6.4 最小方差无偏估计

1. 设总体概率函数是 $p(x;\theta), x_1, x_2, \cdots, x_n$ 是其样本，$T = T(x_1, x_2, \cdots, x_n)$ 是 θ 的充分统计量，则对 $g(\theta)$ 的任一估计 $\hat{g}$，令 $\tilde{g} = E(\hat{g}|T)$，证明：$MSE(\tilde{g}) \leqslant MSE(\hat{g})$。这说明，在均方误差准则下，人们只需要考虑基于充分统计量的估计。

证明：

$$
\begin{aligned}
\mathrm{MSE}(\hat{g}) &= E((\hat{g} - g(\theta))^2) = E\left((\hat{g} - \tilde{g} + \tilde{g} - g(\theta))^2\right) \\
&= E(\hat{g} - \tilde{g})^2 + 2E(\hat{g} - \tilde{g})(\tilde{g} - g(\theta)) + E(\tilde{g} - g(\theta))^2 \\
&= E(\hat{g} - \tilde{g})^2 + 2E(\hat{g} - \tilde{g})(\tilde{g} - g(\theta)) + \mathrm{MSE}(\tilde{g})
\end{aligned}
$$

其中

$$
E(\hat{g} - \tilde{g})(\tilde{g} - g(\theta)) = E(E((\hat{g} - \tilde{g})(\tilde{g} - g(\theta))|T))
$$

由于 T 是充分统计量，所以 $E(\tilde{g} - g(\theta))$ 与 T 无关，所以

$$
\begin{aligned}
E(\hat{g} - \tilde{g})(\tilde{g} - g(\theta)) &= E(\tilde{g} - g(\theta))E(E(\hat{g} - \tilde{g}|T)) \\
&= [E(\tilde{g}) - E(g(\theta))]E(E(\hat{g} - \tilde{g}|T)) = 0
\end{aligned}
$$

所以

$$
\mathrm{MSE}(\hat{g}) = E(\hat{g} - \tilde{g})^2 + \mathrm{MSE}(\tilde{g}) \geqslant \mathrm{MSE}(\tilde{g})
$$

□

3. 设 T 是 $g(\theta)$ 的 UMVUE，$\hat{g}$ 是 $g(\theta)$ 的无偏估计，证明：若 $\mathrm{Var}(\hat{g}) < \infty$，则 $\mathrm{Cov}(T, \hat{g}) \geqslant 0$。

证明：

由于 T 是 $g(\theta)$ 的 UMVUE，所以 $E(T) = g(\theta), \mathrm{Var}(T) < \infty$；由于 $\hat{g}$ 是 $g(\theta)$ 的无偏估计，所以 $E(\hat{g}) = 0$。从而 $E(T - \hat{g}) = E(T) - E(\hat{g}) = 0$，且 $\mathrm{Var}(T - \hat{g}) = \mathrm{Var}(T) + \mathrm{Var}(\hat{g}) + \mathrm{Cov}(T, \hat{g}) < \infty$（应该不存在方差有限但协方差无限的情况吧），所以根据判断准则，

$$
0 = \mathrm{Cov}(T, T - \hat{g}) = \mathrm{Var}(T) - \mathrm{Cov}(T, \hat{g})
$$

所以 $\mathrm{Cov}(T, \hat{g}) = \mathrm{Var}(T) > 0$。 □

5. 设总体 $p(x;\theta)$ 的费希尔信息量存在，若二阶导数 $\dfrac{\partial^2}{\partial\theta^2}p(x;\theta)$ 对一切的 $\theta \in \Theta$ 存在，证明费希尔信息量

$$
I(\theta) = -E\left(\frac{\partial^2}{\partial\theta^2}\ln p(x;\theta)\right)
$$

证明：

$$
\begin{aligned}
-E\left(\frac{\partial^2}{\partial\theta^2}\ln p(x;\theta)\right) &= -\int_{-\infty}^{+\infty}\frac{\partial^2 \ln p(x;\theta)}{\partial\theta^2}p(x;\theta)\mathrm{d}x \\
&=\int_{-\infty}^{+\infty}\left(\frac{\partial \ln p(x;\theta)}{\partial\theta}\right)^2 p(x;\theta)\mathrm{d}x \\
&=\int_{-\infty}^{+\infty}\frac{\partial \ln p(x;\theta)}{\partial\theta}\frac{\partial \ln p(x;\theta)}{\partial p(x;\theta)}\frac{\partial p(x;\theta)}{\partial\theta}p(x;\theta)\mathrm{d}x \\
&=\int_{-\infty}^{+\infty}\frac{\partial \ln p(x;\theta)}{\partial\theta}\frac{1}{p(x;\theta)}\frac{\partial p(x;\theta)}{\partial\theta}p(x;\theta)\mathrm{d}x \\
&=\int_{-\infty}^{+\infty}\frac{\partial \ln p(x;\theta)}{\partial\theta}\frac{\partial p(x;\theta)}{\partial\theta}\mathrm{d}x \\
&=E_x\left(\frac{\partial \ln p(x;\theta)}{\partial\theta}\right)^2 = I(\theta)
\end{aligned}
$$

□

6. 设总体密度函数为 $p(x;\theta)=\theta x^{\theta-1}, 0<x<1, \theta>0,\ x_1,x_2,\cdots,x_n$ 是样本。

(1) 求 $g(\theta)=\dfrac{1}{\theta}$ 的最大似然估计；

解：

对数似然函数为

$$\ln L(\theta)=\sum_{i=1}^n \ln\theta x^{\theta-1}=\sum_{i=1}^n(\ln\theta+(\theta-1)\ln x_i)=n\ln\theta+(\theta-1)\sum_{i=1}^n\ln x_i$$

对 θ 求导并令其为 0，

$$\frac{\partial L(\theta)}{\partial\theta}=\frac{n}{\theta}+\sum_{i=1}^n x_i=0$$

则 θ 的最大似然估计为 $\hat\theta=\dfrac{n}{\sum_{i=1}^n x_i}$，根据最大似然估计的不变性，$g(\theta)=\dfrac{1}{\theta}$ 的最大似然估计为

$$\widehat{g(\theta)}=\frac{1}{\hat\theta}=\frac{1}{n}\sum_{i=1}^n x_i$$

□

(2) 求 $g(\theta)$ 的有效估计。

解：

可以猜测上一小题中的 $\widehat{g(\theta)}=\frac{1}{n}\sum_{i=1}^n x_i$ 为有效估计，接下来验证一下。

可以计算得到总体的方差为 $\dfrac{1}{\theta^2}$，因此 $g(\hat\theta)=\bar x$ 的方差为 $\dfrac{1}{n\theta^2}$。

由于 $\ln p(x;\theta)=\ln\theta+(\theta-1)\ln x$,

$$\frac{\partial \ln p(x;\theta)}{\partial\theta}=\frac{1}{\theta}+\ln x,\quad \frac{\partial^2 \ln p(x;\theta)}{\partial\theta^2}=-\frac{1}{\theta^2},\quad I(\theta)=-E\left(\frac{\partial^2 \ln p(x;\theta)}{\partial\theta^2}\right)=\frac{1}{\theta^2}$$

所以 $I(\frac{1}{\theta})=\theta^2$，所以

$$\mathrm{Var}(\widehat{g(\theta)})=\frac{1}{n\theta^2}=\frac{1}{I(\frac{1}{\theta})}=\frac{1}{I(g(\theta))}$$

因此 $\widehat{g(\theta)}=\frac{1}{n}\sum_{i=1}^n x_i$ 为 $g(\theta)$ 的有效估计。 $\square$

7. 设总体密度函数为 $p(x;\theta)=\dfrac{2\theta}{x^3}e^{-\frac{\theta}{x^2}},x>0,\theta>0$，求 θ 的费希尔信息量 $I(\theta)$。

解：

$$\ln p(x;\theta)=\ln 2+\ln\theta-3\ln x-\frac{\theta}{x^2}$$
$$\frac{\partial\ln p(x;\theta)}{\partial\theta}=\frac{1}{\theta}-\frac{1}{x^2},\quad \frac{\partial^2\ln p(x;\theta)}{\partial\theta^2}=-\frac{1}{\theta^2}$$

所以

$$I(\theta)=-E\left(\frac{\partial^2\ln p(x;\theta)}{\partial\theta^2}\right)=\frac{1}{\theta^2}$$

$\square$

10. 设 $x_1,x_2,\cdots,x_n$ 是来自 $\mathrm{Ga}(\alpha,\lambda)$ 的样本，$\alpha>0$ 已知，试证明 $\dfrac{\bar{x}}{\alpha}$ 是 $g(\lambda)=\dfrac{1}{\lambda}$ 的有效估计，从而也是 UMVUE。

证明：

$$p(x;\lambda)=\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x},x>0;\quad \ln p(x;\lambda)=\alpha\ln\lambda-\ln\Gamma(\alpha)+(\alpha-1)\ln x-\lambda x$$
$$\frac{\partial\ln p(x;\lambda)}{\partial\lambda}=\frac{\alpha}{\lambda}-x;\quad \frac{\partial^2\ln p(x;\lambda)}{\partial\lambda^2}=-\frac{\alpha}{\lambda^2}$$

所以

$$I(\lambda)=-E\left(\frac{\partial^2\ln p(x;\lambda)}{\partial\lambda^2}\right)=\frac{\alpha}{\lambda^2};\quad \text{C-R 下界}=\frac{(g'(\lambda))^2}{nI(\lambda)}=\frac{(-\frac{1}{\lambda^2})^2}{n\frac{\alpha}{\lambda^2}}=\frac{1}{\alpha\lambda^2 n}$$

由于总体的方差为 $\dfrac{\alpha}{\lambda^2}$，所以 $\bar{x}$ 的方差为 $\dfrac{\alpha}{n\lambda^2}$，所以 $\dfrac{\bar{x}}{\alpha}$ 的方差为 $\dfrac{1}{n\alpha\lambda^2}$，等于 C-R 下界。 $\square$