基础数理统计

第九章　参数推断

9.1 关注参数

参数估计

问题：假设样本 $X_1, \cdots, X_n$ 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 μ 或者标准差 σ:
  $\theta = \theta(F) = \int x dF(x) = \mu$ 或者 $\theta(F) = \sigma$;
- 3-σ 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:
  $\theta(F) = P(X > 1) = \int_1^\infty dF(x)$.

分布族与点估计

假定样本 $X_1, \cdots, X_n$ 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 μ 或者标准差 σ: $g(\theta) = \mu$ 或者 $g(\theta) = \sigma$.
- 3-σ 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: $g(\theta) = P(X > 1) = P(\frac{X-\mu}{\sigma} > \frac{1-\mu}{\sigma}) = 1 - \Phi(\frac{1-\mu}{\sigma})$.

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(λ), 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-σ 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: $g(\theta) = P(X > 1) = \exp(-\lambda)$.

9.2 矩估计

k 阶矩

样本矩是统计中最常用的统计量:

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

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

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

矩估计

定义 1

θ 的矩估计 $\widehat{\theta}_n$ 满足

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

矩估计

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

$$
\begin{aligned}
\alpha_1 &= E(X) = \int x df(x, \boldsymbol{\theta}) = \alpha_1(\theta_1, \cdots, \theta_m), \\
&\vdots \\
\alpha_k &= E(X^k) = \int x^k df(x, \boldsymbol{\theta}) = \alpha_k(\theta_1, \cdots, \theta_m)
\end{aligned}
$$

如需要可以继续计算下去.

矩估计--续

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

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

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

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

对于简单样本

- 每个样本矩都是总体矩的无偏估计，即

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

- 对于一般的矩估计，

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

所以矩估计不一定具有无偏性。

矩估计的相合性

由大数定律，每个样本矩都是总体矩的相合估计，即

$$\widehat{\alpha}_k \xrightarrow{P} \alpha_k,$$

所以对应连续的参数函数 $\theta(\cdot)$，

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

所以一般的矩估计是相合估计。

矩估计的性质

定理 1

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

1. 矩估计 $\widehat{\theta}_n$ 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

$$\begin{aligned}
\Sigma &= \boldsymbol{g} V_\theta(\boldsymbol{y}) \boldsymbol{g}^\top, \\
\boldsymbol{y} &= (X, X^2, \ldots, X^k)^\top, \\
\boldsymbol{g} &= (g_1, g_2, \ldots, g_k), \\
g_j &= \partial \alpha_j^{-1}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}.
\end{aligned}$$

作业：1

9.3 极大似然估计

似然函数

令 $X_1, X_2, \ldots, X_n$ 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 Θ 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 $\widehat{\theta}_n$，是使得似然函数 $L_n(\theta)$ 达到最大值的 θ 的值。

例子

例 1

假设 $X_1, X_2, \ldots, X_n$ 独立同分布于 $N(\mu, \sigma^2)$，参数为 $\boldsymbol{\theta} = (\mu, \sigma)$，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

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

例子--续

例 2

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

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

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

多元正态总体

定理 2

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

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

多元正态总体--续

注意到

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

引理 1

设 B 为 p 阶正定矩阵，则 $tr(B) - \ln|B| \geq p$, 且等号成立当且仅当 $B = \boldsymbol{I}_p$.

证明.

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

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

9.4 极大似然估计的相合性

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 $D(f(x; \theta), f(x; \phi))$.

极大似然估计的相合性--续

定理 3

令 $\theta_\star$ 为 θ 的真实值，定义

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

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

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

并且对任意 $\epsilon > 0$，

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

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

9.5 极大似然估计的同变性

极大似然估计的同变性

设参数 θ 的变化范围是 $\Theta, L(\theta)$ 是似然函数。设 $\tau = g(\boldsymbol{\theta})$ 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 $\widehat{\tau}$ 满足 $M(\widehat{\tau}) = \sup_{\tau\in\Theta^\star} M(\tau)$, 则称 $\widehat{\tau}$ 是 $g(\theta)$ 的最大似然估计。

极大似然估计的同变性--续

定理 4

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

证明.

根据定义,$\forall \tau \in \Theta^\star$,

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

$\square$

极大似然估计的同变性--续

例 3

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

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

ρ 的最大似然估计为

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

9.6 极大似然估计的渐近正态性

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

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

渐近正态性

定理 6

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

1. $se \approx \sqrt{1/I_n(\theta)}$ 且

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

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

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

渐近正态性--续

定理 7
令 $C_n = \left(\widehat{\theta} - z_{\alpha/2}\widehat{se},\ \widehat{\theta} + z_{\alpha/2}\widehat{se}\right)$，则当 $n \to \infty$ 时，$P_\theta(\theta \in C_n) \to 1-\alpha$。

例子

例 4

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

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

θ 的最大似然估计是 $\widehat{\theta} = \bar{X}$。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)。

例子--续

例 5

令 $X_1, X_2, \ldots, X_n \sim Poisson(\lambda)$。则

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

λ 的最大似然估计是 $\widehat{\lambda} = \bar{X}$。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

9.7 极大似然估计的最优性

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

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

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 $\widehat{\theta}_n$ 是极大似然估计，$\widetilde{\theta}_n$ 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。

9.8 Delta 方法

Delta 方法

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

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

其中 $\widehat{\tau}_n = g(\widehat{\theta}_n)$ 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，$P_\theta(\tau \in C_n) \to 1-\alpha$。

例子

例 6

令 $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$，假设 μ 已知，σ 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

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

最大似然估计为

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

例子

故 ψ 的近似 $1-\alpha$ 置信区间为

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

作业：6

9.9 多参数模型

多参数模型

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

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

定义 Fisher 信息矩阵为

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

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

定理 10

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

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

定理 11 (多参数 Delta 方法)

假设 ∇g 在 $\widehat{\boldsymbol{\theta}}_n$ 处不等于 0，令 $\hat{\tau}_n = g(\widehat{\boldsymbol{\theta}}_n)$，则

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

其中

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

例子

例 7

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

对数似然函数为

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

(μ, σ^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

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

例子

故

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

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

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

作业：7(a)(b)(c) ((d) 选做)

9.10 参数 Bootstrap 方法

例子

例 8

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

1. 计算

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

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

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

例子--续

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

$$\widehat{\mathrm{se}}_{\mathrm{boot}} = \sqrt{\frac{\sum_{b=1}^{B} (\widehat{\psi}_b^\star - \widehat{\psi})^2}{B}}$$

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