第2.3节 最小方差无偏估计和有效估计

一、最小方差无偏估计

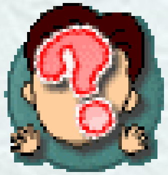
二、有效估计

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达到最优，是一种最优估计.如何寻求此种估计，将变得非常有意义.

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$X=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$.

证 设$\hat{\theta}_1(X)$是θ的一个无偏估计，令$L(X)=\hat{\theta}_1(X)-\hat{\theta}(X)$

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

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

$$=D[L(X)]+D[\hat{\theta}(X)]+2E\{[L(X)-EL(X)][(\hat{\theta}(X)-E\hat{\theta}(X)]\}$$

$$=D[L(X)]+D[\hat{\theta}(X)]\geq D[\hat{\theta}(X)]$$

因而，$\hat{\theta}(X)$是θ的$MVUE$.

注 **1** 此定理是最小方差无偏估计的判别法，但无法寻求最小方差无偏估计的存在性.

2 由于$L(X)$的任意性，因而很难利用定理判别.

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

证　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

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

因而　$E(L(X)\bar{X})=0$

故$\overline{X}$是μ的$MVUE$.

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

求二阶导数，则

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

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

求导数，则

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

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

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

因而 $E(L(X)S_n^{*2})=0$

故S_n^{*2}是σ^2的$MVUE$.

由此例可以看出，利用判别定理进行判别，非常复杂，况且也无法利用此定理去寻求MVUE.

充分完备统计量是解决上述困难的有力工具.

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$\hat{\theta}^*\stackrel{\Delta}{=}E(\hat{\theta}\mid T)$

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

即$\hat{\theta}^*$是θ的$MVUE$.

证明从略

由于$\hat{\theta}^{*}=E(\hat{\theta}\,|\,T)$仍然是充分统计量且作为$\theta$的估计量，可称之为充分估计量，上述定理表明，要寻找θ的最小方差无偏估计，只需在无偏的充分估计量类中寻找就足够了。假若θ的充分无偏估计量是唯一的，则这个充分无偏估计量一定是最小方差无偏估计量，那么在什么情况下，它才是唯一的呢？显然，如果它又是完备统计量，便可保证其唯一性。

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$\hat{\theta}^*\overset{\Delta}{=}E(\hat{\theta}\mid T)$

则$\hat{\theta}^*$是θ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

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

以及

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

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

又T是完备统计量，由完备统计量定义可知

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

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

注　最小方差无偏估计计算方法

1、构造一个充分完备统计量$T(X_1,\cdots,X_n)$和一个θ的无偏估计$\hat{\theta}$.

2、计算$E(\hat{\theta}\mid T)$,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$E(\bar{X}\mid\bar{X})=\bar{X}$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,$(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$是来自$X$的一个样本,求$\mu$和$\sigma^2$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

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

所以

$\bar{X}$和S_n^{*2}分别是μ和σ^2的最小方差无偏估计量.

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,$(X_1, X_2, \cdots, X_n)^{\mathrm{T}}$ 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

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

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

其中 $I_{(0,\theta)}(x)=1$ 当 $0<x<\theta$.

由因子分解定理，$X_{(n)}$是θ 的充分统计量，其分布密度为

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

利用完备分布族定义可以验证该分布族具有完备性.

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

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

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

二、有效估计

上一节介绍了最小方差无偏估计以及相应的寻求方法。自然会引入另一个问题：无偏估计的方差是否可以任意的小？是否有下界？事实上，**Rao-Cramer不等式**可以回答此问题。

1、Fisher信息量

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

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

为Fisher信息量.

Fisher信息量的另外一种表达式为：

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

2、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$f(x;\theta)$，$\theta \in \Theta$，$(X_1, X_2, \cdots X_n)^{\mathrm{T}}$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$L(x,\theta) = \prod_{i=1}^{n} f(x_i;\theta)$;

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

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

$\frac{(g'(\theta))^2}{nI(\theta)}$为罗－克拉美下界，$I(\theta)$称为*Fisher*信息量。

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

由此可见，统计量的方差不可以无限的小，存在下界。当无偏估计的方差达到下界，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

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

又由于 $\int L(x,\theta)\mathrm{d}x=1$

因而 $$\int\frac{\partial L(x,\theta)}{\partial\theta}\mathrm{d}x=0$$

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

改写上式为

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

由施瓦兹不等式可知

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

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

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

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

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

因而有

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

又因为

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

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

这是因为

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

$$=E(\frac{\partial \ln f(x_i,\theta)}{\partial\theta})E(\frac{\partial \ln f(x_j,\theta)}{\partial\theta})$$

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

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

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

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

例4(p58例2.23) 设 $(X_1, X_2, \cdots, X_n)^{\mathrm{T}}$ 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
$$I(\lambda) = E\left[\frac{\partial \ln f(X,\lambda)}{\partial\lambda}\right]^2 = E\left(\frac{\partial (X\ln\lambda - \lambda - \ln X!)}{\partial\lambda}\right)^2$$

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

因此
$$D(\hat{\lambda}) \geq \frac{1}{nI(\lambda)} = \frac{\lambda}{n}$$

所以 $\bar{X}$ 是 λ 的最小方差无偏估计。

例5(p58例2.24) 设$(X_1, X_2, \cdots, X_n)^{\mathrm{T}}$是来自正态总体$N(\mu, \sigma^2)$的样本，试求$\mu$和$\sigma^2$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
$$\frac{\partial \ln f(x,\mu,\sigma^2)}{\partial \sigma^2} = -\frac{1}{2\sigma^2} + \frac{(x-\mu)^2}{2\sigma^4}$$

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

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

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

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

即S_n^{*2}的方差达不到罗-克拉美下界。但是由例2.20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

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

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$0<e(\hat{\theta})\le 1.$

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$(X_1, X_2, \cdots, X_n)^{\mathrm{T}}$是来自$N(\mu, \sigma^2)$的一个样本，证明：$\bar{X}$是$\mu$ 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 $$I(p) = E\left(\frac{\partial \ln f(X,p)}{\partial p}\right)^2$$

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

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

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

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

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，$X=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$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θ 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。

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

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

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

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

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

又因为 $\frac{\partial\ln L(x,p)}{\partial p}=C(p)[\hat{p}(x_1,x_2,\cdots,x_n)-p]$，其中

$C(p)=\frac{n}{p(1-p)}$，同时 $\overline{X}$ 是 p 的充分统计量，因而由

定理2.11可知，$\overline{X}$ 是 p 的有效估计.

Thank You!