# 基础数理统计

# 第九章 参数推断

- 1 9.1 关注参数
- 2 9.2 矩估计
- ③ 9.3 极大似然估计
- 4 9.4 极大似然估计的相合性
- 5 9.5 极大似然估计的同变性
- 6 9.6 极大似然估计的渐近正态性
- 9.7 极大似然估计的最优性
- 8 9.8 Delta 方法
- 9.9 多参数模型
- 10 9.10 参数 Bootstrap 方法

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 注

- 9.1 关注参数

9.1 关注参数

9.2 矩估计

9.3 极大似然估计

正态性

9.7 极大似然估计的最优

9.8 Delta 方法

9.9 多参数模型

问题: 假设样本  $X_1, \dots, X_n$  来源于某一个总体 F, 通过样本来推断总体的某个特征  $\theta = \theta(F)$ 。例如

• 估计均值  $\mu$  或者标准差  $\sigma$ :

$$\theta = \theta(F) = \int x dF(x) = \mu$$
 或者  $\theta(F) = \sigma$ ;

- 3- $\sigma$  准则我们需要估计  $\theta(F) = \mu \pm 3\sigma$ ;
- 中位数  $\theta(F) = F^{-1}(1/2)$ ;
- 新的观察值大于 1 的概率:  $\theta(F) = P(X > 1) = \int_{1}^{\infty} dF(x).$

#### 9.1 关注参数

- 9.2 矩估计
- 0.3 极大似然估计
- 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方法

假定样本  $X_1, \dots, X_n$  来自于正态分布族  $N(\mu, \sigma^2)$ , 这里的参数  $\theta = (\mu, \sigma)$ . 感兴趣的问题可能是

- 估计均值  $\mu$  或者标准差  $\sigma$ :  $g(\theta) = \mu$  或者  $g(\theta) = \sigma$ .
- 3- $\sigma$  准则我们需要估计  $g(\theta) = \mu \pm 3\sigma$ ;
- 中位数  $g(\theta) = \mu$ ;
- 新的观察值大于 1 的概率:  $g(\theta) = P(X>1) = P(\frac{X-\mu}{\sigma} > \frac{1-\mu}{\sigma}) = 1 \Phi(\frac{1-\mu}{\sigma}).$

#### 9.1 关注参数

- 9.2 矩估计
- .3 极大似然估
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方法

# 分布族与点估计--续

同样的一组样本如果假定来自于指数分布族  $Exp(\lambda)$ , 其中参数为  $\theta = \lambda \in \{\lambda > 0\}$ . 感兴趣的问题为:

- 估计均值  $g(\theta) = 1/\lambda$  或标准差  $g(\theta) = 1/\lambda$ .
- 3- $\sigma$  准则我们需要估计  $g(\theta) = (-2/\lambda, 4/\lambda)$ ;
- 中位数  $g(\theta) = \log(2)/\lambda$ ;
- 大于 1 的概率:  $g(\theta) = P(X > 1) = \exp(-\lambda)$ .

#### 9.1 关注参数

- 9.2 矩估计
- Im 1 WIADAL
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐过 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 注

- 2 9.2 矩估计

9.2 矩估计

9.1 关注参数

9.3 极大似然估计

9.6 极大似然估计的渐近

正态性 9.7 极大似然估计的最优

9.8 Delta 方法

9.9 多参数模型

样本矩是统计中最常用的统计量:

$$\widehat{\alpha}_k = \frac{1}{n} \sum_{i=1}^n X_i^k, \ k = 1, 2, \cdots,$$

根据大数定律,其是总体矩  $\alpha_k = \int x^k dF(x)$  的相合估计,即

$$\widehat{\alpha}_k = \frac{1}{n} \sum_{i=1}^n X_i^k \stackrel{P}{\to} \alpha_k = \int x^k dF(x).$$

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 法

# 定义 1

 $\theta$  的矩估计  $\hat{\theta}_n$  满足

$$\alpha_i(\widehat{\theta}_n) = \widehat{\alpha}_i, \ i = 1, 2, \dots, k.$$

9.1 关注参数

#### 9.2 矩估计

- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 法

假设简单样本  $X_1, \ldots, X_n$  来自于分布族  $\{f(x, \theta)\}$  其中有 m 个未知参数  $\theta = (\theta_1, \cdots, \theta_m)$ ,可以计算出其与总体矩  $\alpha_k$  之间的数学关系:

$$lpha_1 = E(X) = \int x df(x, \theta) = lpha_1(\theta_1, \cdots, \theta_m),$$
 $\vdots$ 
 $\alpha_k = E(X^k) = \int x^k df(x, \theta) = lpha_k(\theta_1, \cdots, \theta_m)$ 
如需要可以继续计算下去.

9.1 关注参数

9.2 矩估计

3 极士们然估计

9.4 极大似然估计的相合 性

9.5 极大似然估计的同变 性

9.6 极大似然估计的渐近 正态性 9.7 极大似然估计的最优

注

9.8 Delta 方法

9.9 多参数模型

通过未知参数  $\theta_1, \dots, \theta_m$  和各阶矩  $\alpha_1, \dots, \alpha_k (k \ge m)$  之间关系方程,可以解出

$$\theta_j = \theta_j(E(X), \dots, E(X^k)) = \theta_j(\alpha_1, \dots, \alpha_k), j = 1, \dots, m.$$

用样本矩  $\hat{\alpha}_j$  代替入上式中的  $\alpha_j$ , 即得到原始参数  $\theta_1, \dots, \theta_m$  的估计

$$\hat{\theta}_j = \theta_j(\widehat{\alpha}_1, \cdots, \widehat{\alpha}_k), \ j = 1, \cdots, m.$$

- 9.1 关注参数
- 9.2 矩估计
  - .3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方

### 对于简单样本

• 每个样本矩都是总体矩的无偏估计, 即

$$\mathsf{E}\widehat{\alpha}_k = \frac{1}{n} \sum_{i=1}^n \mathsf{E}(X_i^k) = \int x^k dF(x) = \alpha_k,$$

• 对于一般的矩估计,

$$\mathsf{E}\theta(\widehat{\alpha}_1,\cdots,\widehat{\alpha}_k)\neq\theta(\alpha_1,\cdots,\alpha_k)$$

所以矩估计不一定具有无偏性。

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方法

由大数定律,每个样本矩都是总体矩的相合估计,即

$$\widehat{\alpha}_{\mathbf{k}} \stackrel{P}{\to} \alpha_{\mathbf{k}},$$

所以对应连续的参数函数  $\theta(\cdot)$ ,

$$\theta(\widehat{\alpha}_1, \cdots, \widehat{\alpha}_k) \stackrel{P}{\to} \theta(\alpha_1, \cdots, \alpha_k),$$

所以一般的矩估计是相合估计。

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方法

# 矩估计的性质

### 定理 1

令  $\hat{\theta}_n$  表示矩估计。在适当的条件下,下述成立:

- 1. 矩估计  $\hat{\theta}_n$  以接近概率 1 存在;
- 2. 矩估计是相合估计;
- 3. 矩估计是渐近正态的:  $\sqrt{n}(\widehat{\theta}_n \theta) \rightsquigarrow N(0, \Sigma)$ , 其中

$$\Sigma = \mathbf{g} V_{\theta}(\mathbf{y}) \mathbf{g}^{\top},$$

$$\mathbf{y} = (X, X^{2}, \dots, X^{k})^{\top},$$

$$\mathbf{g} = (g_{1}, g_{2}, \dots, g_{k}),$$

$$g_{j} = \partial \alpha_{j}^{-1}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}.$$

作业: 1

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 法

- 1 9.1 关注参数
- 2 9.2 矩估计
- 3 9.3 极大似然估计
- 4 9.4 极大似然估计的相合性
- 5 9.5 极大似然估计的同变性
- 6 9.6 极大似然估计的渐近正态性
- 7 9.7 极大似然估计的最优性
- 8 9.8 Delta 方法

9.1 关注参数

2 担估订

9.3 极大似然估计

极大似然估计的

9.6 极大似: 正态性

9.7 极大似然估计的最优性

9.8 Delta 方法

9.9 多参数模型

令  $X_1, X_2, ..., X_n$  独立同分布于概率密度函数  $f(x; \theta)$ 。

# 定义 2

似然函数定义为:

$$L_n(\theta) = \prod_{i=1}^n f(X_i, \theta), \theta \in \Theta,$$

其中  $\Theta$  是参数空间。对数似然函数为  $\ell(\theta) = \log L_n(\theta)$ 。

# 定义 3

极大似然估计 (maximum likelihood estimation, MLE) 记为  $\hat{\theta}_n$ ,是使得似然函数  $L_n(\theta)$  达到最大值的  $\theta$  的值。

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的问变性
- 9.6 极大似然估计的渐近 正态性 9.7 极大似然估计的最优
- 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方法

# 例 1

假设  $X_1, X_2, \ldots, X_n$  独立同分布于  $N(\mu, \sigma^2)$ , 参数为  $\theta = (\mu, \sigma)$ , 对数似然函数为

$$\ell(\mu, \sigma) = -n \log \sigma - \frac{nS^2}{2\sigma^2} - \frac{n(\bar{X} - \mu)^2}{2\sigma^2},$$

其中 
$$\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$$
,  $S^2 = n^{-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ 。最大似然估计为

$$\hat{\mu} = \bar{X}, \ \hat{\sigma} = S.$$

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 性
- 9.5 极大似然估计的同变 性
- 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 法

# 例 2

令  $X_1, X_2, ..., X_n$  独立同分布于  $Uniform[0, \theta]$ , 似然函数为

$$L_n(\theta) = \theta^{-n} I(\theta \ge X_{(n)}),$$

故最大似然估计为  $\hat{\theta}_n = X_{(n)}$ .

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方

### **定理** 2

设  $X_{(i)}(i=1,2,\cdots,n)$  是多元正态总体  $N_p(\mu,\Sigma)$  的随机样本, n>p, 则  $\mu,\Sigma$  的最大似然估计为  $\widehat{\mu}=\bar{X},\widehat{\Sigma}=\frac{1}{n}A$ , 这里

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_{(i)}, \ A = \sum_{i=1}^{n} (X_{(i)} - \bar{X})(X_{(i)} - \bar{X})^{\top}.$$

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 注

#### 注意到

$$\begin{split} \log \ell(\pmb{\mu}, \Sigma) &= -\frac{np}{2} \ln(2\pi) - \frac{n}{2} (\bar{X} - \pmb{\mu})^\top \Sigma^{-1} (\bar{X} - \pmb{\mu}) \\ &- \frac{n}{2} \ln |\Sigma| - \frac{1}{2} \mathrm{tr}(\Sigma^{-1} A). \\ \log \ell(\bar{X}, \Sigma) &= -\frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln |\frac{A}{n}| \\ &- \frac{n}{2} \left[ \mathrm{tr} \left( \Sigma^{-1/2} \frac{A}{n} \Sigma^{-1/2} \right) \right. \\ &- \log \left| \Sigma^{-1/2} \frac{A}{n} \Sigma^{-1/2} \right| \right]. \end{split}$$

- 9.1 关注参数
- 0 2 毎估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性 9.7 极大似然估计的最优
- 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 法

#### 引理 1

设 B 为 p 阶正定矩阵,则  $tr(B) - \ln |B| \ge p$ ,且等号成立当且仅当  $B = I_p$ .

### 证明.

利用 
$$\ln(1+x) \le x$$
,  $\forall x > -1$ .  $x = 0$ ,  $\ln(1+x) = x$ ;  $-1 < x < 0$ ,  $(\ln(1+x))' > 1$ ;  $x > 0$ ,  $(\ln(1+x))' < 1$ .

利用该引理即得到多元正态总体中参数的最大似然估计。

作业: 2, 3, 4

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 注

- 1 9.1 关注参数
- 2 9.2 矩估计
- ③ 9.3 极大似然估计
- 4 9.4 极大似然估计的相合性
- ⑤ 9.5 极大似然估计的同变性
- 6 9.6 极大似然估计的渐近正态性
- 7 9.7 极大似然估计的最优性
- ® 9.8 Delta 方法

9.1 关注参数

2 矩估计

9.3 极大似然估计 9.4 极大似然估计的相合

性 9.5 极大似然估计的同变

9.6 极大似然估计的渐近 正态性

9.7 极大似然估计的最优 性

9.8 Delta 方法

9.9 多参数模型

# 极大似然估计的相合性

f,g 为概率密度函数, 定义 f 和 g 间的 KL 距离为

$$D(f,g) = \int f(x) \log \left(\frac{f(x)}{g(x)}\right) dx.$$

对任意  $\theta, \phi \in \Theta$ , 令  $D(\theta, \phi)$  表示  $D(f(x; \theta), f(x; \phi))$ .

- 9.1 关注参数
- 92 钜估计
- 9.3 极大似然估计

#### 9.4 极大似然估计的相合 性

- 9.5 极大似然估计的同变性
- 9.6 极大似然估计的渐近 正态性 9.7 极大似然估计的最优
- 生
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方法

# 极大似然估计的相合性--续

#### 定理 3

令  $\theta_{\star}$  为  $\theta$  的真实值, 定义

$$M_n(\theta) = \frac{1}{n} \sum_i \log \left( \frac{f(X_i; \theta)}{f(X_i; \theta_*)} \right),$$

且  $M(\theta) = -D(\theta_{\star}, \theta)$ 。 假设

$$\sup_{\theta\in\Theta}|M_n(\theta)-M(\theta)|\stackrel{P}{\to}0,$$

并且对任意  $\epsilon > 0$ ,

$$\sup_{\theta: |\theta - \theta_{\star}| \ge \epsilon} M(\theta) < M(\theta_{\star}).$$

令  $\hat{\theta}_n$  表示极大似然估计,则  $\hat{\theta}_n \stackrel{P}{\to} \theta_{\star}$ 。

9.1 关注参数

9.2 矩估计

9.3 极大似然估计

9.4 极大似然估计的相合 性

9.5 极大似然估计的同变 性

9.6 极大似然估计的渐近 正态性 9.7 极大似然估计的最优

±

9.8 Delta 方法

9.9 多参数模型

- 1 9.1 关注参数
- 2 9.2 矩估计
- ③ 9.3 极大似然估计
- 4 9.4 极大似然估计的相合性
- ⑤ 9.5 极大似然估计的同变性
- 6 9.6 极大似然估计的渐近正态性
- 7 9.7 极大似然估计的最优性
- ⑧ 9.8 Delta 方法

9.1 关注参数

担怕订

9.3 极大似然估计 9.4 极大似然估计的相合

9.5 极大似然估计的同变 性

9.6 极大似然估计的渐近 正态性 9.7 极大似然估计的最优

9.8 Delta 方法 9.9 多参数模型

9.9 多参数模型

# 极大似然估计的同变性

设参数  $\theta$  的变化范围是  $\Theta, L(\theta)$  是似然函数。设  $\tau=g(\pmb{\theta})$  的取值范围是  $\Theta^\star$ 。对任何  $\tau\in\Theta^\star$ ,令  $M(\tau)=\sup_{\theta:g(\theta)=\tau}L(\theta)$ 。

### 定义 4

称  $M(\tau)$  为函数  $g(\theta)$  诱导出的似然函数.

### 定义 5

若  $\hat{\tau}$  满足  $M(\hat{\tau}) = \sup_{\tau \in \Theta^*} M(\tau)$ , 则称  $\hat{\tau}$  是  $g(\theta)$  的最大似然估计。

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 法

# 定理 4

令  $\tau = g(\theta)$  是  $\theta$  的函数。  $\hat{\theta}_n$  是  $\theta$  的极大似然估计,则  $\hat{\tau}_n = g(\hat{\theta}_n)$  是  $\tau$  的极大似然估计。

## 证明.

根据定义. $\forall \tau \in \Theta^*$ .

$$M(\tau) = \sup_{\theta: g(\theta) = \tau} L(\theta) \le L(\widehat{\theta}_n) \le \sup_{\theta: g(\theta) = \widehat{\tau}_n} L(\theta) = M(\widehat{\tau}_n).$$

9.1 关注参数

99 钜估计

9.3 极大似然估计

9.4 极大似然估计的相合 性

9.5 极大似然估计的同变性

9.6 极大似然估计的渐近 正态性

9.7 极大似然估计的最优 性

9.8 Delta 方法

9.9 多参数模型

 $\Box$ 

# 极大似然估计的同变性--续

### 例 3

 $x_1, x_2, \cdots, x_n$  是来自  $N_2(\mu, \Sigma)$  的简单随机样本,  $\Sigma = (\sigma_{ij}), \sigma_{ij} = \sigma_i \sigma_j \rho(i, j = 1, 2)$ 。 $\mu, \sigma_i^2 (i = 1, 2)$  的最大 似然估计为

$$\widehat{\boldsymbol{\mu}} = \bar{\mathbf{x}}, \ \widehat{\sigma}_i^2 = \frac{1}{n} \sum_{\alpha=1}^n (\mathbf{x}_{\alpha i} - \bar{\mathbf{x}}_i)^2.$$

 $\rho$  的最大似然估计为

$$\widehat{\rho} = \frac{\sum_{\alpha=1}^{n} (x_{\alpha 1} - \bar{x}_1)(x_{\alpha 2} - \bar{x}_2)}{\sqrt{\sum_{\alpha=1}^{n} (x_{\alpha 1} - \bar{x}_1)^2} \sqrt{\sum_{\alpha=1}^{n} (x_{\alpha 2} - \bar{x}_2)^2}}$$

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性 9.7 极大似然估计的最优
- III.
- 9.8 Delta 方法
- 9.9 多参数模型

- 1 9.1 关注参数
- 2 9.2 矩估计
- ③ 9.3 极大似然估计
- 4 9.4 极大似然估计的相合性
- 5 9.5 极大似然估计的同变性
- 6 9.6 极大似然估计的渐近正态性
- 7 9.7 极大似然估计的最优性
- ⑧ 9.8 Delta 方法

- 9.1 关注参数
  - 矩估计
- 9.3 极大似然估计 9.4 极大似然估计的相合
  - 5 极士心然体计位
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方法

# Fisher 信息量

## 定义 6

记分函数定义为

$$s(X; \theta) = \frac{\partial \log f(X; \theta)}{\partial \theta}.$$

Fisher 信息量定义为

$$I_n(\theta) = V_{\theta}\left(\sum_{i=1}^n s(X_i; \theta)\right) = \sum_{i=1}^n V_{\theta}(s(X_i; \theta)).$$

## 定理 5

$$I_n(\theta) = nI(\theta)$$
, 这里

$$I(\theta) = -E_{\theta} \left( \frac{\partial^2 \log f(X; \theta)}{\partial \theta^2} \right) = E_{\theta} s^2(X; \theta).$$

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 法

# 渐近正态性

令 
$$se = \sqrt{V(\hat{\theta}_n)}$$
,在适当的正则条件下,下述结论成立

1. se ≈  $\sqrt{1/I_n(\theta)}$  且

$$\frac{\widehat{\theta}_n - \theta}{\text{se}} \rightsquigarrow N(0, 1).$$

2. 
$$\diamondsuit \widehat{se} = \sqrt{1/I_n(\widehat{\theta})} \blacksquare$$

$$\frac{\widehat{\theta}_n - \theta}{\widehat{se}} \rightsquigarrow N(0,1).$$

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性 9.7 极大似然估计的最优
- 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 法

# 渐近正态性--续

定理 7

令 
$$C_n = (\widehat{\theta} - z_{\alpha/2}\widehat{se}, \ \widehat{\theta} + z_{\alpha/2}\widehat{se})$$
, 则当  $n \to \infty$  时,  $P_{\theta}(\theta \in C_n) \to 1 - \alpha_{\bullet}$ 

9.1 关注参数

9.2 矩估计

9.3 极大似然估计

9.4 放入以然值订的相管 性

9.5 极大似然估计的同变 性

9.6 极大似然估计的渐近 正态性

9.7 极大似然估计的最优性

9.8 Delta 方法

9.9 多参数模型

例 4

令 
$$X_1, X_2, \ldots, X_n \sim N(\theta, \sigma^2), \sigma^2$$
 已知。则

$$f(X;\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(X-\theta)^2}{2\sigma^2}\right\},\,$$

 $\theta$  的最大似然估计是  $\hat{\theta} = \bar{X}$ 。记分函数为

$$s(X; \theta) = \frac{\partial \log f(X; \theta)}{\partial \theta} = \frac{X - \theta}{\sigma^2},$$

Fisher 信息量定义为

$$I_n(\theta) = nEs^2(X; \theta) = \frac{n}{\sigma^2}.$$

故  $\bar{X}$  近似服从  $N(\theta, \frac{\sigma^2}{n})$  (本例中为精确分布)。

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然怕订的问受 性
- 9.6 极大似然估计的渐近 正态性 9.7 极大似然估计的最优
- 生
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方法

例 5

$$\diamondsuit$$
  $X_1, X_2, \ldots, X_n \sim \textit{Poisson}(\lambda)$ 。则

$$f(X; \lambda) = \frac{\lambda^X}{X!} \exp(-\lambda),$$

 $\lambda$  的最大似然估计是  $\hat{\lambda} = \bar{X}$ 。记分函数为

$$s(X; \lambda) = \frac{\partial \log f(X; \lambda)}{\partial \lambda} = \frac{X}{\lambda} - 1,$$

Fisher 信息量定义为

$$I_n(\theta) = nEs^2(X; \theta) = \frac{n}{\lambda}.$$

故  $\bar{X}$  近似服从  $N(\lambda, \frac{\lambda}{n})$ 。

作业:5

9.1 关注参数

9.2 矩估计

9.3 极大似然估计

9.4 极大似然估计的相合 性

9.5 极大似然估计的同变 性

9.6 极大似然估计的渐近 正态性

9.7 极大似然估计的最优 性

9.8 Delta 方法 9.9 多参数模型

# 9.7 极大似然估计的最优性

- 9.1 关注参数

- 9.7 极大似然估计的最优性

9.1 关注参数

9.3 极大似然估计

9.4 极大似然估计的相合

正态性

9.7 极大似然估计的最优

9.8 Delta 方法

9.9 多参数模型

# 极大似然估计的最优性

### 定义 7

考虑两个估计量  $T_n$  和  $U_n$ ,并假设

$$\sqrt{n}(T_n - \theta) \rightsquigarrow N(0, t^2), \sqrt{n}(U_n - \theta) \rightsquigarrow N(0, u^2).$$

定义 U 对于 T 的渐近相对效为  $ARE(U,T)=t^2/u^2$ 。

## 定理8

如果  $\hat{\theta}_n$  是极大似然估计, $\tilde{\theta}_n$  是其它任意估计,则

$$ARE(\widetilde{\theta}_n, \widehat{\theta}_n) \leq 1.$$

即极大似然估计具有最小(渐近)方差,称极大似然估 计是最有效的或渐近最优的。

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 性 9.5 极大似然估计的同变
- 9.6 极大似然估计的渐近
- 9.7 极大似然估计的最优
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方

- 1 9.1 关注参数
- 2 9.2 矩估计
- ③ 9.3 极大似然估计
- 4 9.4 极大似然估计的相合性
- 5 9.5 极大似然估计的同变性
- ⑤ 9.6 极大似然估计的渐近正态性
- 7 9.7 极大似然估计的最优性
- ଃ 9.8 Delta 方法

- 9.1 关注参数
  - 矩估计
- 9.3 极大似然估计
- 9.6 极大似然估计的渐近 正态性 9.7 极大似然估计的最优
  - Dolto 古注
- 9.8 Delta 方法 0.0 冬糸粉档刑
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 法

# Delta 方法

令  $\tau = g(\theta)$ , 这里 g 是光滑函数。 $\tau$  的极大似然估计为  $\hat{\tau}_n = g(\hat{\theta}_n)$ ,  $\hat{\theta}_n$  为  $\theta$  的极大似然估计。

## 定理 9 (Delta 方法)

如果  $\tau = g(\theta)$ , 其中, g 可微, 且  $g'(\theta) \neq 0$ , 则

$$\frac{\widehat{\tau}_{\mathbf{n}} - \tau}{\widehat{\mathbf{se}}(\widehat{\tau}_{\mathbf{n}})} \leadsto \mathbf{N}(0, 1),$$

其中 
$$\hat{\tau}_n = g(\hat{\theta}_n)$$
 且

$$\widehat{se}(\widehat{\tau}_n) = |g'(\widehat{\theta}_n)|\widehat{se}(\widehat{\theta}_n).$$

记

$$C_n = (\widehat{\tau}_n - z_{\alpha/2}\widehat{se}(\widehat{\tau}_n), \widehat{\tau}_n + z_{\alpha/2}\widehat{se}(\widehat{\tau}_n)),$$

则当  $n \to \infty$  时, $P_{\theta}(\tau \in C_n) \to 1 - \alpha$ 。

9.1 关注参数

9.2 矩估计

9.3 极大似然估计

9.4 极大似然估计的相合 性

9.5 极大似然估计的同变性

9.6 极大似然估计的渐近 正态性

9.7 极大似然估计的最优 性

9.8 Delta 方法

9.9 多参数模型

9.10 参数 Bootstrap 方 法

#### 例 6

令  $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$ ,假设  $\mu$  已知, $\sigma$  未知,希望估计  $\psi = \log \sigma$ 。

对数似然函数为

$$\ell(\sigma) = -n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2.$$

Fisher 信息量为

$$I_n(\sigma) = nI(\sigma) = V\left(\frac{\partial \ell(\sigma)}{\partial \sigma}\right) = \frac{2n}{\sigma^2}.$$

最大似然估计为

$$\widehat{\sigma}_n = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}}.$$

9.1 关注参数

9.2 矩估计

9.3 极大似然估计

9.4 极大似然估计的相合

9.5 极大似然估计的同变性

9.6 极大似然估计的渐近 正态性

9.7 极大似然估计的最优 性

9.8 Delta 方法

9.9 多参数模型

9.10 参数 Bootstrap 方 法 故  $\psi$  的近似  $1-\alpha$  置信区间为

$$\left(\log\left(\widehat{\sigma}_{n}\right) - \frac{z_{\alpha/2}}{\sqrt{2n}}, \log\left(\widehat{\sigma}_{n}\right) + \frac{z_{\alpha/2}}{\sqrt{2n}}\right).$$

作业: 6

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方法

9.1 关注参数

9.3 极大似然估计

9.6 极大似然估计的渐近 正态性

9.7 极大似然估计的最优

9.8 Delta 方法

9.9 多参数模型

9.10 参数 Bootstrap 方

令 
$$\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k)^{\top}$$
 且令  $\widehat{\boldsymbol{\theta}}_n = (\widehat{\theta}_1, \widehat{\theta}_2, \dots, \widehat{\theta}_k)$  为极大似然估计。令  $\ell_n = \sum_{i=1}^n \log f(X_i; \boldsymbol{\theta}),$ 

$$H_{jj} = \frac{\partial^2 \ell_n}{\partial \theta_i^2}, \ H_{jk} = \frac{\partial^2 \ell_n}{\partial \theta_j \partial \theta_k}.$$

定义 Fisher 信息矩阵为

$$I_{n}(\theta) = -\begin{pmatrix} \mathsf{E}_{\theta}(H_{11}) & \mathsf{E}_{\theta}(H_{12}) & \dots & \mathsf{E}_{\theta}(H_{1k}) \\ \mathsf{E}_{\theta}(H_{21}) & \mathsf{E}_{\theta}(H_{22}) & \dots & \mathsf{E}_{\theta}(H_{2k}) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{E}_{\theta}(H_{k1}) & \mathsf{E}_{\theta}(H_{k2}) & \dots & \mathsf{E}_{\theta}(H_{kk}) \end{pmatrix}$$

令  $J_n(\theta) = I_n^{-1}(\theta)$  是  $I_n(\theta)$  的逆矩阵。

9.1 关注参数

9.2 矩估计

9.3 极大似然估计

.4 极大似然估计的相合 生

9.5 极大似然估计的同变 性

9.6 极大似然估计的渐近 正态性 9.7 极大似然估计的最优

性

9.8 Delta 方法

9.9 多参数模型

9.10 参数 Bootstrap 方 法

#### 定理 10

在适当的正则条件下, $\hat{\theta}_n - \theta \rightsquigarrow N(\mathbf{0}, J_n)$ . 令  $\tau = g(\theta_1, \dots, \theta_k)$  为一个函数,令

$$\nabla g = \left(\frac{\partial g}{\partial \theta_1}, \dots, \frac{\partial g}{\partial \theta_k}\right)^{\top}.$$

# 定理 11 (多参数 Delta 方法)

假设  $\nabla g$  在  $\hat{\theta}_n$  处不等于  $\theta$ ,  $\hat{\tau}_n = g(\hat{\theta}_n)$ , 则

$$\frac{\hat{\tau}_n - \tau}{\widehat{\mathsf{se}}(\widehat{\tau}_n)} \rightsquigarrow \mathsf{N}(0,1).$$

其中

$$\widehat{se}(\widehat{\tau}_n) = \sqrt{(\nabla g|_{\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}_n})^{\top} J_n(\widehat{\boldsymbol{\theta}}_n)(\nabla g|_{\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}_n})}.$$

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合性
- 9.5 极大似然估计的同变性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
  - 9.9 多参数模型
  - 9.10 参数 Bootstrap 方 法

#### 例 7

令  $X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$ ,假设  $\mu, \sigma$  未知,希望估计  $\psi = \sigma/\mu$ 。

对数似然函数为

$$\ell(\sigma) = -n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \mu)^2.$$

 $(\mu, \sigma^2)$  的极大似然估计为

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

Fisher 信息矩阵为

$$I_n(\mu,\sigma) = \begin{pmatrix} \frac{n}{\sigma^2} & 0\\ 0 & \frac{2n}{\sigma^2} \end{pmatrix}.$$

- 9.1 关注参数
- 9.2 矩估计
- 0.2 tr+////////
- 9.4 极大似然估计的相合性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性 9.7 极大似然估计的最优
- 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 法

故

$$J_n = \frac{1}{n} \begin{pmatrix} \sigma^2 & 0 \\ 0 & \frac{\sigma^2}{2} \end{pmatrix}.$$

g 的梯度为  $\nabla g = (-\frac{\sigma}{\mu^2} \ \frac{1}{\mu})^{\top}$ ,故

$$\widehat{\mathsf{se}}(\widehat{\tau}) = \frac{1}{\sqrt{n}} \sqrt{\frac{\widehat{\sigma}^4}{\widehat{\mu}^4} + \frac{\widehat{\sigma}^2}{2\widehat{\mu}^2}}.$$

作业: 7(a)(b)(c) ((d) 选做)

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 法

- 1 9.1 关注参数
- 2 9.2 矩估计
- ③ 9.3 极大似然估计
- 4 9.4 极大似然估计的相合性
- 5 9.5 极大似然估计的同变性
- 6 9.6 极大似然估计的渐近正态性
- 9.7 极大似然估计的最优性
- 8 9.8 Delta 方法

9.1 关注参数 9.2 矩估计

9.3 极大似然估计

极大似然估

正态性 9.7 极大似然估计的最优 性

9.8 Delta 方法 9.9 多参数模型

9.10 参数 Bootstrap 方

例 8

令  $X_1,X_2,\ldots,X_n\sim N(\mu,\sigma^2)$ ,假设  $\mu,\sigma$  未知,希望估计  $\psi=\sigma/\mu$ 。

1. 计算

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

2. 随机模拟  $X_1^\star, X_2^\star, \ldots, X_n^\star \sim N(\widehat{\mu}, \widehat{\sigma}^2)$ , 计算

$$\widehat{\mu}^{\star} = \frac{1}{n} \sum_{i=1}^{n} X_i^{\star}, \quad \widehat{\sigma}^{\star 2} = \frac{1}{n} \sum_{i=1}^{n} (X_i^{\star} - \bar{X}^{\star})^2. \quad \widehat{\psi}^{\star} = \widehat{\sigma}^{\star} / \widehat{\mu}^{\star}.$$

- 9.1 关注参数
- 9.2 矩估计
- 3.3 极大似然估计
- 9.4 极大似然估计的相合 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方法

3. 重复上述过程 B 次得到 Bootstrap 复本  $\widehat{\psi}_1^\star,\ldots,\widehat{\psi}_B^\star$ 。标准误差的估计值为

$$\widehat{\mathsf{se}}_{boot} = \sqrt{\frac{\sum\limits_{b=1}^{B}(\widehat{\psi}_{b}^{\star} - \widehat{\psi})^{2}}{B}}$$

作业: 9(a) 用 Delta 方法计算  $\widehat{se}$  和  $\theta$  的 90% 的置信区 间。

- 9.1 关注参数
- 9.2 矩估计
- 9.3 极大似然估计
- 性
- 9.5 极大似然估计的同变 性
- 9.6 极大似然估计的渐近 正态性
- 9.7 极大似然估计的最优 性
- 9.8 Delta 方法
- 9.9 多参数模型
- 9.10 参数 Bootstrap 方 法