第 14 周　参数点估计

14.1 参数的矩估计法

参数估计和假设检验

20 名某地区高中男生的身高数据（单位：cm）：

170.1 179.0 171.5 173.1 174.1 177.2 170.3 176.2 163.7 175.4

163.3 179.0 176.5 178.4 165.1 179.4 176.3 179.0 173.9 173.7

假设总体服从正态分布，确定正态分布的期望和方差。 参数估计

是否有充分的依据可以相信这组数据是来自正态总体？ 假设检验

****

假设总体的分布的形式已知，分布中含有未知参数θ，

$X_1,X_2,\cdots,X_n$是来自该总体的一个样本，

构造适当的统计量$\hat{\theta}=\hat{\theta}\left(X_1,X_2,\cdots,X_n\right)$，估计总体分布的参数$\theta$，

就称$\hat{\theta}$为参数θ的点估计。

最常用的两种点估计构造方法是矩估计法和极大似然估计法。

****

设一个盒子里装有一定量的白球和黑球，试估计其中黑球比例。

假定进行 10 次有放回的抽取，抽到 3 个黑球：黑球比例大约为 0.3

假设盒子中黑球比例为p，随机变量$X\sim\begin{pmatrix}1 & 0\\ p & 1-p\end{pmatrix}$

$X_1,X_2,\cdots,X_{10}$为来自总体X的样本，　　$E\left(X\right)=p\approx\bar{X}=\dfrac{X_1+X_2+\cdots+X_{10}}{10}$

$$\frac{x_1+\cdots+x_{10}}{10}=0.3$$

****

估计参数θ， $E\left(X^k\right)=f_k(\theta)\ \Rightarrow\ \dfrac{X_1^k+X_2^k+\cdots+X_n^k}{n}=f_k\left(\hat{\theta}\right)$

$$\Rightarrow \bar{\theta}=\bar{\theta}\left(X_1,X_2,\cdots,X_n\right)$$

如总体包含m个未知参数，则列m个方程求解。

取m个不同的数字特征(最常用的是原点矩、中心矩)，用m个参数表示的理论表达式；样本值近似理论值。矩估计法

矩估计不唯一 为了计算简单，尽可能用低阶矩。

****

例 14.1.1 总体服从泊松分布$P(\lambda)$，$X_1,X_2,\cdots,X_n$是来自该总体的一个样本，试求参数λ的矩估计量。

解 泊松分布的期望$E(X)=\lambda$，用样本均值$\bar{X}$近似期望，$\bar{X}\approx E(X)=\lambda$，

得λ的矩估计量为$\bar{\lambda}=\bar{X}$。

解法 2 泊松分布的方差$Var(X)=\lambda$，用样本方差S^2替换总体方差，

得λ的矩估计量为$\hat{\lambda}=S^2$。

****

例 14.1.2 总体服从指数分布$E(\lambda)$，$X_1,X_2,\cdots,X_n$是来自该总体的一个样本，试求参数λ的矩估计。

解 指数总体的期望$E(X)=\dfrac{1}{\lambda}$，用样本均值近似期望，$\bar{X}\approx E(X)=\dfrac{1}{\lambda}$

得到方程$\bar{X}=\dfrac{1}{\bar{\lambda}}$，求解得参数$\lambda$的矩估计量为$\bar{\lambda}=\dfrac{1}{\bar{X}}$。

****

例 14.1.3 设 $X_1, X_2, \cdots, X_n$ 是来自正态总体 $N\left(\mu, \sigma^2\right)$ 的样本，求参数 μ 和 σ^2 的矩估计量。

解 正态总体的期望 $E\left(X\right) = \mu$，方差 $Var\left(X\right) = \sigma^2$

用样本均值 $\bar{X}$ 和样本方差 S^2 近似总体期望和方差，得

μ 的矩估计量为 $\bar{\mu} = \bar{X}$

σ^2 的矩估计量为 $\overline{\sigma^2} = S^2$。

****

例 14.1.4 对于均匀总体 $U(a,b)$，设 $X_1, X_2, \cdots, X_n$ 是来自该总体的样本，试求参数 a,b 的矩估计量。

解 利用总体的期望与方差 $E\left(X\right) = \frac{a+b}{2}$，$Var\left(X\right) = \frac{\left(b-a\right)^2}{12}$，得到方程组

$$\begin{cases} \dfrac{\bar{a}+\bar{b}}{2} = \bar{X} \\ \dfrac{\left(\bar{b}-\bar{a}\right)^2}{12} = S^2 \end{cases}$$

解得参数 a,b 的矩估计量 $\bar{a} = \bar{X} - \sqrt{3}S$，$\bar{b} = \bar{X} + \sqrt{3}S$。

解法 2 考虑总体的 1 阶和 2 阶原点矩

$E\left(X\right) = \frac{a+b}{2}$, $E\left(X^2\right) = \frac{a^2+ab+b^2}{3}$，得到方程组

$$\begin{cases} \dfrac{\tilde{a}+\tilde{b}}{2} = \bar{X} \\ \dfrac{\tilde{a}^2+\tilde{a}\tilde{b}+\tilde{b}^2}{3} = \dfrac{X_1^2+\cdots X_n^2}{n} \end{cases}$$

解得参数 a,b 的矩估计量 $\tilde{a}, \tilde{b}$。

****

14.2 参数的极大似然估计法

设一个盒子里装有一定量的白球和黑球，试估计其中黑球比例 p。

假定进行 10 次有放回的抽取，抽到 3 个黑球。黑球个数 $X \sim b(p,10)$

发生这一结果的概率 $P(X=3)=C_{10}^{3}p^{3}(1-p)^{7}$

$p=0.1$ 时，$P=0.0574$；　$p=0.4$ 时，$P=0.215$；　$p=0.3$ 时，　$P=0.2668$

极大似然估计方法的基本思想是以最大概率解释样本数据，

相对于其他参数，所考虑的样本数据更像是来自于这组参数。

****

极大似然估计(maximum likelihood estimation, MLE)的实现过程

1. 设总体分布的概率函数为 $f(x;\theta)$，$\theta \in \Theta$，其中 θ 是一组未知参数，Θ 称为参数空间，即参数 θ 可能取值的集合。

2. $x_1,x_2,\cdots,x_n$ 是来自该总体的样本观测值，则样本值发生的**联合概率函数**是关于 θ 的函数，用 $L(\theta;x_1,x_2,\cdots,x_n)$ 表示，简记为 $L(\theta)$，

$$L(\theta)=L(\theta;x_1,x_2,\cdots,x_n)=\prod_{k=1}^{n}f(x_k;\theta),$$

$L(\theta)$ 称为样本值的似然函数，

3. 函数 $\hat{\theta}(x_1,x_2,\cdots,x_n)$ 满足 $L(\hat{\theta})=\max_{\theta\in\Theta}L(\theta)$，则称统计量 $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 为参数 θ 的极大似然估计量。

****

极大似然的求解

$$L(\theta)=L(\theta;x_1,x_2,\cdots,x_n)=\prod_{k=1}^{n}f(x_k;\theta),\qquad L(\hat{\theta})=\max_{\theta\in\Theta}L(\theta)$$

$\ln(L(\theta))$ 与 $L(\theta)$ 达到最大值时，θ 取值相同

$$\ln(L(\theta))=\ln\prod_{k=1}^{n}f(x_k;\theta)=\sum_{k=1}^{n}\ln f(x_k;\theta)$$

$$\frac{d\ln L(\theta)}{d\theta}=\frac{d}{d\theta}\sum_{k=1}^{n}\ln f(x_k;\theta)=\sum_{k=1}^{n}\frac{f'(x_k;\theta)}{f(x_k;\theta)},\quad 求解\ \sum_{k=1}^{n}\frac{f'(x_k;\theta)}{f(x_k;\theta)}=0$$

若似然函数不可导时，则不能使用这种方法。

****

例 14.2.1　总体服从泊松分布 $P(\lambda)$，用极大似然估计法估计参数 λ。

解 总体的分布率为 $P(X=j)=\frac{\lambda^j}{j!}\cdot e^{-\lambda}$，$j=0,1,2,\cdots$

设 $x_1,x_2,\cdots,x_n$ 为来自该总体的样本观测值，则似然函数

$L(\lambda)=\prod_{k=1}^{n}P(X=x_k)=\prod_{k=1}^{n}\frac{\lambda^{x_k}}{x_k!}\cdot e^{-\lambda}$，两端取对数，得对数似然函数

$$\ln L(\lambda)=\ln\left(\prod_{k=1}^{n}\frac{\lambda^{x_k}}{x_k!}\cdot e^{-\lambda}\right)=\sum_{k=1}^{n}\ln\left(\frac{\lambda^{x_k}}{x_k!}\cdot e^{-\lambda}\right)=\sum_{k=1}^{n}\left(x_k\ln\lambda+(-\lambda)-\ln\prod x_k!\right)$$

$$=\sum_{k=1}^{n}x_k\cdot\ln\lambda-n\lambda-n\ln\prod x_k!=n\bar{x}\ln\lambda-n\lambda-n\ln\prod x_k!$$

将上式对 λ 求导，并令其等于 0

$$\frac{d(\ln L(\lambda))}{d\lambda}=\frac{d\left(n\bar{x}\ln\lambda-n\lambda-n\ln\prod x_k!\right)}{d\lambda}=\frac{n\bar{x}}{\lambda}-n=0$$

解得 $\lambda=\bar{x}$，根据题意知必为对数似然函数的最大值点，所以参数 λ 的极大似然估计量为 $\hat{\lambda}=\bar{X}$。

****

例 14.2.2 总体服从指数分布 $E(\lambda)$，求参数 λ 的极大似然估计量。

解 设 $x_1,x_2,\cdots,x_n$ 为来自该总体的样本观测值，

指数分布的密度函数为 $f(x)=\lambda\cdot e^{-\lambda x}\ (x>0)$，

则似然函数 $L(\lambda)=\prod_{k=1}^{n}\lambda\cdot e^{-\lambda x_k}=\lambda^n\cdot e^{-\lambda(x_1+\cdots+x_n)}$，

对数似然函数 $\ln(L(\lambda))=n\ln\lambda-\lambda(x_1+\cdots+x_n)$，对 λ 求导，并令其等于 0

$$\frac{d\ln(L(\lambda))}{d\lambda}=\frac{n}{\lambda}-(x_1+\cdots+x_n)=0,$$

解得 $\lambda=\dfrac{n}{x_1+\cdots+x_n}=\dfrac{n}{n\bar{x}}=\dfrac{1}{\bar{x}}$，根据题意知必为对数似然函数的最大值点，

所以参数 λ 的极大似然估计量为 $\hat{\lambda}=\dfrac{1}{\bar{X}}$。

****

例 14.2.3 设 $X_1,X_2,\cdots,X_n$ 是正态总体 $N(\mu,\sigma^2)$ 的样本，其中 σ^2 已知，求参数 μ 的极大似然估计量。

解：设 $x_1,x_2,\cdots,x_n$ 为样本观测值，

则似然函数 $L(\mu)=\prod_{k=1}^{n}f(x_k;\mu)=\prod_{k=1}^{n}\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_k-\mu)^2}{2\sigma^2}}=\left(\dfrac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$，

对数似然函数为

$$\ln L(\mu)=\ln\left(\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}\right)=n\ln\left(\frac{1}{\sqrt{2\pi}}\right)-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2,$$

将上式对 μ 求导，并令其等于 0

$$\frac{d\ln L(\mu)}{d\mu}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)=0$$

解得 $\mu=\dfrac{x_1+\cdots+x_n}{n}=\bar{x}$，根据题意知必为对数似然函数的极大值点，所以参数 μ 的极大似然估计量为 $\hat{\mu}=\bar{X}$。

****

例 14.2.4 $X_1,X_2,\cdots,X_n$ 为来自均匀分布总体 $U(a,b)$ 的样本，试利用极大似然估计给出参数 a,b 的估计量。

解：设 $x_1,x_2,\cdots,x_n$ 为样本观测值，

总体分布的密度函数为 $f(x;a,b)=\dfrac{1}{b-a}, a\le x\le b$

似然函数为 $L(a,b)=\prod_{k=1}^{n} f(x_k;a,b)=(\dfrac{1}{b-a})^n$， $a\le x_k\le b, k=1,2,\cdots,n$，

显然 $L(a,b)$ 关于 a 是单调递增函数，关于 b 是单调递减函数。使得 $L(a,b)$ 最大，必须使 $b-a$ 达到最小，即使 b 尽可能小，a 尽可能大。

考虑到 $a\le x_1,x_2,\cdots,x_n\le b$， $a\le \min\{x_1,x_2,\cdots,x_n\}$， $b\ge \max\{x_1,x_2,\cdots,x_n\}$，

参数 a,b 的极大似然估计 $\hat{a}=\min(X_1,X_2\cdots,X_n)$， $\widehat{b}=\max(X_1,X_2\cdots,X_n)$。

****

例 14.2.5 估计 Cauchy 分布 $f(x;\theta)=\dfrac{1}{\pi\left[1+(x-\theta)^2\right]}$ （$x\in R$）的参数 θ。

$$\int_{-\infty}^{+\infty}|x|\cdot f(x)dx=2\int_0^{+\infty}x\cdot f(x)dx=2\int_0^{+\infty}\frac{x}{\pi\left[1+(x-\theta)^2\right]}dx$$

$$=2\int_0^{+\infty}\frac{x-\theta}{\pi\left[1+(x-\theta)^2\right]}dx+2\int_0^{+\infty}\frac{\theta}{\pi\left[1+(x-\theta)^2\right]}dx$$

$$=\frac{1}{\pi}\int_0^{+\infty}\frac{d\left[1+(x-\theta)^2\right]}{\left[1+(x-\theta)^2\right]}+\theta=\frac{1}{\pi}\ln\left[1+(x-\theta)^2\right]\Big|_0^{+\infty}+\theta$$

Cauchy 分布随机变量的期望不存在，因此不能用矩估计法对参数 θ 进行估计。

****

例 14.2.5 估计 Cauchy 分布 $f(x;\theta)=\dfrac{1}{\pi\left[1+(x-\theta)^2\right]}$ （$x\in R$）的参数 θ。

解：设 $x_1, x_2, \cdots, x_n$ 为来自该总体的样本观测值，

似然函数 $L(\theta) = \prod_{k=1}^{n} f(x_k;\theta) = \prod_{k=1}^{n} \dfrac{1}{\pi\left[1+\left(x_k-\theta\right)^2\right]}$

对数似然函数 $\ln L(\theta) = \sum_{k=1}^{n} -\left(\ln\pi + \ln\left(1+\left(x_k-\theta\right)^2\right)\right)$

将上式对 θ 求导，并令其等于 0， $\dfrac{d\ln L(\theta)}{d\theta} = -\sum_{k=1}^{n} \dfrac{x_k-\theta}{1+\left(x_k-\theta\right)^2} = 0$

此方程无法的到解析解，需要用一定的计算方法近似求解。

****

14.3 参数点估计的无偏性和有效性

参数点估计的优良评价

无偏性定义 设 θ 为未知参数，$\hat{\theta} = \hat{\theta}\left(X_1, X_2, \cdots, X_n\right)$ 是 θ 的一个估计量，若有 $E\left(\hat{\theta}\right) = \theta$，则称 $\hat{\theta}$ 是参数 θ 的无偏估计量。

无偏性准则是希望保证估计量没有系统偏差。

泊松分布总体 $X \sim P(\lambda)$，$E\left(\bar{X}\right) = \lambda$，$E\left(S^2\right) = Var\left(X\right) = \lambda$

$E\left(\bar{X}\right) = E\left(X\right)$ ， $E\left(S^2\right) = \sigma^2$

****

如果估计量 $\hat{\theta}_1$ 的期望 $E\left(\hat{\theta}_1\right) = c\theta$，设 $\hat{\theta} = \dfrac{1}{c}\hat{\theta}_1$，则 $E\left(\hat{\theta}\right) = \theta$，

称 $\hat{\theta}$ 为对估计量 $\hat{\theta}_1$ 的无偏校正。

$X_1, X_2, \cdots, X_n$ 是来自均匀总体 $U(0,\theta)$ 的样本，可验证 $E\left(\max\limits_{1\le k\le n} X_k\right) = \dfrac{n}{n+1}\theta$，

统计量 $\max\limits_{1\le k\le n} X_k$ 不是参数 θ 的无偏估计，$\dfrac{n+1}{n}\max\limits_{1\le k\le n} X_k$ 即为无偏校正。

****

有效性定义　设$\hat{\theta}_1$，$\hat{\theta}_2$是参数θ的无偏估计，如果有$Var\left(\hat{\theta}_1\right)<Var\left(\hat{\theta}_2\right)$，

则称$\hat{\theta}_1$比$\hat{\theta}_2$有效。

有效性准则希望估计量围绕参数真值波动的幅度越小越好。

通过增加样本数量可以提高估计的有效性。例如样本均值 $Var\left(\bar{X}\right)=\frac{\sigma^2}{n}$

****

例 14. 3. 1 设总体$X\sim N\left(\mu,\sigma^2\right)$，$X_1,X_2,X_3,X_4$是一个样本。验证

$$\hat{\theta}_1=\bar{X}，\quad \hat{\theta}_2=2X_1+X_2-X_3-X_4，\quad \hat{\theta}_3=\frac{3}{8}\left(X_1+X_2\right)+\frac{1}{8}\left(X_3+X_4\right)$$

均为参数μ的无偏估计量，并分析哪一个更有效。

$$E\left(\hat{\theta}_1\right)=E\left(\bar{X}\right)=E\left(X\right)=\mu,$$

$$E\left(\hat{\theta}_2\right)=E\left(2X_1+X_2-X_3-X_4\right)=2\mu+\mu-\mu-\mu=\mu$$

$$E\left(\hat{\theta}_3\right)=\frac{3}{8}E\left(X_1+X_2\right)+\frac{1}{8}E\left(X_3+X_4\right)=\frac{3}{8}\cdot 2\cdot\mu+\frac{1}{8}\cdot 2\cdot\mu=\mu$$

$$Var\left(\hat{\theta}_1\right)=Var\left(\bar{X}\right)=Var\left(\frac{X_1+X_2+X_3+X_4}{4}\right)=\frac{1}{16}\cdot 4\cdot Var\left(X\right)=\frac{\sigma^2}{4},$$

$$Var\left(\hat{\theta}_2\right)=Var\left(2X_1+X_2-X_3-X_4\right)=7\cdot Var\left(X\right)=7\sigma^2$$

$$Var\left(\hat{\theta}_3\right)=Var\left[\frac{3}{8}\left(X_1+X_2\right)+\frac{1}{8}\left(X_3+X_4\right)\right]=\left[2\cdot\left(\frac{3}{8}\right)^2+2\cdot\left(\frac{1}{8}\right)^2\right]\cdot Var\left(X\right)=\frac{5}{16}\sigma^2$$

$$Var\left(\hat{\theta}_1\right)<Var\left(\hat{\theta}_3\right)<Var\left(\hat{\theta}_2\right)$$

****

14. 4 参数点估计应用实例

例 14.4.1 假设某小学的学生周末玩游戏的时间服从正态分布 $N\left(\mu,\sigma^2\right)$，下面是对 10 名该小学的小学生的抽样调查数据（单位：小时）

2.2, 3.5，0.8，1.5，2.4，
1.2，4.3，2.8，2.4，1.7

试估计该小学学生每天玩游戏时间的规律 $N\left(\mu,\sigma^2\right)$，$X\sim N\left(2.28,1.06^2\right)$

周末玩游戏超过 5 个小时的学生比例；

大约有多少比例的学生周末没有玩游戏的时间；

会不会有学生玩游戏的时间超过 10 小时 ……

****

例 14.4.2 $X_1,X_2,\cdots,X_n$ 是来自均匀总体 $U(0,\theta)$ 的样本，验证 $\hat{\theta}_1=2\bar{X}$ 和 $\hat{\theta}_2=\dfrac{n+1}{n}\max\limits_{1\le k\le n}X_k$ 都是参数 θ 的无偏估计，并比较它们的有效性。

解 均匀总体 $U(0,\theta)$ 的期望、方差分别为 $E\left(X\right)=\dfrac{\theta}{2}, Var\left(X\right)=\dfrac{\theta^2}{12}$

$E\left(\hat{\theta}_1\right)=E\left(2\bar{X}\right)=E\left(2X\right)=2\times\dfrac{\theta}{2}=\theta$，所以 $\hat{\theta}_1=2\bar{X}$ 是参数 θ 的无偏估计

$$Var(\hat{\theta}_1)=Var\left(2\bar{X}\right)=4Var\left(\bar{X}\right)=4\frac{Var\left(X\right)}{n}=4\frac{\theta^2}{12n}=\frac{\theta^2}{3n}$$

$\hat{\theta}_2=\dfrac{n+1}{n}\max\limits_{1\le k\le n}X_k$， 计算 $E\left(\hat{\theta}_2\right)$， $Var\left(\hat{\theta}_2\right)$

1. 记 $\tilde{\theta}=\max\limits_{1\le k\le n}X_k$， 计算 $\tilde{\theta}$ 的分布函数和密度函数

2. 计算 $\tilde{\theta}$ 的期望和方差

计算 $\tilde{\theta}=\max\limits_{1\le k\le n}X_k$ 的分布函数，

当 $0 \le y \le \theta$ 时，由样本的独立同分布性质，可知其分布函数为

$$F_{\tilde{\theta}}(y) = P\left(\tilde{\theta} \le y\right) = P\left(\max_{1\le k\le n} X_k \le y\right) = \prod_{k=1}^{n} P\left(X_k \le y\right) = \left(\frac{y}{\theta}\right)^n$$

$\tilde{\theta} = \max\limits_{1\le k\le n} X_k$ 的分布函数 $F_{\tilde{\theta}}(y) = \begin{cases} 0, y < 0 \\ \left(\dfrac{y}{\theta}\right)^n, 0 \le y \le \theta \\ 1, y > \theta \end{cases}$，

于是 $\tilde{\theta}$ 的概率密度为 $f_{\tilde{\theta}}(y) = \begin{cases} \dfrac{n}{\theta^n} y^{n-1}, y \in [0, \theta] \\ 0, \text{其它} \end{cases}$，　因此，我们有

$$E(\tilde{\theta}) = \int_{-\infty}^{+\infty} y f_{\tilde{\theta}}(y) dy = \int_0^{\theta} y \frac{n}{\theta^n} y^{n-1} dy$$

$$= \frac{n}{\theta^n} \int_0^{\theta} y^n dy = \frac{n}{\theta^n} \cdot \left.\frac{y^{n+1}}{n+1}\right|_0^{\theta} = \frac{n}{n+1}\theta$$

故 $E\left(\hat{\theta}_2\right) = E\left(\dfrac{n+1}{n} \max\limits_{1\le k\le n} X_k\right) = \dfrac{n+1}{n} E\left(\tilde{\theta}\right) = \theta$。

所以 $\hat{\theta}_2 = \dfrac{n+1}{n} \max\limits_{1\le k\le n} X_k$ 也是参数 θ 的无偏估计。

$$E(\tilde{\theta}^2) = \int_0^{\theta} y^2 f_{\tilde{\theta}}(y) dy = \int_0^{\theta} y^2 \frac{n}{\theta^n} y^{n-1} dy = \frac{n}{\theta^n} \int_0^{\theta} y^{n+1} dy = \frac{n}{\theta^n} \cdot \frac{\theta^{n+2}}{n+2} = \frac{n}{n+2}\theta^2,$$

$$Var\left(\tilde{\theta}\right) = E(\tilde{\theta}^2) - \left(E(\tilde{\theta})\right)^2 = \frac{n}{n+2}\theta^2 - \frac{n^2}{\left(n+1\right)^2}\theta^2 = \frac{n}{(n+1)^2(n+2)}\theta^2$$

因此 $Var(\hat{\theta}_2) = Var\left(\dfrac{n+1}{n}\tilde{\theta}\right) = \dfrac{(n+1)^2}{n^2} Var\left(\tilde{\theta}\right) = \dfrac{1}{n(n+2)}\theta^2$。

$$Var(\hat{\theta}_1) = Var\left(2\bar{X}\right) = 4Var\left(\bar{X}\right) = 4\frac{Var\left(X\right)}{n} = 4\frac{\theta^2}{12n} = \frac{\theta^2}{3n};$$

当 $n > 1$ 时，$\dfrac{1}{n(n+2)}\theta^2 < \dfrac{1}{3n}\theta^2$，所以 $\hat{\theta}_2 = \dfrac{n+1}{n} \max\limits_{1\le k\le n} X_k$ 比 $\hat{\theta}_1 = 2\bar{X}$ 更有效。

****