第 13 周讲稿 UMVUE 和有效估计

UMVUE 和有效估计

★ θ^* 是 θ 的无偏估计，若对于 θ 的任意一个无偏估计量 θ，有 $D\theta^* \le D\theta$，则 θ^* 是 θ 的最小方差无偏估计，记 **MVUE** 或 **UMVUE**（**uniformly minimum-variance unbiased estimator**）

注：这里一致是指 $\forall\theta\in\Theta$。

★ 重要的定理（**Rao-Blackwell** 定理）

定理：随机变量 X 的方差存在，令 $\varphi(Y)=E(X\mid Y)$，则

$$E[\varphi(Y)]=EX, D[\varphi(Y)]\le DX$$

且等号成立当且仅当 $P(X=\varphi(Y))=1$。

证明：不妨设 $EX=0$，则

$$\begin{aligned}
&D[\varphi(Y)]=E[\varphi(Y)]^2\\
&=E[E(X\mid Y)]^2\\
&=E\{[E(X\mid Y)][E(X\mid Y)]\}\\
&=E[E(XE(X\mid Y)\mid Y)]\\
&=E(XE(X\mid Y))\\
&\le\sqrt{EX^2E[E(X\mid Y)]^2}
\end{aligned}$$

从而 $D[\varphi(Y)]=E[E(X\mid Y)]^2\le EX^2=DX$

注：可以直接利用 $DX=D[E(X\mid Y)]+E[D(X\mid Y)]$。

★ 推论：若 $T=T(X_1,X_2,\cdots,X_n)$ 为 θ 的充分统计量，$\hat\theta$ 为 θ 的任一无偏估计量，则 $\tilde\theta=E(\theta\mid T$ 也为 θ 的无偏估计量，且 $D(\tilde\theta)\le D\,\theta$ 。

例：设 $X_1,X_2,\cdots,X_n$ 是总体 $X\sim B(1,p), 0<p<1$（即 **Bernoulli** 分布）的一个样本，显然估计量 X_1 是 p 的无偏估计。我们用 **Rao-Blackwell** 定理求 p

的改进的无偏估计量。

由于$T=\sum_{i=1}^{n}X_i$为p的充分统计量，由于X_1只取 **0** 和 **1** 两个值，故

$$E[X_1|T=t]=P(X_1=1|T=t)$$
$$=\frac{P(X_1=1,T=t)}{P(T=t)}$$
$$=\frac{P(X_1=1,X_2+X_3+\cdots+X_n=t-1)}{P(T=t)}$$
$$=\frac{P(X_1=1)P(X_2+X_3+\cdots+X_n=t-1)}{P(T=t)}$$
$$=\frac{pC_{n-1}^{t-1}p^{t-1}(1-p)^{n-1-(t-1)}}{C_n^t p^t(1-p)^{n-t}}$$
$$=\frac{t}{n}$$

故由 **Rao-Blackwell** 定理知

$E[X_1|T]=\dfrac{T}{n}(=\bar{X})$是$p$的改进的无偏估计量。

注：利用对称性直接可知，$E[X_1|T]=E[X_1|\sum_{i=1}^{n}X_i]=\dfrac{\sum_{i=1}^{n}X_i}{n}$。

★ **UMVUE** 的一个判别准则：零无偏估计定理

如果 **UMVUE** 存在，则由 **Rao-Blackwell** 定理的推论知，它一定是充分统计量的函数。

零无偏估计量是指期望为 **0** 的统计量。

定理（零无偏估计定理）如果$\hat{\theta}$是θ的方差有限的无偏估计量，且对任何零无偏估计量$\varphi=\varphi(X_1,\cdots,X_n)$，都有

$$Cov_{\theta}(\hat{\theta},\varphi)=0,\forall\theta\in\Theta$$

则$\hat{\theta}$是θ的 **UMVUE**。

证明：对θ的任意一个无偏估计量$\tilde{\theta}$，显然$\varphi\triangleq\tilde{\theta}-\hat{\theta}$为零无偏估计量，且

$$D(\tilde{\theta}) = D(\hat{\theta}+\varphi) = D(\hat{\theta}) + D(\varphi) + 2Cov(\hat{\theta},\varphi)$$
$$= D(\hat{\theta}) + D(\varphi) \ge D(\hat{\theta}), \forall \theta \in \Theta$$

例：设 $X \sim E(\frac{1}{\theta})$，参数$\theta$ 未知，$X_1, X_2, \ldots, X_n$ 是其容量为 n 的样本. 则 $\bar{X}$ 为θ 的 UMVUE。

解：由因子分解定理，$T = T(X_1, X_2, ..., X_n) = \sum_{i=1}^{n} X_i$ 为θ 的充分统计量，

$\bar{X} = \frac{T}{n}$ 为θ 的无偏估计量。设任何零无偏估计量$\varphi = \varphi(X_1, \cdots, X_n)$，由于

$$E\varphi = \int_0^\infty \cdots \int_0^\infty \varphi(x_1, \cdots, x_n) \prod_{i=1}^{n} [\frac{e^{-\frac{x_i}{\theta}}}{\theta}] dx_1 \cdots dx_n = 0$$

即 $\int_0^\infty \cdots \int_0^\infty \varphi(x_1, \cdots, x_n) e^{-\sum_{i=1}^{n} \frac{x_i}{\theta}} dx_1 \cdots dx_n = 0$

两边对θ 求导得

$$\int_0^\infty \cdots \int_0^\infty \frac{n\bar{x}}{\theta^2} \varphi(x_1, \cdots, x_n) e^{-\sum_{i=1}^{n} \frac{x_i}{\theta}} dx_1 \cdots dx_n = 0$$

即 $E(\bar{X}\varphi) = 0$，即 $Cov(\bar{X}, \varphi) = 0$，故 $\bar{X}$ 为θ 的 **UMVUE**。

Cremer-Rao 不等式和有效估计

★ **Fisher** 信息量：

设总体的密度函数（或 **pmf**）$f(x;\theta)$，$\theta\in\Theta$满足下列条件：

1) 参数空间Θ是直线上的一个开区间；

2) 支撑$S=\{x: f(x;\theta)>0\}$与θ无关；

3) 导数$\dfrac{\partial}{\partial\theta}f(x;\theta)$对一切$\theta\in\Theta$都存在；

4) 对 $f(x;\theta)$，积分与微分运算可交换次序，即

$$\frac{\partial}{\partial\theta}\int_{-\infty}^{\infty}f(x;\theta)dx=\int_{-\infty}^{\infty}\frac{\partial}{\partial\theta}f(x;\theta)dx,$$

5) 期望$I(\theta)=E\left[\dfrac{\partial}{\partial\theta}\ln f(X;\theta)\right]^2$存在。

则称该期望$I(\theta)$为总体分布的 **Fisher** 信息量。

注：如果二阶导数对一切$\theta\in\Theta$都存在，则$I(\theta)$还可以用下式计算

$$I(\theta)=-E\left[\frac{\partial^2}{\partial\theta^2}\ln f(X;\theta)\right]$$

★ **Cremer-Rao** 不等式

设$X_1,X_2,\cdots,X_n$ 为取自具有 **pdf**（或 **pmf**）$f(x;\theta)$，$\theta\in\Theta=\{\theta:a<\theta<b\}$的总体$X$的一个样本，$a$，$b$为已知常数，$a$可以取$-\infty$，$b$可以取$+\infty$。又$\eta=\eta(X_1,X_2,\cdots,X_n)$是$g(\theta)$的一个无偏估计，且满足正则条件：

1) 集合$\{x: f(x;\theta)>0\}$与θ无关；

2) $g'(\theta)$与$\dfrac{\partial f(x;\theta)}{\partial\theta}$存在，且对一切$\theta\in\Theta$，

$$\frac{\partial}{\partial\theta}\int f(x;\theta)dx=\int\frac{\partial f(x;\theta)}{\partial\theta}dx,$$

$$\begin{aligned}&\frac{\partial}{\partial\theta}\int\int\cdots\int\eta(x_1,x_2,\cdots,x_n)f(x_1;\theta)f(x_2;\theta)\cdots f(x_n;\theta)dx_1dx_2\cdots dx_n\\&=\int\int\cdots\int\eta(x_1,x_2,\cdots,x_n)\frac{\partial}{\partial\theta}\left[\prod_{i=1}^{n}f(x_i;\theta)\right]dx_1dx_2\cdots dx_n;\end{aligned}$$

3) 令$I\left(\theta\right)=E_{\theta}\left(\dfrac{\partial \ln f\left(X;\theta\right)}{\partial\theta}\right)^{2}>0$，

则 $D_{\theta}\eta\geq\dfrac{\left[g'\left(\theta\right)\right]^{2}}{nI\left(\theta\right)}$ （*）

并且存在一个有可能依赖于θ但不依赖于$X_1,X_2,\cdots,X_n$的数K，使得等式

$\sum_{i=1}^{n}\dfrac{\partial\ln f\left(X_i;\theta\right)}{\partial\theta}=K\left(\eta-g\left(\theta\right)\right)$以概率 **1** 成立，以上这个条件为式（*）中等式成立的充要条件。特别地当$g\left(\theta\right)=\theta$时，不等式（*）化为

$$D_{\theta}\eta\geq\frac{1}{nI\left(\theta\right)}.$$

[注]1.若ξ是离散型随机变量，$f(x;\theta)$则表示为$P(\xi=x;\theta)$，相应的积分号改为求和号。

2.在使用 **R-C** 不等式时可不必验证 **2**）是否成立，因为在一般情况下，当 **1**）成立时 **2**）自动满足。

3. 满足 **1**）、**2**）假定的估计量称为正则估计。

4. R-C 下界$\dfrac{\left[g'\left(\theta\right)\right]^{2}}{nI\left(\theta\right)}$不是所有无偏估计的下界，而是无偏估计类中一个子集——正则无偏估计的方差下界。

5. R-C 的重要作用—达到 **R-C** 下界的估计量一定是 **UMVUE.**反之不然。即 **UMVUE** 不一定达到 **R-C** 下界。

6. $I(\theta)$信息量的意义

当 $D\eta=\dfrac{1}{nI(\theta)}$ $I(\theta)$越大，$D(\eta)$越小估计精度高，而$I(\theta)$大，则认为模型本身所含的信息量较多，或者说θ易认识，所以可视$I(\theta)$反应了模型中含有信息的量。

7.若 **C-R** 不等式的等号成立（达到 **C-R** 下界），则称$\eta=\eta\left(X_1,X_2,\cdots,X_n\right)$是$g\left(\theta\right)$的有效估计，有效估计一定是 **UMVUE**。

8. 设T_1,T_2为θ的两个无偏估计，其方差均存在，称$eff_{\theta}(T_1\,|\,T_2)\triangleq\dfrac{D_{\theta}(T_2)}{D_{\theta}(T_1)}$

为T_1关于T_2的效率，称T_1比T_2有效，若$eff_\theta(T_1|T_2)>1$。

若T为θ的有效估计（即$D_\theta(T)=\dfrac{1}{nI(\theta)}$），则定义$\theta$的无偏估计$T_1$的效率为

$$eff_\theta(T_1)=eff_\theta(T_1|T)=\frac{D_\theta(T)}{D_\theta(T_1)}=\frac{1}{nI(\theta)D_\theta(T_1)}$$

故有效估计的效率为 1，任一无偏估计的效率不超过 1.

例 设样本$X_1,X_2,...,X_n$**是来自**$B(1,p)$**，求**p**的 UMVUE。**

解：设总体分布为$f(x;p)=p^x(1-p)^{1-x}$，$x=0,1,0<p<1$.

易证$\{f(x;p)|p\in p(0,1)\}$满足正则条件，因为

$$I(p)=E_p\left[\frac{\partial}{\partial p}\ln f(X,p)\right]^2=E_p\left[\frac{X-p}{p(1-p)}\right]^2=\frac{1}{p^2(1-p)^2}E_p(X-p)^2=\frac{1}{p(1-p)},$$

故$\dfrac{1}{nI(p)}=\dfrac{p(1-p)}{n}$为$p$的无偏估计的 **C-R** 下界；

$\bar{X}$作为p的无偏估计，有：$D_p(\bar{X})=\dfrac{1}{n^2}\sum_{i=1}^{n}D_p(X_i)=\dfrac{1}{n^2}np(1-p)=\dfrac{p(1-p)}{n}$,

$\bar{X}$达到了 **C-R** 下界，故$\bar{X}$为p的 **UMVUE**。

例 设总体X**的密度函数为**$f(x;\lambda)=\begin{cases}\lambda e^{-ax}, & x>0;\\ 0, & x\le 0;\end{cases}$**未知参数**$\lambda>0$**，总体的一个样本为**$(X_1,X_2,\cdots,X_n)$**，证明**$\bar{X}$**是**$\dfrac{1}{\lambda}$**的一个 UMVUE。**

证明：由指数分布的总体满足正则条件可得：

$$I(\lambda)=-E\left[\frac{\partial^2}{\partial\lambda^2}\ln f(X;\lambda)\right]=-E\left(\frac{-1}{\lambda^2}\right)=\frac{1}{\lambda^2},$$

$\dfrac{\left[\left(\frac{1}{\lambda}\right)'\right]^2}{nI(\lambda)}=\dfrac{\left[\frac{-1}{\lambda^2}\right]^2}{n\frac{1}{\lambda^2}}=\dfrac{1}{n\lambda^2}$为$\dfrac{1}{\lambda}$的无偏估计方差的 **C-R** 下界；

另一方面$E(\bar{X})=\dfrac{1}{\lambda}$，$V_{ar}(\bar{X})=\dfrac{1}{n\lambda^2}$；即$\bar{X}$的方差达到了 **C-R** 下界，

故$\bar{X}$是$\dfrac{1}{\lambda}$的一个 **UMVUE**。

完备统计量与 Lemann-Scheffe 定理

定义：称分布族$\{f_\theta(x):\theta\in\Theta\}$是完备的，若$\forall\theta\in\Theta$,

$$E_\theta(g(X))=0\Rightarrow P_\theta(g(X)=0)=1$$

定义：称统计量$T=T(X_1,X_2,\cdots,X_n)$为完备的，若T 的分布族是完备的。

注：完备分布族条件下，$E_\theta(g_1(X))=E_\theta(g_2(X))$ ，则有

$$P_\theta(g_1(X)=g_2(X))=1$$

例：二项分布族$\{B(n,p):0<p<1\}$（n已知）是完备分布族。

证明：若$E_p(g(X))=\sum_{k=0}^{n}g(k)C_n^k p^k(1-p)^{n-k}=0$， $0<p<1$

则$\sum_{k=0}^{n}g(k)C_n^k(\frac{p}{1-p})^k=0$， $0<p<1$

等式左面是$\frac{p}{1-p}$的n次多项式，故$g(k)=0,k=0,1,2,\cdots,n$，证毕。

★ $\prod_{i=1}^{n}f(x_i;\theta)=C(\theta)h(x_1,x_2,...,x_n)\exp\{b(\theta)T(x_1,x_2,...,x_n)\}$

则$T=T(x_1,x_2,...,x_n)$是θ的充分完备统计量

★ $\prod_{i=1}^{n}f(x_i;\theta)=C(\theta)h(x_1,x_2,...,x_n)\exp\{b_1(\theta)T_1(x_1,x_2,...,x_n)+b_2(\theta)T_2(x_1,x_2,...,x_n)\}$

则(T_1,T_2)是$\theta=(\theta_1,\theta_2)$的充分完备统计量

★ **Lehmann-Scheffe 定理：**

若 T 是θ的充分完备统计量，θ是θ的一个无偏估计，则$\theta^*=E(\theta|T)$为θ的惟一的 **UMVUE**

★ **UMVUE 的求解步骤：**

① 求出参数θ的充分完备统计量 **T**

② 求出$ET=g(\theta)$，则$\theta=g^{-1}(T)$是θ的一个无偏估计

或求出一个无偏估计，然后改写成用 **T** 表示的函数

③　综合，$E[g^{-1}(T)\big|T]=g^{-1}(T)$ 是 θ 的 **UMVUE**

或者：求出 θ 的矩估计或 ML 估计，再求效率，为 1 则必为 UMVUE

例：**设样本 $X_1,X_2,...,X_n$ 是来自 $B(1,p)$，求 p 的 UMVUE**

解：（**1**）由上例知，$T=\sum_{i=1}^{n}X_i\sim B(n,p)$ 为充分完备统计量（也是指数族分布），且 X_1 为 p 的无偏估计。由 **Lehmann-Scheffe** 定理，p 的 **UMVUE** 为

$$E(X_1|T)=E(X_1|\sum_{i=1}^{n}X_i)=\frac{1}{n}\sum_{i=1}^{n}X_i=\bar{X}$$

例：设样本 $X_1,X_2,...,X_n$ 是来自 $N(\theta,1)$，求 θ 的 UMVUE

解：$T=\bar{X}\sim N(\theta,1)$ 是充分完备统计量（指数族分布），且 X_1 为 θ 的无偏估计。由 **Lehmann-Scheffe** 定理，θ 的 **UMVUE** 为 $E(X_1|T)$。

由于 $\begin{pmatrix}X_1\\ \bar{X}\end{pmatrix}\sim N(\begin{pmatrix}\theta\\ \theta\end{pmatrix},\begin{pmatrix}1 & \frac{1}{n}\\ \frac{1}{n} & \frac{1}{n}\end{pmatrix})$，故

$$E(X_1|T)=E(X_1|\bar{X})=\theta+\frac{Cov(X_1,\bar{X})}{D\bar{X}}(\bar{X}-\theta)$$

$$=\theta+(\bar{X}-\theta)=\bar{X}$$

即 $\bar{X}$ 为 θ 的 **UMVUE**。