2.3 最小方差无偏估计和有效估计

一、最小方差无偏估计

二、有效估计

一、最小方差无偏估计

最小方差无偏估计在均方误差意义下是一种最优估计.如何寻求此种估计，非常有意义.

1 最小方差无偏估计的判别法

定理2.7 设$\hat{\theta}(X)$是θ的一个无偏估计，$D(\hat{\theta})<\infty$，若对任何满足条件：$EL(X)=0, DL(X)<\infty$的统计量$L(X)$,有

$$E(L(X)\hat{\theta}(X))=0$$

则$\hat{\theta}(X)$是θ的$MVUE$,其中$X=(X_1,X_2,\cdots,X_n)^T$.

证 设$\hat{\theta}_1(X)$是θ的任一个方差有限的无偏估计，

令$L(X)=\hat{\theta}_1(X)-\hat{\theta}(X)$

显然$EL(X)=E(\hat{\theta}_1(X)-\hat{\theta}(X))=0$,且方差有限，同时

$$D\hat{\theta}_1(X)=D(L(X)+\hat{\theta}(X))$$

$$=DL(X)+D\hat{\theta}(X)+2\operatorname{cov}(L(X),\hat{\theta}(X))$$

$$=DL(X)+D\hat{\theta}(X)+2\underline{\left(EL(X)\hat{\theta}(X)\text{-}EL(X)E\hat{\theta}(X)\right)}$$

$$=D[L(X)]+D[\hat{\theta}(X)]\geq D[\hat{\theta}(X)] \qquad \boxed{=\underbrace{EL(X)\hat{\theta}(X)}_{0}\text{-}\theta\underbrace{EL(X)}_{0}=0}$$

因而，$\hat{\theta}(X)$是θ的$MVUE$.

例1（p54例2. 20） 设$X=(X_1,X_2,\cdots,X_n)^T$是来自总体N(μ,σ^2)的一个样本，已知$\overline{X}$和S_n^{*2}是μ和σ^2的无偏估计，证明$\overline{X}$和S_n^{*2}分别是μ和σ^2的$MVUE$.

证 先证明$\overline{X}$是μ的$MVUE$.

设$L(X)$满足$EL(X)=0$,则

$$\int\cdots\int L\cdot\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\}\mathrm{d}x=0$$

两边关于μ求导，则

$$D\hat{\theta}_1(X)=D(L(X)+\hat{\theta}(X)) \qquad \textbf{(1)}$$

因而 $E(L(X)\overline{X})=0$ 得证。

再证明S_n^{*2}是σ^2的$MVUE$.

对 $$\int\cdots\int L\cdot\underline{\sum_{i=1}^{n}x_i}\cdot\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\}\,\mathrm{d}x=\mathbf{0}\quad\mathbf{(1)}$$

关于μ再求导数，则得

$$\int\cdots\int L\cdot\underline{\left(\sum_{i=1}^{n}x_i\right)^2}\cdot\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\}\,\mathrm{d}x=0\quad\mathbf{(2)}$$

对$\int\cdots\int L\cdot\mathbf{exp}\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\}\mathbf{d}x=0$关于$\sigma^2$

求导数，则

$$\int\cdots\int L\cdot\underline{\sum_{i=1}^{n}(x_i-\mu)^2}\,\mathbf{exp}\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\}\mathbf{d}x=0\quad\mathbf{(5)}$$

又由于$\sum_{i=1}^{n}(x_i-\overline{x})^2=\sum_{i=1}^{n}(x_i-\mu)^2-n(\overline{x}-\mu)^2$,可得

$$\sum_{i=1}^{n}(x_i-\overline{x})^2=\underline{\sum_{i=1}^{n}(x_i-\mu)^2}-\underline{n\overline{x}^2}+2n\underline{\mu}\overline{x}-n\underline{\mu}^2.$$

(5) (2) (1) 条件

$$\int\cdots\int L\cdot\sum_{i=1}^{n}(x_i-\overline{x})^2\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\}dx=0$$

因而 $E(L(X)S_n^{*2})=0.$ 得证。

由此例可看出，利用判别定理进行判别，非常复杂.

注 定理2.7只是*MVUE*的一个判别法，但无法寻求 *MVUE*.

为了寻求更好的方法，需借助充分统计量甚至是完备统计量的概念。

补充条件期望

定义　条件分布的数学期望（若存在）称为条件期望，其定义如下：

$$E(X|Y=y)=\begin{cases}\sum_i x_i P(X=x_i|Y=y), & (X,Y)\text{为离散型；}\\ \int_{-\infty}^{+\infty} xp(x|y)\mathrm{d}x, & (X,Y)\text{为连续型.}\end{cases}$$

$$E(Y|X=x)=\begin{cases}\sum_j y_j P(Y=y_j|X=x), & (X,Y)\text{为离散型；}\\ \int_{-\infty}^{+\infty} y\,p(y|x)\mathrm{d}y, & (X,Y)\text{为连续型.}\end{cases}$$

注意：①条件期望$E(X|Y=y)$是y的函数，它与无条件期望$E(X)$在计算和意义上都有区别.

②因为条件期望是条件分布的数学期望，因此它具有数学期望的一切性质. 如

$$E(aX_1+bX_2|Y=y)=aE(X_1|Y=y)+bE(X_2|Y=y)$$

当X_1与X_2相互独立时，

$$E(X_1X_2|Y=y)=E(X_1|Y=y)E(X_2|Y=y).$$

等等.

(3) $E\{\phi(X,Y)|Y=y\}=E\{\phi(X,y)|Y=y\}$

$E\{\psi(X)\phi(X,Y)|X\}=\psi(X)E\{\phi(X,Y)|X\}$

定理（**重期望公式**）设(X,Y)是二维随机变量且$E(X)$存在，则

$$E(X)=E(E(X|Y)).$$

说明：①重期望$E(E(X|Y))$公式的具体使用为

- 如果Y是离散型随机变量则$E(E(X|Y))$公式为

$$E(X)=\sum_j p(Y=y_j)E(X|Y=y_j);$$

如果Y是连续型随机变量则$E(E(X|Y))$公式为

$$E(X)=\int_{-\infty}^{+\infty} p_Y(y)E(X|Y=y)dy.$$

定理2.8 设总体X的分布函数为$F(x,\theta),\theta\in\Theta$是未知参数，$X=(X_1,X_2,\cdots,X_n)^T$是来自总体$X$的一个样本，如果$T=T(X_1,X_2,\cdots,X_n)$是$\theta$的充分统计量，$\hat{\theta}$是$\theta$的任一无偏估计，记$\hat{\theta}^*\overset{\Delta}{=}E(\hat{\theta}\,|\,T)$

则有
$$E\hat{\theta}^*=\theta,\qquad \theta\in\Theta,$$
$$D\hat{\theta}^*\le D\hat{\theta},\qquad \theta\in\Theta,$$

证明：由于$T=T(x_1, x_2, \ldots, x_n)$是充分统计量，故其条件期望与参数θ无关，从而条件期望

$\hat{\theta}^* = E(\hat{\theta} \mid T)$也与$\theta$无关，并且$\hat{\theta}^*$是T的函数，

因此，$\hat{\theta}^*$可作为θ的估计。

由于 $E\hat{\theta}^* = E[(E\hat{\theta} \mid T)] = E\hat{\theta} = \theta.$

则$\hat{\theta}^*$为θ的无偏估计。

又由于 $D\hat{\theta} = E(\hat{\theta}-\theta)^2 = E(\hat{\theta}-\hat{\theta}^* + \hat{\theta}^* - \theta)^2$

$$= E(\hat{\theta}-\hat{\theta}^*)^2 + E(\hat{\theta}^*-\theta)^2 + 2E[(\hat{\theta}-\hat{\theta}^*)(\hat{\theta}^*-\theta)]$$

$$= E(\hat{\theta}-\hat{\theta}^*)^2 + D(\hat{\theta}^*) + 2E[(\hat{\theta}-\hat{\theta}^*)(\hat{\theta}^*-\theta)]$$

$$E[(\hat{\theta}-\hat{\theta}^*)(\hat{\theta}^*-\theta)] = E\{E[(\hat{\theta}-\hat{\theta}^*)(\hat{\theta}^*-\theta) \mid T]\}$$

$$= E\{(\hat{\theta}^*-\theta)E[(\hat{\theta}-\hat{\theta}^*) \mid T]\}$$

$$= E\{(\hat{\theta}^*-\theta)[E(\hat{\theta} \mid T) - E(\hat{\theta}^* \mid T)]\}$$

$$= E\{(\hat{\theta}^*-\theta)[E(\hat{\theta} \mid T) - \hat{\theta}^*]\}$$

$$= E\{(\hat{\theta}^*-\theta)[\hat{\theta}^* - \hat{\theta}^*]\} = 0$$

$$\Longrightarrow \quad D(\hat{\theta}) = E(\hat{\theta}-\hat{\theta}^*)^2 + D(\hat{\theta}^*) \quad \ge D(\hat{\theta}^*)$$

注： 定理2.8提供了一种改善无偏估计的方法，即，一个无偏估计$\hat{\theta}$对充分统计量的条件期望将能导出一个新的无偏估计，且它的方差更小。并且此无偏估计还一定是θ 的充分统计量的函数。

? 什么时候这么构造来的无偏估计是MVUE呢？

定理2.9 设总体X的分布函数为$F(x,\theta),\theta\in\Theta$是未知参数，$X=(X_1,X_2,\cdots,X_n)^T$是来自总体$X$的一个样本，如果$T=T(X_1,X_2,\cdots,X_n)$是$\theta$的充分完备统计量，$\hat{\theta}$是$\theta$的任一无偏估计，记$\hat{\theta}^*\overset{\Delta}{=}E(\hat{\theta}\,|\,T)$则$\hat{\theta}^*$是$\theta$的唯一的$MVUE$.

证　设$\hat{\theta}_1$和$\hat{\theta}_2$是θ的任意两个无偏估计,T是充分统计量，由定理2.8可知

$$E_\theta[E(\hat{\theta}_1|T)]=E\hat{\theta}_1=\theta,\quad E_\theta[E(\hat{\theta}_2|T)]=E\hat{\theta}_2=\theta \quad (*)$$

及 $D_\theta[E(\hat{\theta}_1|T)]\le D\hat{\theta}_1,\quad D_\theta[E(\hat{\theta}_2|T)]\le D\hat{\theta}_2$

由此$(*)$式可得 $E_\theta[E(\hat{\theta}_1|T)-E(\hat{\theta}_2|T)]=0$

又由于T是完备统计量，由定义1.6

$$P\{E(\hat{\theta}_1|T)=E(\hat{\theta}_2|T)\}=1$$

即θ的充分无偏估计唯一，由定理2.8可知$\hat{\theta}^*=E(\hat{\theta}_1|T)$是$MVUE$.

注 最小方差无偏估计寻找方法

1、构造一个充分完备统计量$T(X_1,\cdots,X_n)$和一个

利用定理1.5寻找。

θ的无偏估计$\hat{\theta}$.

利用矩估计法或最
大似然估计法寻找。

2、计算$E(\hat{\theta}\,|\,T)$,即得θ的一个$MVUE$.

例如

$\bar{X}$是泊松分布λ的充分完备统计量，同时也是λ的

无偏估计，则$E(\bar{X}\,|\,\bar{X})=\bar{X}$是$\lambda$的一个$MVUE$.

例2(p56例2.21)设总体 $X \sim N(\mu,\sigma^2)$, μ,σ^2为未知参数, $X_1,X_2,\cdots,X_n$ 是来自X的一个样本, 求μ和σ^2的最小方差无偏估计量.

解 由例1.10可知

$T=(\bar{X},S_n^2)$是(μ,σ^2)的充分完备统计量，因而

$$\hat{\mu}=E(\bar{X}\mid T)=\bar{X},\quad \hat{\sigma}^2=E(S_n^{*2}\mid T)=S_n^{*2}$$

所以

$\bar{X},S_n^{*2}$分别是μ和σ^2的最小方差无偏估计.

例3(p56例2.22) 设总体 X 在 $[0,\theta]$ 上服从均匀分布,其中 $\theta(\theta>0)$ 未知,$(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的样本,求 θ 的最小方差无偏估计量.

解 首先寻求充分完备统计量，样本的联合分布为

$$L(\theta)=\prod_{i=1}^{n} f(x_i)=\begin{cases}\dfrac{1}{\theta^n}, & 0<x_{(1)}<\cdots<x_{(n)}<\theta \\ 0, & \text{其他}\end{cases}$$

$$=\theta^{-n} I_{(0,\theta)}(x_{(n)})$$

其中 $I_{(0,\theta)}(x)=1$ 当 $0<x<\theta$,显然 $X_{(n)}$ 是 θ 的充分统计量

再验证$X_{(n)}$是完备统计量。

可通过验证$X_{(n)}$的分布族是完备族来得到结论.

又由于$X_{(n)}$的分布密度为

$$f_{X_{(n)}}(x)=\begin{cases}\dfrac{n}{\theta^n}x^{n-1} & 0<x<\theta \\ 0 & \text{其他}\end{cases}$$

利用定义1.5，若对任意一个满足

$Eg(X)=0,\quad \forall\theta\in\Theta,$ 的$g(X)$,有

$$\int_{-\infty}^{+\infty} g(x)f_{X_{(n)}}(x)dx=\int_0^{\theta} g(x)\frac{n}{\theta^n}x^{n-1}dx=0,\quad \forall\theta\in\Theta.$$

即, $\int_0^{\theta} g(x)x^{n-1}dx=0,\quad \forall\theta\in\Theta.$

对于$\int_0^{\theta} g(x)x^{n-1}dx = 0, \quad \forall \theta \in \Theta.$

两边对θ求导，得$g(\theta)\theta^{n-1} = 0, \quad \forall \theta \in \Theta.$

在这里，$\Theta=\{\theta > 0\}$，因此，$g(\theta)=0$，当$\theta > 0$时.

因此，当$X > 0$时，$g(X)=0.$

从而$X_{(n)}$的分布是完备分布族。$X_{(n)}$是充分完备统计量。

又由于 $E(X_{(n)}) = \int_0^{\theta} \frac{n}{\theta^n} x^n dx = \frac{n}{n+1}\theta$

$$E\left(\frac{n+1}{n}X_{(n)}\right) = \theta$$

无偏估计

所以 $E\left(\frac{n+1}{n}X_{(n)} \mid X_{(n)}\right) = \frac{n+1}{n}X_{(n)}$是$\theta$的$MVUE$.

二、有效估计

问题：无偏估计的方差是否可以任意的小？是否有下界？Rao-Cramer不等式可回答此问题。

1、Rao-Cramer不等式

定理2.10 设Θ是实数轴上的一个开区间，总体X的分布密度为$f(x;\theta)$，$\theta\in\Theta$，$(X_1,X_2,\cdots X_n)^T$是来自总体X的一个样本，$T=T(X_1,X_2,\cdots,X_n)$是$g(\theta)$的一个无偏估计量，且满足条件：

(1) 集合$S=\{x\mid f(x;\theta)\neq 0\}$与$\theta$无关；

(2) $\dfrac{\partial f(x;\theta)}{\partial \theta}$存在且对$\Theta$中一切$\theta$有

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

$$\frac{\partial}{\partial \theta}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} T(x_1,x_2,\cdots,x_n)L(x,\theta)dx_1dx_2\cdots dx_n$$

$$=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} T(x_1,x_2,\cdots,x_n)\frac{\partial}{\partial \theta}L(x,\theta)dx_1dx_2\cdots dx_n$$

其中$L(x,\theta)=\prod_{i=1}^{n} f(x_i;\theta)$;

积分和微分可交换运算次序

(3) $I(\theta)=E\left(\frac{\partial \ln f(X;\theta)}{\partial \theta}\right)^2>0$

则对一切 $\theta\in\Theta$，有 $D(T(X))\ge\frac{(g'(\theta))^2}{nI(\theta)}$，其中

$\frac{(g'(\theta))^2}{nI(\theta)}$为Rao – Cramer下界。

特别是当$g(\theta)=\theta$时，有 $D(T(X))\ge\frac{1}{nI(\theta)}$.

结论：无偏估计量的方差不可以无限小，存在下界。当其方差达到下界，它一定是MVUE. 但最小方差无偏估计不一定达到下界.

证　由统计量$T(X)$的无偏性可知：

$$ET(X) = \int T(x)L(x,\theta)\mathrm{d}x = g(\theta)$$

因而 $$\int T(x)\frac{\partial L(x,\theta)}{\partial\theta}\mathrm{d}x = g'(\theta) \qquad (1)$$

又由于 $$\int L(x,\theta)\mathrm{d}x = 1$$

$$\Longrightarrow g(\theta)\int \frac{\partial L(x,\theta)}{\partial\theta}\mathrm{d}x = 0 \qquad (2)$$

联立(1),(2) $\Longrightarrow$ $$\int (T(x)-g(\theta))\frac{\partial L(x,\theta)}{\partial\theta}\mathrm{d}x = g'(\theta)$$

改写 $$\int (T(x)-g(\theta))\frac{\partial L(x,\theta)}{\partial \theta}\mathrm{d}x = g'(\theta)$$

$$\Rightarrow \quad g'(\theta)=\int \{(T(x)-g(\theta))\sqrt{L(x,\theta)}\}\{\frac{\sqrt{L(x,\theta)}}{L(x,\theta)}\frac{\partial L(x,\theta)}{\partial \theta}\}\mathrm{d}x$$

由施瓦兹不等式可知 $E|XY| \le \sqrt{E(|X|^2)E(|Y|^2)}$

$$(g'(\theta))^2 \le \int \{(T(x)-g(\theta))^2 L(x,\theta)\mathrm{d}x$$

$$\times \int \frac{1}{L(x,\theta)}(\frac{\partial L(x,\theta)}{\partial \theta})^2 \mathrm{d}x$$

$$= D(T(X))\times \int \frac{1}{L(x,\theta)}(\frac{\partial L(x,\theta)}{\partial \theta})^2 \mathrm{d}x$$

$$= D(T(X))\times \int (\frac{\partial \ln L(x,\theta)}{\partial \theta})^2 L(x,\theta)\mathrm{d}x$$

$$= D(T(X)) \times \int \left(\frac{\partial \ln L(x,\theta)}{\partial \theta}\right)^2 L(x,\theta)\mathrm{d}x$$

$$(g'(\theta))^2 \le D(T(X)) \times E\left(\frac{\partial \ln L(x,\theta)}{\partial \theta}\right)^2$$

因而有

$$D(T(X)) \ge \frac{(g'(\theta))^2}{E\left(\frac{\partial \ln L(x,\theta)}{\partial \theta}\right)^2} \longrightarrow = nI(\theta)$$

其中$L(x,\theta) = \prod_{i=1}^{n} f(x_i;\theta)$;

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

证 $$E\left(\frac{\partial \ln L(x,\theta)}{\partial \theta}\right)^2 = nI(\theta)$$

$$E\left(\frac{\partial \ln L(x,\theta)}{\partial \theta}\right)^2 = E\left(\frac{\partial \sum_{i=1}^{n} \ln f(x_i,\theta)}{\partial \theta}\right)^2 = E\left(\sum_{i=1}^{n} \frac{\partial \ln f(x_i,\theta)}{\partial \theta}\right)^2$$

$$= E\left(\sum_{i=1}^{n} \left(\frac{\partial \ln f(x_i,\theta)}{\partial \theta}\right)^2 + 2\sum_{i=1}^{n}\sum_{j\neq i}^{n} \frac{\partial \ln f(x_i,\theta)}{\partial \theta}\frac{\partial \ln f(x_j,\theta)}{\partial \theta}\right)$$

$$= \underbrace{\sum_{i=1}^{n} E\left(\frac{\partial \ln f(x,\theta)}{\partial \theta}\right)^2}_{= nI(\theta)} + \underbrace{\left(2\sum_{i=1}^{n}\sum_{j\neq i}^{n} E\,\frac{\partial \ln f(x_i,\theta)}{\partial \theta}\frac{\partial \ln f(x_j,\theta)}{\partial \theta}\right)}_{0}$$

$$\text{当} i \neq j \text{时，} E\left(\frac{\partial \ln f(x_i,\theta)}{\partial \theta}\right)\left(\frac{\partial \ln f(x_j,\theta)}{\partial \theta}\right)$$

$$=E\left(\frac{\partial \ln f(x_i,\theta)}{\partial \theta}\right)E\left(\frac{\partial \ln f(x_j,\theta)}{\partial \theta}\right) \quad \text{独立}$$

$$=E\left(\frac{\partial \ln f(x_i,\theta)}{\partial \theta}\right)\int \frac{\partial \ln f(x_j,\theta)}{\partial \theta} f(x_j,\theta)\mathrm{d}x_j$$

$$=E\left(\frac{\partial \ln f(x_i,\theta)}{\partial \theta}\right)\int \frac{\partial f(x_j,\theta)}{\partial \theta}\mathrm{d}x_j = 0$$

$$\text{则有} \quad E\left(\frac{\partial \ln L(x,\theta)}{\partial \theta}\right)^2 = \sum_{i=1}^{n} E\left(\frac{\partial \ln f(x,\theta)}{\partial \theta}\right)^2 = nI(\theta)$$

$$\text{综上所述} \quad D(T(X)) \geq \frac{(g'(\theta))^2}{nI(\theta)}$$

2、Fisher信息量

设总体X的分布密度为$f(x,\theta), X=(X_1,X_2,\cdots,X_n)^T$为其样本，则称

$$\left.\begin{aligned} & I(\theta)=E\left(\frac{\partial \ln f(X;\theta)}{\partial\theta}\right)^2 \\ \text{或}\quad & I(\theta)=-E\left[\frac{\partial^2 \ln f(X;\theta)}{\partial\theta^2}\right]\end{aligned}\right\}\text{为Fisher信息量.}$$

例4(p58例2.23) 设$X_1, X_2, \cdots, X_n$是来自泊松分布$P(\lambda)$的一个样本，试求λ的无偏估计的方差下界。

解 由于$f(x;\lambda)=\dfrac{\lambda^x}{x!}e^{-\lambda}, E(\bar{X})=\lambda, D(\bar{X})=\dfrac{\lambda}{n}$,

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

$$=E\left(\frac{X}{\lambda}-1\right)^2=\frac{1}{\lambda^2}E(X-\lambda)^2=\frac{1}{\lambda^2}D(X)=\frac{1}{\lambda}$$

罗克拉美的下界为 $$\frac{1}{nI(\lambda)}=\frac{\lambda}{n}=D(\bar{X})$$

所以λ的最小方差无偏估计为$\bar{X}$.

例5(p58例2.24) 设 $X_1, X_2, \cdots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$的样本，试求 μ和σ^2的无偏估计的方差下界.

解 $$\ln f(x,\mu,\sigma^2) = -\ln\sqrt{2\pi} - \frac{1}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}(x-\mu)^2$$

$$I(\mu) = E\left\{\frac{\partial \ln f(X,\mu,\sigma^2)}{\partial\mu}\right\}^2 = E\,\frac{(X-\mu)^2}{\sigma^4} = \frac{1}{\sigma^2}$$

其罗克拉美的下界为 $\dfrac{1}{nI(\mu)} = \dfrac{\sigma^2}{n} = D(\bar{X})$

所以样本均值$\bar{X}$为μ的最小方差无偏估计.

又因为 $$\frac{\partial \ln f(x,\mu,\sigma^2)}{\partial \sigma^2} = -\frac{1}{2\sigma^2} + \frac{(x-\mu)^2}{2\sigma^4}$$

$$\frac{\partial^2 \ln f(x,\mu,\sigma^2)}{\partial (\sigma^2)^2} = \frac{1}{2\sigma^4} - \frac{(x-\mu)^2}{\sigma^6}$$

$$I(\sigma^2) = -E\{\frac{\partial^2 \ln f(X,\mu,\sigma^2)}{\partial (\sigma^2)^2}\} = \frac{E(X-\mu)^2}{\sigma^6} - \frac{1}{2\sigma^4} = \frac{1}{2\sigma^4}$$

其罗--克拉美下界为 $$\frac{1}{nI(\sigma^2)} = \frac{2\sigma^4}{n} < D(S_n^{*2}) = \frac{2\sigma^4}{n-1}$$

这是因为$D(\frac{(n-1)S_n^{*2}}{\sigma^2}) = 2(n-1)$,因而$D(S_n^{*2}) = \frac{2\sigma^4}{n-1}$.

即S_n^{*2}的方差达不到罗-克拉美下界。但是由例2知，S_n^{*2}是σ^2的最小方差无偏估计，这表明最小方差无偏估计量的方差不一定能够达到罗-克拉美下界。为此，引入有效估计的概念。

3、有效估计

定义2.8 设$\hat{\theta}$(或$T(X)$)是θ(或$g(\theta)$)的一个无偏估计，若

$$D(\hat{\theta})=\frac{1}{nI(\theta)}\text{（或}D(T(X))=\frac{[g'(\theta)]^2}{nI(\theta)}\text{）}$$

则称$\hat{\theta}$(或$T(X)$)是θ(或$g(\theta)$)的有效估计

定义2.9 设$\hat{\theta}$是θ的任一无偏估计，称$e(\hat{\theta})=\dfrac{(1/nI(\theta))}{D(\hat{\theta})}$

为估计量$\hat{\theta}$的效率。显然$0<e(\hat{\theta})\leq 1$.

定义2.10 如果θ的无偏估计量$\hat{\theta}$的效率满足$\lim\limits_{n\to\infty}e(\hat{\theta})=1$

则称$\hat{\theta}$为θ的渐近有效估计（量）。

如果$\hat{\theta}$是θ的有效估计，则它也是最小方差无偏估计。但反之却不成立。

例6 设$X_1, X_2, \cdots, X_n$是来自$N(\mu, \sigma^2)$的一个样本，证明：$\overline{X}$是μ的有效估计量；S_n^{*2}是σ^2的渐近有效估计量。

证 由信息量计算公式可知：

$$I(\mu) = E\left(\frac{\partial \ln f(X, \mu, \sigma^2)}{\partial \mu}\right)^2 = E\left(\frac{X-\mu}{\sigma^2}\right)^2$$

$$= \frac{1}{\sigma^2}$$

$$e(\bar{X})=\frac{1/(nI(\mu))}{D(\bar{X})}=\frac{\sigma^2/n}{\sigma^2/n}=1$$

$$I(\sigma^2)=-E\{\frac{\partial^2\ln f(X,\mu,\sigma^2)}{\partial(\sigma^2)^2}\}=\frac{E(X-\mu)^2}{\sigma^6}-\frac{1}{2\sigma^4}=\frac{1}{2\sigma^4}$$

$$D(\frac{(n-1)S_n^{*2}}{\sigma^2})=2(n-1)\text{，因此}D(S_n^{*2})=\frac{2\sigma^4}{n-1}$$

所以 $$e(S_n^{*2})=\frac{1/nI(\sigma^2)}{D(S_n^{*2})}=\frac{n-1}{n}\to 1$$

例7(p60例2.25) 设$X \sim B(N,p)$，$X_1, X_2, \cdots, X_n$为总体X的一个样本，试证$\hat{p} = \frac{1}{N}\bar{X}$是$p$的有效估计量。

证 $I(p) = E(\frac{\partial \ln f(X,p)}{\partial p})^2$

$$= E\left[\frac{\mathrm{d}[\ln C_N^X + X\ln p + (N-X)\ln(1-p)]}{\mathrm{d}p}\right]^2$$

$$= E(\frac{X}{p} - \frac{N-X}{1-p})^2 = \frac{1}{p^2(1-p)^2}E(X-Np)^2$$

$$= \frac{Np(1-p)}{p^2(1-p)^2} = \frac{N}{p(1-p)}$$

$D(\hat{p}) = \frac{D(\bar{X})}{N^2} = \frac{p(1-p)}{Nn}$ $\quad e(\hat{p}) = \frac{1/(nI(p))}{D(\hat{p})} = 1$ 得证。

定理2.11 设总体X的分布函数为$F(x,\theta),\theta\in\Theta$是未知参数，$X=(X_1,X_2,\cdots,X_n)^T$是来自总体$X$的一个样本，如果$\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$是$\theta$的无偏估计量，则$\hat{\theta}$是$\theta$的有效估计的充分必要条件为：

1、$\hat{\theta}$是θ的充分估计量；

2、$$\frac{\partial \ln L(x,\theta)}{\partial\theta}=C(\theta)[\hat{\theta}(x_1,x_2,\cdots,x_n)-\theta]$$

其中$L(x,\theta)$是样本的联合分布密度，$C(\theta)$仅依赖参数θ.

证明从略。

定理表明，最大似然估计如果是充分统计量，则一定是有效估计量。

例8(p61例2.26) 设 X 服从两点分布$B(1, p)$, X_1, $X_2, \cdots, X_n$是来自X的一个样本,证明p的最大似然估计量是有效估计.

解 因为X的分布律为

$$P\{X = x\} = p^x(1-p)^{1-x} \quad (x = 0,1)$$

所以 p的似然函数为

$$L(p) = \prod_{i=1}^{n}\left(p^{x_i}(1-p)^{1-x_i}\right) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

$$\ln L(p)=\left(\sum_{i=1}^{n}x_i\right)\ln p+\left(n-\sum_{i=1}^{n}x_i\right)\ln(1-p),$$

$$\frac{\mathrm{d}}{\mathrm{dp}}\ln L(p)=\frac{\sum_{i=1}^{n}x_i}{p}-\frac{n-\sum_{i=1}^{n}(x_i)}{1-p}=\frac{n}{p(1-p)}(\bar{x}-p)=0$$

解得 p 的最大似然估计值 $\hat{p}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x},$

$$\frac{\mathrm{d}}{\mathrm{dp}}\ln L(p)=\frac{\sum_{i=1}^{n}x_i}{p}-\frac{n-\sum_{i=1}^{n}(x_i)}{1-p}=\frac{n}{p(1-p)}(\bar{x}-p)$$

$C(p)$

$\bar{\mathrm{X}}$是p的充分统计量。 有效估计量

Thank You!