^j) = \frac{1}{n} \sum_{i=1}^n X_i^j$  là mô men mẫu cấp  $j, j = 1, 2, \dots$

# Ước lượng mô men

## Định nghĩa

Thống kê  $\hat{\theta}$  được gọi là ước lượng mô men của tham số  $\theta$ , nếu nó là nghiệm của hệ phương trình

$$\hat{\mu}_j = \mu_j$$

- ở đây,  $\hat{\mu}_j = E_n(X^j) = \frac{1}{n} \sum_{i=1}^n X_i^j$  là mô men mẫu cấp  $j, j = 1, 2, \dots$
- và  $\mu_j = E(X^j) = \sum_{i=1}^n x_i^j p_i$  là mô men tổng thể cấp  $j, j = 1, 2, \dots$

# Ước lượng mô men

## Ví dụ 1

- 1 Trung bình mẫu là ước lượng mô men của trung bình tổng thể  $\mu$
- 2 Phương sai mẫu là ước lượng mô men của phương sai tổng thể  $\sigma^2$

# Lời giải

- Với  $j=1$ , ta có phương trình

$$\hat{\mu}_1 = \mu_1 \iff \bar{X} = \mu = E(X)$$

# Lời giải

- Với  $j=1$ , ta có phương trình

$$\hat{\mu}_1 = \mu_1 \iff \bar{X} = \mu = E(X)$$

- Với  $j=2$ , ta có phương trình

$$\hat{\mu}_2 = \mu_2 \iff S_n^2 + (\bar{X})^2 = E(X^2) = \sigma^2 + (E(X))^2$$



# Lời giải

- Với  $j=1$ , ta có phương trình

$$\hat{\mu}_1 = \mu_1 \iff \bar{X} = \mu = E(X)$$

- Với  $j=2$ , ta có phương trình

$$\hat{\mu}_2 = \mu_2 \iff S_n^2 + (\bar{X})^2 = E(X^2) = \sigma^2 + (E(X))^2$$

- Suy ra,

# Lời giải

- Với  $j=1$ , ta có phương trình

$$\hat{\mu}_1 = \mu_1 \iff \bar{X} = \mu = E(X)$$

- Với  $j=2$ , ta có phương trình

$$\hat{\mu}_2 = \mu_2 \iff S_n^2 + (\bar{X})^2 = E(X^2) = \sigma^2 + (E(X))^2$$

- Suy ra,

- 1 Trung bình mẫu  $\bar{X}$  là ước lượng mô men của trung bình tổng thể  $\mu$ .

# Lời giải

- Với  $j=1$ , ta có phương trình

$$\hat{\mu}_1 = \mu_1 \iff \bar{X} = \mu = E(X)$$

- Với  $j=2$ , ta có phương trình

$$\hat{\mu}_2 = \mu_2 \iff S_n^2 + (\bar{X})^2 = E(X^2) = \sigma^2 + (E(X))^2$$

- Suy ra,

- ① Trung bình mẫu  $\bar{X}$  là ước lượng mô men của trung bình tổng thể  $\mu$ .
- ② Phương sai mẫu  $S_n^2$  là ước lượng mô men của phương sai tổng thể  $\sigma^2$ .

## 6.2 Ước lượng khoảng

- 1 Đặt vấn đề
- 2 Khoảng ước lượng cho trung bình tổng thể
- 3 Khoảng ước lượng cho phương sai tổng thể
- 4 Khoảng ước lượng cho tần suất tổng thể
- 5 Khoảng ước lượng và cỡ mẫu

# Đặt vấn đề

## Bài toán

Giả sử  $X \sim F(x, \theta)$ ,  $\theta$  là tham số chưa biết, cần ước lượng. Từ mẫu  $\omega_n = \{X_1, X_2, \dots, X_n\}$ , xác định hai thống kê  $\hat{\theta}_1$  và  $\hat{\theta}_2$ , sao cho

$$P\left(\hat{\theta}_1 < \theta < \hat{\theta}_2\right) = \gamma$$

- Khoảng  $(\hat{\theta}_1; \hat{\theta}_2)$  được gọi là khoảng ước lượng tham số  $\theta$

# Đặt vấn đề

## Bài toán

Giả sử  $X \sim F(x, \theta)$ ,  $\theta$  là tham số chưa biết, cần ước lượng. Từ mẫu  $\omega_n = \{X_1, X_2, \dots, X_n\}$ , xác định hai thống kê  $\hat{\theta}_1$  và  $\hat{\theta}_2$ , sao cho

$$P\left(\hat{\theta}_1 < \theta < \hat{\theta}_2\right) = \gamma$$

- Khoảng  $\left(\hat{\theta}_1; \hat{\theta}_2\right)$  được gọi là khoảng ước lượng tham số  $\theta$
- Xác suất  $\gamma \in (0, 1)$  được gọi là độ tin cậy cho khoảng ước lượng  $\left(\hat{\theta}_1; \hat{\theta}_2\right)$

# Chú ý

- Khoảng  $(\hat{\theta}_1; \hat{\theta}_2)$  còn được gọi là khoảng ước lượng đối xứng

# Chú ý

- Khoảng  $(\hat{\theta}_1; \hat{\theta}_2)$  còn được gọi là khoảng ước lượng đối xứng
- Xác suất  $\gamma \in (0, 1)$  thường được chọn gần 1 như 0.99, 0.999, 0.95, 0.90



# Chú ý

- Khoảng  $(\hat{\theta}_1; \hat{\theta}_2)$  còn được gọi là khoảng ước lượng đối xứng
- Xác suất  $\gamma \in (0, 1)$  thường được chọn gần 1 như 0.99, 0.999, 0.95, 0.90
- Một số tài liệu gọi  $(\hat{\theta}_1; \hat{\theta}_2)$  với độ tin cậy  $\gamma$  là  $\gamma\%$ — khoảng tin cậy

# Chú ý

- Khoảng  $(\hat{\theta}_1; \hat{\theta}_2)$  còn được gọi là khoảng ước lượng đối xứng
- Xác suất  $\gamma \in (0, 1)$  thường được chọn gần 1 như 0.99, 0.999, 0.95, 0.90
- Một số tài liệu gọi  $(\hat{\theta}_1; \hat{\theta}_2)$  với độ tin cậy  $\gamma$  là  $\gamma\%$ — khoảng tin cậy
- Các thống kê  $\hat{\theta}_1$  và  $\hat{\theta}_2$  thường được chọn nếu chúng là ước lượng tốt cho tham số cần ước lượng  $\theta$

# Khoảng ước lượng cho trung bình tổng thể

## Bài toán 1

Giả sử  $X \sim \mathcal{N}(\mu, \sigma^2)$ ,  $\mu$  là tham số chưa biết, cần ước lượng. Từ mẫu  $\omega_n = \{X_1, X_2, \dots, X_n\}$ , xác định hai thống kê  $\hat{\theta}_1$  và  $\hat{\theta}_2$ , sao cho

$$P\left(\hat{\theta}_1 < \mu < \hat{\theta}_2\right) = \gamma$$

- Giả sử tham số thứ hai  $\sigma^2$  đã biết

# Khoảng ước lượng cho trung bình tổng thể

## Bài toán 1

Giả sử  $X \sim \mathcal{N}(\mu, \sigma^2)$ ,  $\mu$  là tham số chưa biết, cần ước lượng. Từ mẫu  $\omega_n = \{X_1, X_2, \dots, X_n\}$ , xác định hai thống kê  $\hat{\theta}_1$  và  $\hat{\theta}_2$ , sao cho

$$P\left(\hat{\theta}_1 < \mu < \hat{\theta}_2\right) = \gamma$$

- Giả sử tham số thứ hai  $\sigma^2$  đã biết
- Mẫu được chọn với cỡ  $n$  đủ lớn, trong thực tế mẫu có  $n \geq 30$  được coi là mẫu lớn

## Định lý 1

Giả sử  $(X_1, X_2, \dots, X_n)$  là mẫu sinh ra từ  $n$  quan sát độc lập của một biến ngẫu nhiên  $X \sim \mathfrak{N}(\mu, \sigma^2)$ . Khi đó, biến ngẫu nhiên  $Z = \frac{\bar{X} - \mu}{\sigma} \sqrt{n}$  có phân phối tiệm cận chuẩn khi  $n \rightarrow \infty$ , tức là

$$\lim_{n \rightarrow \infty} P\left(\frac{\bar{X} - \mu}{\sigma} \sqrt{n} < x\right) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}y^2} dy$$

# Khoảng ước lượng cho trung bình tổng thể

## Bài toán 1

Khoảng ước lượng cho tham số  $\mu$  với độ tin cậy  $\gamma$  xác định bởi

$$\left( \bar{X} - z_{\frac{\gamma}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\frac{\gamma}{2}} \frac{\sigma}{\sqrt{n}} \right)$$

- Giả sử tham số thứ hai  $\sigma^2$  đã biết

# Khoảng ước lượng cho trung bình tổng thể

## Bài toán 1

Khoảng ước lượng cho tham số  $\mu$  với độ tin cậy  $\gamma$  xác định bởi

$$\left( \bar{X} - z_{\frac{\gamma}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\frac{\gamma}{2}} \frac{\sigma}{\sqrt{n}} \right)$$

- Giả sử tham số thứ hai  $\sigma^2$  đã biết
- Trung bình mẫu  $\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$  được tính từ mẫu  $\omega_n = \{X_1, X_2, \dots, X_n\}$

# Khoảng ước lượng cho trung bình tổng thể

- Phân vị chuẩn mức  $\gamma$  là  $z_{\frac{\gamma}{2}}$  được xác định sao cho

$$\Phi_0(z_{\frac{\gamma}{2}}) = \frac{\gamma}{2}$$



# Khoảng ước lượng cho trung bình tổng thể

- Phân vị chuẩn mức  $\gamma$  là  $z_{\frac{\gamma}{2}}$  được xác định sao cho

$$\Phi_0(z_{\frac{\gamma}{2}}) = \frac{\gamma}{2}$$

- Hàm Laplace

$$\Phi_0(x) = \Phi(x) - \frac{1}{2} = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{1}{2}y^2} dy$$

# Phân vị chuẩn mức $\frac{\gamma}{2}$

① nếu  $\gamma=0.99$  thì  $z_{\frac{\gamma}{2}}=2.58$

# Phân vị chuẩn mức $\frac{\gamma}{2}$

- 1 nếu  $\gamma=0.99$  thì  $z_{\frac{\gamma}{2}}=2.58$
- 2 nếu  $\gamma=0.95$  thì  $z_{\frac{\gamma}{2}}=1.96$

# Phân vị chuẩn mức $\frac{\gamma}{2}$

- ① nếu  $\gamma=0.99$  thì  $z_{\frac{\gamma}{2}}=2.58$
- ② nếu  $\gamma=0.95$  thì  $z_{\frac{\gamma}{2}}=1.96$
- ③ nếu  $\gamma=0.90$  thì  $z_{\frac{\gamma}{2}}=1.65$

# Khoảng ước lượng cho trung bình tổng thể

## Ví dụ 1

Thống kê số phút đi làm muộn của một cơ quan, có kết quả sau X số phút muộn: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9

$n_j$ —tần số: 183, 15, 16, 14, 13, 2, 1, 1, 1, 1

- Cho biết  $X \sim \mathcal{N}(\mu, \sigma = 0.1)$ . Với độ tin cậy 0.95 hãy xác định khoảng ước lượng đối xứng cho số phút đi làm muộn trung bình.

# Khoảng ước lượng cho trung bình tổng thể

## Ví dụ 1

Thống kê số phút đi làm muộn của một cơ quan, có kết quả sau

X số phút muộn:	0	1	2	3	4	5	6	7	8	9
$n_j$ —tần số:	183	15	16	14	13	2	1	1	1	1

- Cho biết  $X \sim \mathcal{N}(\mu, \sigma = 0.1)$ . Với độ tin cậy 0.95 hãy xác định khoảng ước lượng đối xứng cho số phút đi làm muộn trung bình.
- Để sai số ước lượng không vượt quá  $\epsilon = 0.01$  với độ tin cậy 0.99 thì cần quan sát một mẫu có cỡ bao nhiêu?

# Khoảng ước lượng cho trung bình tổng thể

## Bài toán 2

Giả sử  $X \sim \mathcal{N}(\mu, \sigma^2)$ ,  $\mu$  và  $\sigma^2$  là hai tham số chưa biết, cần ước lượng tham số  $\mu$ . Từ mẫu  $\omega_n = \{X_1, X_2, \dots, X_n\}$ , xác định hai thống kê  $\hat{\theta}_1$  và  $\hat{\theta}_2$ , sao cho

$$P(\hat{\theta}_3 < \mu < \hat{\theta}_4) = \gamma$$

- Giả sử tham số thứ hai  $\sigma^2$  chưa biết

# Khoảng ước lượng cho trung bình tổng thể

## Bài toán 2

Giả sử  $X \sim \mathcal{N}(\mu, \sigma^2)$ ,  $\mu$  và  $\sigma^2$  là hai tham số chưa biết, cần ước lượng tham số  $\mu$ . Từ mẫu  $\omega_n = \{X_1, X_2, \dots, X_n\}$ , xác định hai thống kê  $\hat{\theta}_1$  và  $\hat{\theta}_2$ , sao cho

$$P\left(\hat{\theta}_3 < \mu < \hat{\theta}_4\right) = \gamma$$

- Giả sử tham số thứ hai  $\sigma^2$  chưa biết
- Mẫu được chọn với cỡ  $n$  đủ lớn, trong thực tế mẫu có  $n \geq 30$  được coi là mẫu lớn



# Giải thích

- 1 Do tham số thứ hai là  $\sigma^2$  cũng chưa biết, nên phải sử dụng một thống kê là ước lượng tốt cho  $\sigma^2$

# Giải thích

- 1 Do tham số thứ hai là  $\sigma^2$  cũng chưa biết, nên phải sử dụng một thống kê là ước lượng tốt cho  $\sigma^2$
- 2 Chọn thống kê là phương sai mẫu điều chỉnh  $\hat{S}_n^2$  vì nó là ước lượng không chệch của tham số  $\sigma^2$

# Giải thích

- 1 Do tham số thứ hai là  $\sigma^2$  cũng chưa biết, nên phải sử dụng một thống kê là ước lượng tốt cho  $\sigma^2$
- 2 Chọn thống kê là phương sai mẫu điều chỉnh  $\hat{S}_n^2$  vì nó là ước lượng không chệch của tham số  $\sigma^2$
- 3 Như vậy, xét thống kê

$$T = \frac{\bar{X} - \mu}{\hat{S}_n^2} \sqrt{n} \sim T_{n-1}$$

có phân phối Student với  $n-1$  bậc tự do

## Định lý 2

Nếu  $X \in \mathfrak{N}(\mu, \sigma^2)$ , thì biến ngẫu nhiên  $T = \frac{\bar{X} - \mu}{\hat{S}_n} \sqrt{n}$  có phân phối Student với  $(n-1)$  bậc tự do

# Khoảng ước lượng cho trung bình tổng thể

## Bài toán 2

Khoảng ước lượng cho tham số  $\mu$  với độ tin cậy  $\gamma$  xác định bởi

$$\left( \bar{X} - t_{\frac{\gamma}{2}}(n-1) \frac{\hat{S}_n}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\gamma}{2}}(n-1) \frac{\hat{S}_n}{\sqrt{n}} \right)$$

- Giả sử tham số thứ hai  $\sigma^2$  chưa biết, được thay bởi phương sai mẫu điều chỉnh  $S_n^2$

# Khoảng ước lượng cho trung bình tổng thể

## Bài toán 2

Khoảng ước lượng cho tham số  $\mu$  với độ tin cậy  $\gamma$  xác định bởi

$$\left( \bar{X} - t_{\frac{\gamma}{2}}(n-1) \frac{\hat{S}_n}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\gamma}{2}}(n-1) \frac{\hat{S}_n}{\sqrt{n}} \right)$$

- Giả sử tham số thứ hai  $\sigma^2$  chưa biết, được thay bởi phương sai mẫu điều chỉnh  $S_n^2$
- Trung bình mẫu  $\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$  được tính từ mẫu  $\omega_n = \{X_1, X_2, \dots, X_n\}$

## Khoảng ước lượng cho trung bình tổng thể khi chưa biết $\sigma^2$

- Phân vị Student mức  $\gamma$  là  $t_{\frac{\gamma}{2}}(n-1)$  được tra từ bảng thống kê Student ứng với độ tin cậy  $\frac{\gamma}{2}$  và độ tự do  $(n-1)$

# Khoảng ước lượng cho trung bình tổng thể khi chưa biết $\sigma^2$

- Phân vị Student mức  $\gamma$  là  $t_{\frac{\gamma}{2}}(n-1)$  được tra từ bảng thống kê Student ứng với độ tin cậy  $\frac{\gamma}{2}$  và độ tự do  $(n-1)$
- Trường hợp mẫu nhỏ,  $n < 30$  cũng thường sử dụng phân phối Student để ước lượng



# Khoảng ước lượng cho trung bình tổng thể khi chưa biết $\sigma^2$

- Phân vị Student mức  $\gamma$  là  $t_{\frac{\gamma}{2}}(n-1)$  được tra từ bảng thống kê Student ứng với độ tin cậy  $\frac{\gamma}{2}$  và độ tự do  $(n-1)$
- Trường hợp mẫu nhỏ,  $n < 30$  cũng thường sử dụng phân phối Student để ước lượng
- Chú ý khi  $n$  lớn thì các phân vị chuẩn và student trùng nhau (theo định lý tiệm cận chuẩn của phân phối Student), hay

$$t_{\frac{\gamma}{2}}(+\infty) = z_{\frac{\gamma}{2}}$$

# Khoảng ước lượng cho trung bình tổng thể

## Ví dụ 2

Thống kê số phút ( $X$ ) đi vào lớp muộn của sinh viên một trường đại học, có kết quả thống kê sau

$$X_j : 0, \quad 1, \quad 2, \quad 3, \quad 4, \quad 5, \quad 6, \quad 7, \quad 8, \quad 9$$

$$n_j : 103, \quad 25, \quad 36, \quad 4, \quad 3, \quad 2, \quad 1, \quad 1, \quad 1, \quad 1$$

- Cho biết  $X \sim \mathcal{N}(\mu, \sigma^2)$ . Với độ tin cậy 0.95 hãy xác định khoảng ước lượng đối xứng cho số phút vào lớp muộn trung bình của sinh viên.

# Khoảng ước lượng cho trung bình tổng thể

## Ví dụ 2

Thống kê số phút ( $X$ ) đi vào lớp muộn của sinh viên một trường đại học, có kết quả thống kê sau

$$X_j : 0, \quad 1, \quad 2, \quad 3, \quad 4, \quad 5, \quad 6, \quad 7, \quad 8, \quad 9$$

$$n_j : 103, \quad 25, \quad 36, \quad 4, \quad 3, \quad 2, \quad 1, \quad 1, \quad 1, \quad 1$$

- Cho biết  $X \sim \mathcal{N}(\mu, \sigma^2)$ . Với độ tin cậy 0.95 hãy xác định khoảng ước lượng đôi xứng cho số phút vào lớp muộn trung bình của sinh viên.
- Xét trường hợp tần suất  $n_j$  có giá trị 1 cho tất cả giá trị của  $X$

# Khoảng ước lượng cho xác suất tổng thể $p$

## Bài toán 3

Giả sử  $X \sim B_n(p)$ ,  $p \in (0, 1)$ . Từ mẫu  $\omega_n = \{X_1, X_2, \dots, X_n\}$ , xác định hai thống kê  $\hat{\theta}_5$  và  $\hat{\theta}_6$ , sao cho

$$P\left(\hat{\theta}_5 < p < \hat{\theta}_6\right) = \gamma$$

- Mẫu sinh ra từ phép thử Bernoulli, với xác suất thành công  $p \in (0, 1)$

# Khoảng ước lượng cho xác suất tổng thể $p$

## Bài toán 3

Giả sử  $X \sim B_n(p)$ ,  $p \in (0, 1)$ . Từ mẫu  $\omega_n = \{X_1, X_2, \dots, X_n\}$ , xác định hai thống kê  $\hat{\theta}_5$  và  $\hat{\theta}_6$ , sao cho

$$P\left(\hat{\theta}_5 < p < \hat{\theta}_6\right) = \gamma$$

- Mẫu sinh ra từ phép thử Bernoulli, với xác suất thành công  $p \in (0, 1)$
- Sử dụng định lý tiệm cận chuẩn (chương 5)

## Định lý 3

Giả sử tiến hành  $n$  quan sát độc lập, với xác suất thành công của mỗi quan sát là  $p$ ,  $p \in (0, 1)$ . Gọi  $k$  là số quan sát thành công trong  $n$  quan sát,  $0 \leq k \leq n$ . Khi đó, biến ngẫu nhiên  $Z = \frac{\frac{k}{n} - p}{\sqrt{p(1-p)}} \sqrt{n}$  có phân phối tiệm cận chuẩn khi  $n \rightarrow \infty$ , tức là

$$\lim_{n \rightarrow \infty} P\left(\frac{\frac{k}{n} - p}{\sqrt{p(1-p)}} \sqrt{n} < x\right) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}y^2} dy$$

# Khoảng ước lượng cho xác suất tổng thể $p$

## Bài toán 3

Giả sử  $X \sim B_n(p)$ ,  $p \in (0, 1)$ . Khi đó, khoảng ước lượng đối xứng cho xác suất tổng thể  $p$  với độ tin cậy  $\gamma$  xác định bởi

$$\left( f_n - z_{\frac{\gamma}{2}} \sqrt{\frac{f_n(1-f_n)}{n}} < p < f_n + z_{\frac{\gamma}{2}} \sqrt{\frac{f_n(1-f_n)}{n}} \right)$$

- Tần suất mẫu  $f_n = \frac{k}{n}$  là ước lượng tốt cho xác suất tổng thể  $p$

# Khoảng ước lượng cho xác suất tổng thể $p$

## Bài toán 3

Giả sử  $X \sim B_n(p)$ ,  $p \in (0, 1)$ . Khi đó, khoảng ước lượng đối xứng cho xác suất tổng thể  $p$  với độ tin cậy  $\gamma$  xác định bởi

$$\left( f_n - z_{\frac{\gamma}{2}} \sqrt{\frac{f_n(1-f_n)}{n}} < p < f_n + z_{\frac{\gamma}{2}} \sqrt{\frac{f_n(1-f_n)}{n}} \right)$$

- Tần suất mẫu  $f_n = \frac{k}{n}$  là ước lượng tốt cho xác suất tổng thể  $p$
- Phân vị chuẩn  $z_{\frac{\gamma}{2}}$  xác định như trong bài toán 1.



# Khoảng ước lượng cho trung bình tổng thể

## Ví dụ 3

Thăm dò 1500 sinh viên học năm thứ nhất hệ tín chỉ, có 30 sinh viên gặp khó khăn trong phương pháp học. Với độ tin cậy 0.90 hãy ước lượng tỷ lệ sinh viên gặp khó khăn trong phương pháp học trong toàn trường.

- gọi  $p$  là tỷ lệ sinh viên gặp khó khăn trong phương pháp học của toàn trường,  $p \in (0, 1)$ . Tần suất mẫu  $f_{1500} = \frac{30}{1500} = 0.02$  Với  $\gamma = 0.9$  có  $z_{0.45} = 1.65$

# Khoảng ước lượng cho trung bình tổng thể

## Ví dụ 3

Thăm dò 1500 sinh viên học năm thứ nhất hệ tín chỉ, có 30 sinh viên gặp khó khăn trong phương pháp học. Với độ tin cậy 0.90 hãy ước lượng tỷ lệ sinh viên gặp khó khăn trong phương pháp học trong toàn trường.

- gọi  $p$  là tỷ lệ sinh viên gặp khó khăn trong phương pháp học của toàn trường,  $p \in (0, 1)$ . Tần suất mẫu  $f_{1500} = \frac{30}{1500} = 0.02$  Với  $\gamma = 0.9$  có  $z_{0.45} = 1.65$
- Khoảng ước lượng cho  $p$  với độ tin cậy 0.9 là

$$0.02 - 1.65\sqrt{\frac{0.02(1 - 0.02)}{1500}}; 0.02 + 1.65\sqrt{\frac{0.02(1 - 0.02)}{1500}}$$

# Khoảng ước lượng cho phương sai tổng thể $\sigma^2$

## Bài toán 4

Giả sử  $X \sim \mathfrak{N}(\mu, \sigma^2)$ , với tham số cần ước lượng là  $\sigma^2$ . Khi đó, khoảng ước lượng đối xứng cho phương sai tổng thể  $\sigma^2$  với độ tin cậy  $\gamma$  xác định bởi

$$\left( \frac{(n-1)\hat{S}_n^2}{\chi_{\frac{1-\gamma}{2}}^2(n-1)} < \sigma^2 < \frac{(n-1)\hat{S}_n^2}{\chi_{\frac{\gamma}{2}}^2(n-1)} \right)$$

- Cho trước độ tin cậy  $\gamma \in (0, 1)$ , tra bảng  $\chi^2$ , xác định các giá trị  $\chi_{\frac{1-\gamma}{2}}^2(n-1)$  và  $\chi_{\frac{\gamma}{2}}^2(n-1)$  sao cho

$$P\left(\chi^2 < \chi_{\frac{1-\gamma}{2}}^2(n-1)\right) = P\left(\chi^2 > \chi_{\frac{\gamma}{2}}^2(n-1)\right) = \frac{1-\gamma}{2}$$

# Khoảng ước lượng cho phương sai tổng thể $\sigma^2$

## Bài toán 4

Giả sử  $X \sim \mathfrak{N}(\mu, \sigma^2)$ , với tham số cần ước lượng là  $\sigma^2$ . Khi đó, khoảng ước lượng đối xứng cho phương sai tổng thể  $\sigma^2$  với độ tin cậy  $\gamma$  xác định bởi

$$\left( \frac{(n-1)\hat{S}_n^2}{\chi_{\frac{1-\gamma}{2}}^2(n-1)} < \sigma^2 < \frac{(n-1)\hat{S}_n^2}{\chi_{\frac{\gamma}{2}}^2(n-1)} \right)$$

- Cho trước độ tin cậy  $\gamma \in (0, 1)$ , tra bảng  $\chi^2$ , xác định các giá trị  $\chi_{\frac{1-\gamma}{2}}^2(n-1)$  và  $\chi_{\frac{\gamma}{2}}^2(n-1)$  sao cho

$$P\left(\chi^2 < \chi_{\frac{1-\gamma}{2}}^2(n-1)\right) = P\left(\chi^2 > \chi_{\frac{\gamma}{2}}^2(n-1)\right) = \frac{1-\gamma}{2}$$

- Dựa vào định lý chương 3

## Định lý 4

Nếu  $(X_1, X_2, \dots, X_n)$  là một mẫu sinh từ biến ngẫu nhiên có phân phối chuẩn  $\mathfrak{N}(\mu, \sigma^2)$ , thì  $\frac{n-1}{\sigma^2} \hat{S}_n^2$  có phân phối  $\chi_{n-1}^2$  với  $(n-1)$  bậc tự do

# Khoảng ước lượng cho phương sai tổng thể

## Ví dụ 4

Đo chỉ số trên huyết áp của một số người cao tuổi, có thống kê sau

129, 132, 140, 138, 143, 133, 137, 140, 143, 138, 140

- Tính trung bình mẫu  $\bar{X}$  và phương sai mẫu  $S_n^2$

# Khoảng ước lượng cho phương sai tổng thể

## Ví dụ 4

Đo chỉ số trên huyết áp của một số người cao tuổi, có thống kê sau

129, 132, 140, 138, 143, 133, 137, 140, 143, 138, 140

- Tính trung bình mẫu  $\bar{X}$  và phương sai mẫu  $S_n^2$
- Với độ tin cậy 0.95 hãy ước lượng phương sai tổng thể  $\sigma^2$ .

# Bài tập chương 6

## Bài tập 1

Quan sát tuổi thọ  $X$  giờ của một số bóng đèn ta có thống kê sau

$X_j$  :    1000    1200    1400    1600    2000

$n_j$  :    20    17    18    21    22

- Tính trung bình mẫu  $\bar{X}$  và phương sai mẫu  $S_n^2$



# Bài tập chương 6

## Bài tập 1

Quan sát tuổi thọ  $X$  giờ của một số bóng đèn ta có thống kê sau

$X_j$  :    1000    1200    1400    1600    2000

$n_j$  :    20    17    18    21    22

- Tính trung bình mẫu  $\bar{X}$  và phương sai mẫu  $S_n^2$
- Xác định khoảng ước lượng cho tuổi thọ trung bình của các bóng đèn với độ tin cậy 0.9

# Bài tập chương 6

## Bài tập 1

Quan sát tuổi thọ  $X$  giờ của một số bóng đèn ta có thống kê sau

$X_j$  :    1000    1200    1400    1600    2000

$n_j$  :    20    17    18    21    22

- Tính trung bình mẫu  $\bar{X}$  và phương sai mẫu  $S_n^2$
- Xác định khoảng ước lượng cho tuổi thọ trung bình của các bóng đèn với độ tin cậy 0.9
- Để sai số ước lượng không quá 30 giờ với độ tin cậy 0.95 thì cần quan sát bao nhiêu bóng đèn.

# Bài tập chương 6

## Bài tập 2

Để ước lượng số cá có trong một hồ nuôi, người ta bắt lên 200 con cá, đánh dấu và thả vào lại trong hồ. Một thời gian sau bắt lên 20 con thì thấy có 5 con bị đánh dấu. Với độ tin cậy 0.99 hãy ước lượng số cá có trong hồ

## Bài tập chương 6

### Bài tập 3

Thăm dò 1200 người thì có 768 người ủng hộ ứng cử viên A. Với độ tin cậy 0.90 hãy ước lượng số cử tri ủng hộ ứng cử viên A nếu biết rằng trong khu vực đó có 5000 người