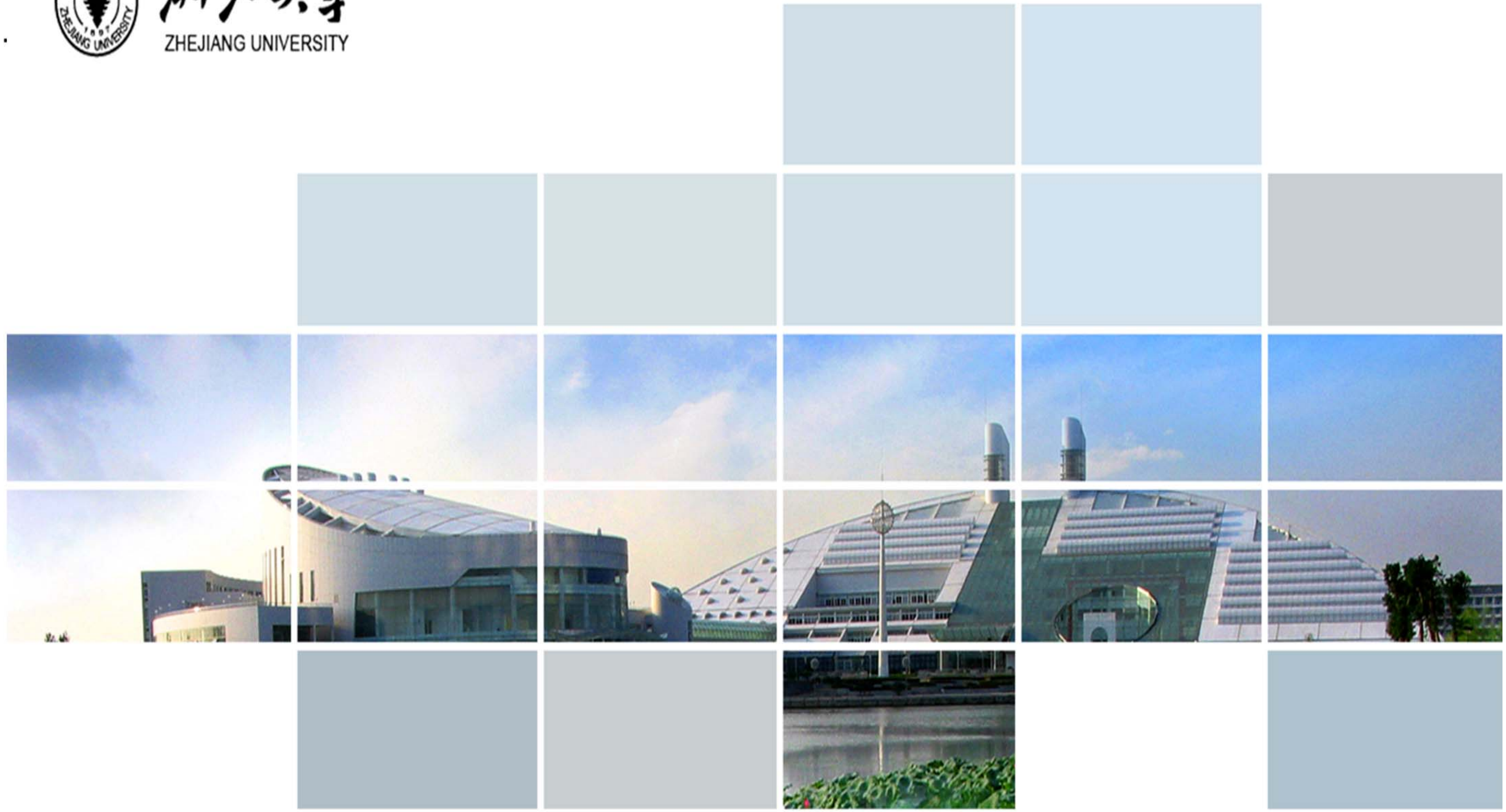
浙江大學
ZHEJIANG UNIVERSITY



第47讲　有效性，均方误差

2. 有效性准则

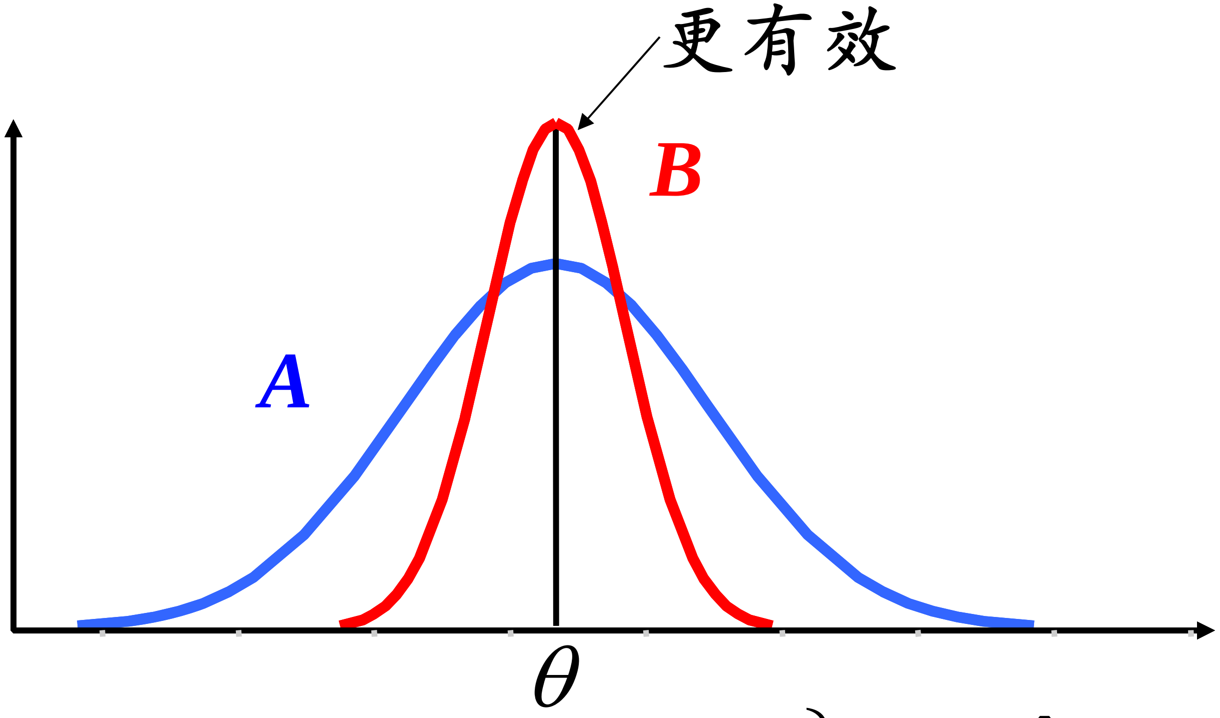
定义: 设$\hat{\theta}_1, \hat{\theta}_2$是$\theta$的两个无偏估计,

如果$D(\hat{\theta}_1) \le D(\hat{\theta}_2)$,对一切$\theta \in \Theta$成立,

且不等号至少对某一$\theta \in \Theta$成立, 则称

$\hat{\theta}_1$比$\hat{\theta}_2$有效.

方差较小的无偏估计量是一个更有效的估计量。



若$\hat{\theta}_1$的概率密度如红线所示,
若$\hat{\theta}_2$的概率密度如蓝线所示,

$D(\hat{\theta}_1) < D(\hat{\theta}_2)$,
$\hat{\theta}_1$比$\hat{\theta}_2$更有效.

例1： 设总体为$X, E(X)=\mu, D(X)=\sigma^2>0$，

$X_1,\cdots,X_n$是一样本. 对$1\le k\le n$，令

$\hat{\theta}_k=\frac{1}{k}(X_1+\ldots+X_k)$，即$\hat{\theta}_k$为前$k$个样本平均值.

显然，对$1\le k\le n$，$\hat{\theta}_k$均是参数μ的无偏估计。

问：这一系列估计量$\hat{\theta}_1,\hat{\theta}_2,\ldots,\hat{\theta}_n(n\ge 2)$中，哪个$\hat{\theta}_k$作为参数$\mu$的估计最有效?

解: $D(\hat{\theta}_k)=\dfrac{1}{k^2}\sum\limits_{i=1}^{k}D(X_i)=\dfrac{\sigma^2}{k}$,

即估计量的方差随着k的增加而减少,

$\therefore\ \hat{\theta}_n$ 最有效.

例2： 设总体$X\sim U[0,\theta]$, $X_1,\cdots,X_n$为一样本.

已知θ的两个无偏估计为

$$\hat{\theta}_1=2\bar{X},\qquad \hat{\theta}_2=\frac{n+1}{n}X_{(n)}.$$

试判别$\hat{\theta}_1$与$\hat{\theta}_2$哪个更有效$(n\geq 2\text{时})$?

解：$D\left(\hat{\theta}_1\right)=D\left(2\bar{X}\right)=4\cdot\frac{1}{n}\cdot\frac{\theta^2}{12}=\frac{\theta^2}{3n}$

根据第46讲中的例2，$X_{(n)}$的密度函数为

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

$$\Rightarrow E\left(X_{(n)}\right)=\int_0^\theta \frac{nx^n}{\theta^n}dx=\frac{n}{n+1}\theta,$$

$$E\left(X_{(n)}^2\right)=\int_0^\theta \frac{nx^{n+1}}{\theta^n}dx=\frac{n}{n+2}\theta^2$$

$$于是D\left(\hat{\theta}_2\right)=\frac{(n+1)^2}{n^2}\left\{E\left(X_{(n)}^2\right)-\left[E\left(X_{(n)}\right)\right]^2\right\}=\frac{\theta^2}{n(n+2)}$$

当$n\geq 2$时，

$$\therefore D\left(\hat{\theta}_1\right)=\frac{\theta^2}{3n}>\frac{\theta^2}{n(n+2)}=D\left(\hat{\theta}_2\right)$$

$\Rightarrow$ $\hat{\theta}_2$比$\hat{\theta}_1$更有效.

3. 均方误差准则

定义：设$\hat{\theta}$是参数θ的点估计，方差存在，则称

$E(\hat{\theta}-\theta)^2$是估计量$\hat{\theta}$的均方误差，记为$Mse(\hat{\theta})$.

若$\hat{\theta}$是θ的无偏估计，则有$Mse(\hat{\theta})=D(\hat{\theta})$.

设$\hat{\theta}_1,\hat{\theta}_2$是$\theta$的点估计，如果$Mse(\hat{\theta}_1)\le Mse(\hat{\theta}_2)$，

对一切$\theta\in\Theta$成立，且不等号至少对某一

$\theta\in\Theta$成立，则称在均方误差准则下$\hat{\theta}_1$优于$\hat{\theta}_2$.

在实际应用中，均方误差准则比无偏性准则更重要.

例3: 利用均方误差准则, 对用样本方差S^2和样本二阶中心矩B_2分别估计正态总体方差σ^2时进行评价.

解：在正态总体下，

$$(n-1)S^2/\sigma^2 \sim \chi^2(n-1).$$

又因S^2是σ^2的无偏估计，因此

$$Mse(S^2)=D(S^2)=\frac{2\sigma^4}{n-1}.$$ —第42讲例2

而 $$Mse(B_2)=E[(B_2-\sigma^2)^2]=\frac{2n-1}{n^2}\sigma^4$$

$$=D(B_2)+[E(B_2)-\sigma^2]^2$$

$$Mse(S^2) = \frac{2\sigma^4}{n-1}.$$

$$Mse(B_2) = \frac{2n-1}{n^2}\sigma^4.$$

当 $n > 1$ 时，有 $\frac{2n-1}{n^2} < \frac{2}{n-1}$,

因此在均方误差准则下，B_2 优于 S^2.

S^2 与 B_2 的选择：

在样本容量较小时，估计方差通常选样本方差 S^2；

当样本容量较大时，两种估计差异很小。