ƯỚC LƯỢNG ĐIỂM

Nguyễn Văn Thìn

BỘ MÔN THỐNG KÊ TOÁN HỌC
KHOA TOÁN - TIN HỌC
ĐẠI HỌC KHOA HỌC TỰ NHIÊN TP.HCM

Tháng 2 năm 2016

Outline

Outline

Giới thiệu

- Giả sử cần khảo sát một đặc tính X thuộc một tổng thể xác định.
- Biến ngẫu nhiên X có phân phối $F(x;\theta)$, trong đó tham số θ chưa biết.
- Bài toán: tìm tham số θ.
  - Chọn một mẫu ngẫu nhiên cỡ n từ X: $X_1, X_2, \ldots, X_n$.
  - Thống kê $\hat{\Theta} = h(X_1, \ldots, X_n)$ gọi là một ước lượng điểm (point estimator) cho θ.
  - Với một mẫu thực nghiệm $x_1, \ldots, x_n$, ta gọi $\hat{\theta} = h(x_1, \ldots, x_n)$ là một giá trị ước lượng điểm (point estimate) cho θ.

Ví dụ

- X = Chiều cao dân số trong một khu vực, $X \sim \mathcal{N}(\mu, \sigma^2)$. Phân phối của X phụ thuộc vào kỳ vọng μ và phương sai σ^2. Thống kê trung bình mẫu và phương sai mẫu

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

là những ước lượng điểm cho μ và σ^2.

- Với một mẫu thực nghiệm $x_1 = 150, x_2 = 155, x_3 = 167$, giá trị ước lượng điểm của μ và σ^2 là $\bar{x} = 157.333$, $s^2 = 76.333$.

Outline

Ước lượng không chệch

Định nghĩa 1

Ước lượng điểm $\hat{\Theta}$ gọi là một ước lượng không chệch (Unbiased estimator) cho tham số θ nếu

$$\mathbb{E}(\hat{\Theta}) = \theta \quad (1)$$

Nếu $\hat{\Theta}$ là ước lượng chệch của θ, độ sai khác

$$\mathbb{E}(\hat{\Theta}) - \theta \quad (2)$$

gọi là độ chệch của ước lượng, ký hiệu $\text{Bias}(\hat{\Theta})$.

Ước lượng không chệch - Ví dụ

1. $\bar{X}$ là một ước lượng không chệch của μ

$$\mathbb{E}(\bar{X}) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{\sum_{i=1}^{n}\mathbb{E}X_i}{n} = \frac{n\mu}{n} = \mu$$

2. S^2 là một ước lượng không chệch của σ^2,

$$\mathbb{E}(S^2) = \sigma^2 \quad \text{(Tại sao?)}$$

3. $\hat{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ là ước lượng chệch của σ^2 vì

$$\mathbb{E}(\hat{S}^2) \neq \sigma^2$$

Ước lượng hiệu quả

Định nghĩa 2

Xét $\hat{\Theta}$ và $\tilde{\Theta}$ là hai ước lượng không chệch của θ, $\hat{\Theta}$ gọi là ước lượng hiệu quả (Efficiency estimator) hơn $\tilde{\Theta}$ nếu với một cỡ mẫu n cho trước

$$\mathbb{V}\text{ar}(\hat{\Theta}) < \mathbb{V}\text{ar}(\tilde{\Theta}) \tag{3}$$

Định nghĩa 3

Xét tất cả các ước lượng không chệch của θ. Ước lượng nào có phương sai bé nhất được gọi là ước lượng không chệch phương sai bé nhất (MVUE).

Định lí 4

Trong một mẫu ngẫu nhiên cỡ n: $X_1, \ldots, X_n$ được chọn từ $X \sim \mathcal{N}(\mu, \sigma^2)$ thì $\bar{X}$ là ước lượng hiệu quả nhất cho μ.

Trung bình của bình phương sai số

- Trong một số trường hợp, ước lượng $\tilde{\Theta}$ là ước lượng chệch (với độ chệch nhỏ), nhưng lại có phương sai nhỏ hơn các ước lượng không chệch $\hat{\Theta}$ khác. Khi đó, ta có thể muốn chọn $\tilde{\Theta}$, mặc dù là ước lượng chệch nhưng nó có độ phân tán nhỏ hơn nhiều so với các ước lượng $\hat{\Theta}$ khác.
- Một độ đo kết hợp giữa độ chệch (Bias) và phương sai mẫu của một ước lượng là trung bình của bình phương sai số (Mean Squared Error - MSE)

$$\text{MSE}(\hat{\Theta}) = \mathbb{E}(\hat{\Theta} - \theta)^2 \tag{4}$$

- $\text{MSE}(\hat{\Theta}) = \mathbb{V}\text{ar}(\hat{\Theta}) + (\text{Bias}(\hat{\Theta}))^2$.
- Nếu $\hat{\Theta}$ là ƯLKC: $\text{MSE}(\hat{\Theta}) = \mathbb{V}\text{ar}(\hat{\Theta})$.

Trung bình của bình phương sai số

• Cho trước hai ước lượng, $\hat{\Theta}$ và $\tilde{\Theta}$, *tiêu chuẩn MSE* cho phép ta chọn $\tilde{\Theta}$ nếu, với cỡ mẫu n

$$\text{MSE}(\tilde{\Theta}) < \text{MSE}(\hat{\Theta})$$

• hoặc $\mathbb{V}\text{ar}(\hat{\Theta}) - \mathbb{V}\text{ar}(\tilde{\Theta}) > (\text{Bias}(\tilde{\Theta}))^2 - (\text{Bias}(\hat{\Theta}))^2$.

• Nếu cả $\hat{\Theta}$ và $\tilde{\Theta}$ là ƯLKC, *tiêu chuẩn MSE* trở thành tiêu chuẩn so sánh dựa trên phương sai mẫu.

• Tiêu chuẩn MSE tương đương với việc so sánh tỷ số

$$\text{Eff}(\hat{\Theta}, \tilde{\Theta}) = \frac{\text{MSE}(\tilde{\Theta})}{\text{MSE}(\hat{\Theta})} \tag{5}$$

và chọn $\tilde{\Theta}$ nếu $\text{Eff}(\hat{\Theta}, \tilde{\Theta}) < 1$.

Sai số chuẩn

Định nghĩa 5

Sai số chuẩn (standard error) của một ước lượng $\hat{\Theta}$ chính là độ lệch tiêu chuẩn của nó, cho bởi

$$\sigma_{\hat{\Theta}} = \sqrt{\mathbb{V}\text{ar}(\hat{\Theta})} \tag{6}$$

Nếu sai số chuẩn chứa các tham số chưa biết có thể được ước lượng, thì khi thay các giá trị đó vào $\sigma_{\hat{\Theta}}$ ta được **sai số chuẩn ước lượng**, kí hiệu là $\hat{\sigma}_{\hat{\Theta}}$.
Đôi khi sai số chuẩn ước lượng được kí hiệu là $s_{\hat{\Theta}}$ hoặc $se(\hat{\Theta})$.

Sai số chuẩn
Một số ví dụ

Tham số	Ước lượng T	$\mathbb{V}ar(T)$	$se(T)$
μ	$\bar{X}$	$\frac{\sigma^2}{n}$	$\frac{S}{\sqrt{n}}$
p	$\hat{p}$	$\frac{p(1-p)}{n}$	$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
σ^2	S^2	$\frac{2\sigma^4}{n-1}$	$S^2\sqrt{\frac{2}{n-1}}$

- Phương sai của S^2 ở trên chỉ đúng trong trường hợp tổng thể có phân phối chuẩn $N(\mu, \sigma^2)$.

Ước lượng vững

ƯỚC LƯỢNG ĐIỂM
Nguyễn Văn Thìn
Ước lượng điểm
Các tiêu chuẩn Ước lượng
Các phương pháp ước lượng điểm

Định nghĩa 6

Gọi $\hat{\Theta}_n = h(X_1, \ldots, X_n)$ là một ước lượng điểm của tham số θ. Ước lượng $\hat{\Theta}_n$ gọi là bền vững (consistency) nếu $\hat{\Theta}_n \xrightarrow{\mathbb{P}} \theta$, tức là

$$\lim_{n\to\infty} \mathbb{P}\left(|\hat{\Theta}_n - \theta| > \epsilon\right) = 0, \quad \forall \epsilon > 0$$

Ví dụ 7

1. S^2 là ước lượng vững của σ^2.
   Thật vậy, theo BĐT Chebyshev,
   $P(|S^2 - \sigma^2| > \epsilon) \leq \frac{Var(S^2)}{\epsilon^2} = \frac{2\sigma^4}{(n-1)\epsilon^2} \xrightarrow{n\to\infty} 0.$
2. Với $X \sim B(n, p)$, $\hat{p} = \frac{X}{n}$ là ước lượng vững cho p. (Tại sao?)

Outline

ƯỚC LƯỢNG ĐIỂM
Nguyễn Văn Thìn
Ước lượng điểm
Các tiêu chuẩn Ước lượng
Các phương pháp ước lượng điểm

Phương pháp Moment

ƯỚC LƯỢNG ĐIỂM
Nguyễn Văn Thìn
Ước lượng điểm
Các tiêu chuẩn Ước lượng
Các phương pháp ước lượng điểm

- Ý tưởng: đồng nhất các moment của tổng thể với các moment mẫu.

Định nghĩa 8

Giả sử tham số $\theta = (\theta_1, \ldots, \theta_k)$ có k thành phần. Với $1 \leq j \leq k$, moment thứ j của tổng thể là

$$\mu'_j = \mathbb{E}(X^j) = \int x^j f(x)dx \tag{7}$$

và moment mẫu thứ j

$$m_j = \frac{1}{n}\sum_{i=1}^{n} X_i^j \tag{8}$$

Phương pháp Moment

Định nghĩa 9

Xét $X_1, X_2, \ldots, X_n$ là một mẫu ngẫu nhiên được chọn từ một phân phối xác suất với k tham số $\theta = (\theta_1, \ldots, \theta_k)$ chưa biết. Ước lượng điểm moment $\hat{\Theta} = (\hat{\Theta}_1, \ldots, \hat{\Theta}_k)$ thu được bởi đồng nhất k moment tổng thể với k moment mẫu và giải hệ phương trình thu được với các tham số chưa biết

$$\begin{aligned} m_1 &= \mu_1'(\theta_1, \ldots, \theta_k) \\ m_2 &= \mu_2'(\theta_1, \ldots, \theta_k) \\ &\vdots \\ m_k &= \mu_k'(\theta_1, \ldots, \theta_k) \end{aligned}$$

Phương pháp Moment - Ví dụ

ƯỚC LƯỢNG ĐIỂM
Nguyễn Văn Thìn
Ước lượng điểm
Các tiêu chuẩn Ước lượng
Các phương pháp ước lượng điểm

1. Giả sử $X_1, \ldots, X_n$ là mẫu ngẫu nhiên chọn từ $X \sim \mathcal{N}(\mu, \sigma^2)$. Tìm các ước lượng moment cho μ và σ^2.

   Moment tổng thể: $\mu_1' = \mu$, $\mu_2' = \mathbb{E}(X^2) = \mu^2 + \sigma^2$.

   Moment mẫu: $m_1 = \frac{1}{n}\sum_{i=1}^n X_i$, $m_2 = \frac{1}{n}\sum_{i=1}^n X_i^2$.

   Giải hệ:

$$\mu = \frac{1}{n}\sum_{i=1}^n X_i, \quad \mu^2 + \sigma^2 = \frac{1}{n}\sum_{i=1}^n X_i^2$$

   Ta thu được ước lượng moment cho μ và σ^2 là

$$\hat{\mu} = \bar{X}, \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$$

Phương pháp Moment - Ví dụ

ƯỚC LƯỢNG ĐIỂM
Nguyễn Văn Thìn
Ước lượng điểm
Các tiêu chuẩn Ước lượng
Các phương pháp ước lượng điểm

2. Với $X \sim B(k, p)$, tìm ước lượng moment cho các tham số k và p.
3. Với $X \sim Gamma(r, \lambda)$, tìm ước lượng moment cho các tham số r và λ biết

$$\mathbb{E}(X) = \frac{r}{\lambda}, \quad \mathbb{V}\text{ar}(X) = \frac{r}{\lambda^2}$$

Phương pháp hợp lý cực đại (Maximum Likelihood)

ƯỚC LƯỢNG ĐIỂM
Nguyễn Văn Thìn
Ước lượng điểm
Các tiêu chuẩn Ước lượng
Các phương pháp ước lượng điểm

Định nghĩa 10 (Hàm hợp lý)

Xét $X_1, \ldots, X_n$ là một mẫu ngẫu nhiên chọn từ tổng thể có hàm mật độ xác suất (hay hàm xác suất) $f(x|\theta)$, với $\theta = (\theta_1, \ldots, \theta_k) \in \Theta$ chưa biết. Hàm hợp lý $L(\theta)$ được định nghĩa bởi

$$L(\theta|\mathbf{x}) = L(\theta_1, \ldots, \theta_k | x_1, \ldots, x_n) = \prod_{i=1}^n f(x_i|\theta_1, \ldots, \theta_k) \qquad (9)$$

- Hàm hợp lý $L(\theta)$ chỉ phụ thuộc vào tham số θ.

Phương pháp hợp lý cực đại (Maximum Likelihood)

Định nghĩa 11

Ước lượng hợp lý cực đại (Maximum Likelihood Estimation - MLE) $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_k)$ là giá trị của $\theta = (\theta_1, \ldots, \theta_k) \in \Theta$ làm cực đại hàm hợp lý.

Nhận xét 12

Bởi vì hàm log là một hàm tăng nên MLE cũng là những giá trị của $\theta = (\theta_1, \ldots, \theta_k) \in \Theta$ làm cực đại hàm log $L(\theta_1, \ldots, \theta_k)$.

Phương pháp hợp lý cực đại - Ví dụ

ƯỚC LƯỢNG ĐIỂM
Nguyễn Văn Thìn
Ước lượng điểm
Các tiêu chuẩn Ước lượng
Các phương pháp ước lượng điểm

1. Xét $X \sim B(1, p)$. Hàm xác suất của X là

$$f(x; p) = \begin{cases} p^x(1-p)^{1-x} & , \quad x = 0, 1 \\ 0 & , \quad \text{nơi khác} \end{cases}$$

Tìm ước lượng hợp lý cực đại cho tham số p.
Hàm hợp lý cho mẫu ngẫu nhiên cỡ n là

$$L(p) = \prod_{i=1}^{n} f(X_i; p) = \prod_{i=1}^{n} p^{X_i}(1-p)^{1-X_i} = p^S(1-p)^{n-S}$$

với $S = \sum_{i=1}^{n} X_i$.
Lấy logarit hàm hợp lý MLE không thay đổi, do đó

$$\ln L(p) = S \ln p + (n - S)\ln(1 - p)$$

Phương pháp hợp lý cực đại - Ví dụ

ƯỚC LƯỢNG ĐIỂM
Nguyễn Văn Thìn
Ước lượng điểm
Các tiêu chuẩn Ước lượng
Các phương pháp ước lượng điểm

Ta có

$$\frac{d \ln L(p)}{dp} = \frac{S}{p} - \frac{(n-S)}{1-p}$$

Giải phương trình trên ta thu được MLE của p là

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

2. Cho $X \sim Exp(\lambda)$, tìm ước lượng hợp lý cực đại cho tham số λ.

3. Cho $X \sim \mathcal{N}(\mu, \sigma^2)$, tìm ước lượng hợp lý cực đại cho kỳ vọng μ và phương sai σ^2.

Tính chất của Ước lượng hợp lý cực đại (MLE)

ƯỚC LƯỢNG ĐIỂM
Nguyễn Văn Thìn
Ước lượng điểm
Các tiêu chuẩn Ước lượng
Các phương pháp ước lượng điểm

- Gọi $\hat{\Theta}_n$ là ước lượng hợp lý cực đại (MLE) của tham số θ, ta có các tính chất sau

1. MLE là ước lượng bền vững: $\hat{\Theta}_n \xrightarrow{\mathbb{P}} \theta$.
2. MLE là ước lượng bất biến: Nếu $\hat{\Theta}_n$ là MLE của θ thì $g(\hat{\Theta}_n)$ là MLE của $g(\theta)$.
3. MLE hội tụ chuẩn:

$$\frac{\hat{\Theta}_n - \theta}{SE(\hat{\Theta}_n)} \rightsquigarrow \mathcal{N}(0, 1)$$

4. MLE là ước lượng hội tụ tối ưu: nghĩa là, trong số tất cả các ước lượng tốt, MLE là ước lượng có phương sai bé nhất, ít nhất là đối với trường hợp mẫu lớn.

Phương pháp Bayes [1]

- Dựa trên thông tin về phân phối của tham số θ.
- Giả sử tham số θ chưa biết là biến ngẫu nhiên có phân phối xác định với hàm mật độ xác suất $f(\theta)$.
- $f(\theta)$ gọi là hàm mật độ xác suất tiên nghiệm (prior probability distribution function).
- Hàm mật độ xác suất đồng thời của mẫu ngẫu nhiên $(X_1, \ldots, X_n)$ và θ được biểu diễn như sau

$$f(x_1, \ldots, x_n, \theta) = f(x_1, \ldots, x_n|\theta)f(\theta) \tag{10}$$

[1] Đây là phần đọc thêm

Phương pháp Bayes [1]

- Hàm mật độ lề của mẫu cho bởi

$$f(x_1, \ldots, x_n) = \int_{R_\Theta} f(x_1, \ldots, x_n, \theta)d\theta \tag{11}$$

  với R_Θ là không gian tham số (các giá trị có thể nhận được của θ).
- Hàm mật độ xác suất có điều kiện của θ cho trước $(x_1, \ldots, x_n)$

$$f(\theta|x_1, \ldots, x_n) = \frac{f(x_1, \ldots, x_n|\theta)f(\theta)}{f(x_1, \ldots, x_n)} \tag{12}$$

[1] Đây là phần đọc thêm

Phương pháp Bayes [1]

- $f(\theta|x_1, \ldots, x_n)$ gọi là hàm mật độ xác suất hậu nghiệm của θ (posterior pdf).
- Hàm mật độ xác suất tiên nghiệm $f(\theta)$ cho biết thông tin về θ khi thực hiện quan trắc để lấy mẫu ngẫu nhiên $(X_1, \ldots, X_n)$.
- Hàm mật độ xác suất hậu nghiệm $f(\theta|x_1, \ldots, x_n)$ cho biết thông tin về θ sau khi có mẫu cụ thể.

[1] Đây là phần đọc thêm

Phương pháp Bayes [1]

- Kỳ vọng có điều kiện của θ, định nghĩa bởi

$$\hat{\theta} = \mathbb{E}(\theta|x_1, \ldots, x_n) = \int_{R_\Theta} \theta f(\theta|x_1, \ldots, x_n)d\theta \tag{13}$$

  là một giá trị ước lượng Bayes của θ, và

$$\hat{\Theta} = \mathbb{E}(\theta|X_1, \ldots, X_n) \tag{14}$$

  gọi là ước lượng Bayes cho θ.

[1] Đây là phần đọc thêm

Phương pháp Bayes [1] - Ví dụ

1. Xét $(X_1, \ldots, X_n)$ là một mẫu ngẫu nhiên chọn từ biến ngẫu nhiên X có phân phối Bernoulli, $X \sim B(1, p)$. Hàm mật độ xác suất của X cho bởi

$$f(x; p) = p^x(1-p)^{1-x}, \quad x = 0, 1$$

với $0 \le p \le 1$ chưa biết. Giả sử tham số p có phân phối đều trên khoảng $(0, 1)$. Tìm ước lượng Bayes cho p.

[1] Đây là phần đọc thêm

Phương pháp Bayes [1] - Ví dụ

Ta có

- Hàm mật độ xác suất tiên nghiệm của tham số p có phân phối đều
$$f(p) = 1, \quad 0 < p < 1$$
- Hàm mật độ xác suất hậu nghiệm của p cho bởi

$$f(p|x_1, \ldots, x_n) = \frac{f(x_1, \ldots, x_n, p)}{f(x_1, \ldots, x_n)}$$

[1] Đây là phần đọc thêm

Phương pháp Bayes [1] - Ví dụ

- Hàm mật độ xác suất đồng thời của $(X_1, \ldots, X_n)$ và p

$$\begin{aligned} f(x_1, \ldots, x_n, p) &= f(x_1, \ldots, x_n|p)f(p) \\ &= \left(\prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}\right) f(p) \\ &= p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i} = p^m(1-p)^{n-m} \end{aligned}$$

với $m = \sum_{i=1}^n x_i$.

$$f(x_1, \ldots, x_n) = \int_0^1 f(x_1, \ldots, x_n, p)dp = \int_0^1 p^m(1-p)^{n-m}dp$$

[1] Đây là phần đọc thêm

Phương pháp Bayes [1] - Ví dụ

Sử dụng kết quả: với các số nguyên m và k, ta có

$$\int_0^1 p^m(1-p)^k dp = \frac{m!k!}{(m+k+1)!}$$

Suy ra

$$f(x_1, \ldots, x_n) \int_0^1 p^m(1-p)^{n-m}dp = \frac{m!(n-m)!}{(n+1)!}$$

Hàm mật độ xác suất hậu nghiệm của p là

$$f(p|x_1, \ldots, x_n) = \frac{f(x_1, \ldots, x_n, p)}{f(x_1, \ldots, x_n)} = \frac{p^m(1-p)^{n-m}(n+1)!}{m!(n-m)!}$$

[1] Đây là phần đọc thêm

Giá trị ước lượng Bayes của p thu được bởi kỳ vọng có điều kiện

$$\begin{aligned}\mathbb{E}(p|x_1,\ldots,x_n) &= \int_0^1 pf(p|x_1,\ldots,x_n)dp \\ &= \frac{(n+1)!}{m!(n-m)!}\int_0^1 p^{m+1}(1-p)^{n-m}dp \\ &= \frac{(n+1)!}{m!(n-m)!}\frac{(m+1)!(n-m)!}{(n+2)!} \\ &= \frac{m+1}{n+2} = \frac{1}{n+2}\left(\sum_{i=1}^n x_i + 1\right)\end{aligned}$$

Ước lượng Bayes là

$$\hat{P} = \mathbb{E}(p|X_1,\ldots,X_n) = \frac{1}{n+2}\left(\sum_{i=1}^n X_i + 1\right)$$

[1] Đây là phần đọc thêm

2. Xét $(X_1,\ldots,X_n)$ là một mẫu ngẫu nhiên chọn từ biến ngẫu nhiên $X \sim Exp(\lambda)$, λ chưa biết. Biết rằng tham số λ là biến ngẫu nhiên có phân phối mũ với tham số α. Tìm ước lượng Bayes cho λ.

[1] Đây là phần đọc thêm