1. 假设 $X_1, X_2, \cdots, X_n$ 是来自总体 X 的简单随机样本,已知 $E(X^k) = \alpha_k (k = 1, 2, 3, 4)$. 证明:当 n 充分大时,随机变量 $Z_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ 近似服从正态分布,并指出其分布参数.

解 因为 $X_1, X_2, \cdots, X_n$ 独立同分布,

所以 $X_1^2, X_2^2, \cdots, X_n^2$ 也独立同分布,

且 $E(X_i^2)=\alpha_2,$

$$D(X_i^2)=E(X_i^4)-(EX_i^2)^2=\alpha_4-\alpha_2^2,$$

根据德莫佛一拉普拉斯定理知

$$V_n=\frac{\sum_{i=1}^{n}X_i{}^2-n\alpha_2}{\sqrt{n(\alpha_4-\alpha_2^2)}}=\frac{\frac{1}{n}\sum_{i=1}^{n}X_i{}^2-\alpha_2}{\sqrt{\frac{1}{n}(\alpha_4-\alpha_2^2)}}=\frac{Z_n-\alpha_2}{\sqrt{\frac{1}{n}(\alpha_4-\alpha_2^2)}}$$

的极限分布是标准正态分布.

故当n充分大时,

V_n近似服从标准正态分布,

从而当n充分大时,

$$Z_n = \sqrt{\frac{1}{n}(\alpha_4 - \alpha_2^2)}V_n + \alpha_2$$ 近似服从

参数为$\mu = \alpha_2$, $\sigma^2 = \dfrac{\alpha_4 - \alpha_2^2}{n}$的正态分布.

2.用机器包装味精,每袋味精净重为随机变量,期望值为100克,标准差为10克,一箱内装200袋味精,求一箱味精净重大于20500克的概率?

解: 设一箱味精净重为X,箱中第i袋味精净重为X_i,(i=1,2,…,200)

则 $X_1,X_2,\dots,X_{200}$独立同分布, $EX_i=100$, $DX_i=10^2=100$,

且 $X=\sum_{i=1}^{200}X_i$ 由独立同分布的中心极限定理得:

X近似服从正态分布, $E(X)=200E(X_i)=20000$, $D(X)=200D(X_i)=20000$,

所求为$P(X>20500)=1-P(X\le 20500)$

$$\approx 1-\Phi\left(\frac{20500-20000}{\sqrt{20000}}\right)=1-\Phi(3.54)=0.0002$$

故 一箱味精净重大于20500的概率为0.0002.

3. 假设测量的随机误差 $X \sim N(0, 10^2)$, 试求在 100 次独立重复测量中, 至少有 3 次测量误差的绝对值大于 19.6 的概率 α, 并利用泊松分布求出 α 的近似值 (要求小数点后取两位有效数字, $\Phi(1.96) = 0.975$).

解 设 p 为每次测量误差的绝对值大于 19.6 的概率,

$$p = P\{|X| > 19.6\} = P\left\{\frac{|X|}{10} > \frac{19.6}{10}\right\}$$

$$= P\left\{\frac{|X|}{10} > 1.96\right\} = 2 \cdot P\left\{\frac{X}{10} > 1.96\right\}$$

$$= 2 - 2\Phi(1.96) = 0.05,$$

设 k 为100次独立测量中事件 $\{|X| > 19.6\}$ 出现的次数,

则 k 服从参数为 $n = 100,\ p = 0.05$ 的二项分布,

故 $\alpha = P\{k \geq 3\} = 1 - P\{k < 3\}$

$$= 1 - 0.95^{100} - 100 \times 0.95^{99} \times 0.05$$

$$- \frac{100 \times 99}{2} \times 0.95^{98} \times 0.05^{2}$$

由**泊松定理**知,

k 近似服从参数 $\lambda = np = 100 \times 0.05$ 的泊松分布,

故 $\alpha = P\{k \geq 3\} = 1 - P\{k < 3\}$

$$\approx 1 - \frac{5^0 e^{-5}}{0!} - \frac{5e^{-5}}{1!} - \frac{5^2 e^{-5}}{2!}$$

$$= 1 - e^{-5}\left(1 + 5 + \frac{25}{2}\right) \approx 0.87 .$$

4. 售报员在报摊上卖报，已知每个过路人在报摊上买报的概率为1/3. 令X 是出售了100份报时过路人的数目，求 $P(280 \le X \le 320)$.

解:令X_i 为售出了第 $i-1$ 份报纸后到售出第i 份报纸时的过路人数，$i = 1,2,\ldots,100$

$$P(X_i = k) = p(1-p)^{k-1}\Big|_{p=1/3}, \quad k = 1,2,\cdots$$

$$E(X_i) = \frac{1}{p}\Big|_{p=1/3} = 3, \; D(X_i) = \frac{1-p}{p^2}\Big|_{p=1/3} = 6$$

$X_1, X_2, \cdots, X_{100}$ 相互独立, $X = \sum_{k=1}^{100} X_k$

$$E(X) = 300, \ \ D(X) = 600$$

由独立同分布中心极限定理, 有

$$X \sim N(300,600) \ \ (\text{近似})$$

$$P(280 \le X \le 320) \approx \Phi\left(\frac{320-300}{\sqrt{600}}\right) - \Phi\left(\frac{280-300}{\sqrt{600}}\right)$$

$$\approx 2\Phi\left(\frac{20}{\sqrt{600}}\right) - 1 \approx 2\Phi(0.8165) - 1 \approx 0.5878$$

5. 检验员逐个检查某产品,每查一个需用10秒钟．但有的产品需重复检查一次，再用去10秒钟．若产品需重复检查的概率为 0.5, 求检验员在 8 小时内检查的产品多于1900个的概率.

解: 若在 8 小时内检查的产品多于1900个,

即检查1900个产品所用的时间小于 8 小时.

设 X 为检查1900 个产品所用的时间(秒)

设 X_k 为检查第 k 个产品所用的时间

(单位：秒)，$k = 1,2,\ldots,1900$

X_k	10	20
P	0.5	0.5

$$E(X_k)=15,\quad D(X_k)=25$$

$X_1, X_2, \cdots, X_{1900}$ 相互独立同分布，$X=\sum_{k=1}^{1900} X_k$

$$E(X)=1900\times 15=28500$$

$$D(X)=1900\times 25=47500$$

$$X \overset{\text{近似}}{\sim} N(28500,47500)$$

$$P(10\times 1900 \le X \le 3600\times 8)$$

$$= p(19000 \le X \le 28800)$$

$$\approx \Phi\left(\frac{28800-28500}{\sqrt{47500}}\right)-\Phi\left(\frac{19000-28500}{\sqrt{47500}}\right)$$

$$\approx \varPhi(1.376)-\varPhi(-43.589)$$

$$\approx 0.9162$$

6. 将一颗骰子连掷100次，则点数之和不少于500的概率是多少？

解 设X_k为第k次掷出的点数,$k=1,2,\ldots,100$,则

$X_1,\ldots,X_{100}$独立同分布.

$$E(X_1)=\frac{7}{2},D(X_1)=\frac{1}{6}\sum_{i=1}^{6}k^2-\frac{49}{4}=\frac{35}{12}$$

由中心极限定理 $$\sum_{i=1}^{100}X_i \overset{\text{近似}}{\sim} N(100\times\frac{7}{2},100\times\frac{35}{12})$$

$$P\{\sum_{i=1}^{100}X_i\geq 500\}\approx 1-\Phi\left(\frac{500-100\times\frac{7}{2}}{10\sqrt{\frac{35}{12}}}\right)=1-\Phi(8.78)\approx 0$$

7. 设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立均服从泊松分布 $P(2)$. 若随机变量 $Y = \sum_{i=1}^{100} X_i$，求 $P(190 < Y < 210)$.

解 因为 X_i 服从 $P(2)$，$i = 1,2,\cdots$

即 $P(X_i = k) = \dfrac{2^k}{k!} e^{-2}, (k = 1,2,\cdots)$

所以 $E(X_i) = 2, D(X_i) = 2$，$i = 1,2,\cdots,100$

Y 近似服从 $N\left(200, (10\sqrt{2})^2\right)$，于是

$$P(190 < Y < 210) \approx \Phi\left(\frac{210-200}{10\sqrt{2}}\right) - \Phi\left(\frac{190-200}{10\sqrt{2}}\right)$$

$$= 2\Phi(0.707) - 1 = 0.52$$

8. 设某妇产医院出生男孩的概率为 0.515，求在 10000 个新生儿中，出生的女孩不少于男孩的概率.

解法1 设X为10000个新生儿中男孩个数

则X服从B(n,p)，其中n=10000，p=0.515

由德莫弗-拉普拉斯中心极限定理，所求概率为

$$P(X \le 5000) = P\left(\frac{X-np}{\sqrt{np(1-p)}} \le \frac{5000-np}{\sqrt{np(1-p)}}\right)$$

$$\approx \Phi\left(\frac{5000-np}{\sqrt{np(1-p)}}\right) = \Phi\left(\frac{5000-10000\times 0.515}{\sqrt{10000\times 0.515\times(1-0.515)}}\right) \approx \Phi(-3) = 0.00135$$

解法2 设X为10000个新生儿中男孩个数

$$X_i=\begin{cases}1 & \text{第i个是男孩}\\ 0 & \text{第i个是女孩}\end{cases}\quad (i=1,2,\cdots,10000)$$

则 $X=\sum_{i=1}^{10000} X_i$， 且 $X_1,X_2,\cdots,X_{10000}$独立同分布

$$\mu=E(X_i)=1\times0.515+0\times(1-0.515)=0.515$$

$$\sigma^2=D(X_i)=E(X_i^2)-\left(E(X_i)\right)^2=1^2\times0.515-0.515^2=0.249775$$

则女孩不少于男孩的概率为

$$P\left(X\le 5000\right)=P\left(\frac{X-10000\times0.515}{\sqrt{10000\times0.249775}}\le\frac{5000-10000\times0.515}{\sqrt{10000\times0.249775}}\right)$$

$$\approx\Phi\left(\frac{5000-10000\times0.515}{\sqrt{10000\times0.249775}}\right)\approx\Phi\left(-3\right)=0.00135$$

9.: 在一家保险公司里有10000个人参加寿命保险，每人每年付12元保险费。在一年内一个人死亡的概率为0.6%，死亡时其家属可向保险公司领得1000元，问：

(1)保险公司亏本的概率有多大?

(2)其他条件不变，为使保险公司一年的利润不少于60000元，赔偿金至多可设为多少?

解 设X表示一年内死亡的人数，则$X \sim B(n, p)$，

其中

$$n = 10000,\quad p = 0.6\%,$$

设Y表示保险公司一年的利润，

$$Y = 10000 \times 12 - 1000X$$

于是由中心极限定理

$$(1) P\{Y<0\} = P\{10000 \times 12 - 1000X < 0\}$$

$$= 1 - P\{X \le 120\}$$

$$\approx 1 - \Phi(7.75) = 0;$$

（**2**）设赔偿金为**a**元，则令

$$P\{Y \geq 60000\} \geq 0.9$$

P{Y>60000}=P{10000×12-a**X>60000}**

=P{X≤60000/a}≥0.9;

由中心极限定理,上式等价于

$$\Phi\left(\frac{\frac{60000}{a}-10000\times 0.006}{\sqrt{10000\times 0.006\times 0.994}}\right) \geq 0.9$$

$$\Rightarrow a \leq 3017$$

10. 某人一次射击，命中环数X的分布列为

X	10	9	8	7	6
P	0.8	0.1	0.05	0.02	0.03

求100次射击中命中环数在900环到930环之间的概率.

解:设X表示100次中命中的总环数,X_i表示第i次命中的环数(i=1,…,100),

则 $X_1,X_2,\ldots,X_{100}$ 相互独立同分布，E(X_i)=9.62, D(X_i)=0.82,

且 $X=\sum_{i=1}^{100} X_i$ E(X)=962, D(X)=82, 故 X~N(962,82)

P{900<X<930}= $\Phi\left(\frac{930-962}{\sqrt{82}}\right)-\Phi\left(\frac{900-962}{\sqrt{82}}\right) = \Phi(-3.53)-\Phi(6.85)$

=0.99979

1. 设有一容量n=8的样本观察值为(8,6,7,5,7,8,9,6)，求样本均值及样本方差的观察值。

解　　由$\overline{X}=\left(\sum_{i=1}^{n}X_i\right)\Big/n$，得

$$\overline{x}=\frac{1}{8}(8+6+7+5+7+8+9+6)=7$$

由$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$，得

$$S^2=\frac{1}{7}[(8-7)^2+(6-7)^2+\cdots+(6-7)^2]=2$$

例2 已知某种纱的强力服从$N(1.56,0.22^2)$（单位：千克）今抽取容量为$n=50$的样本，求样本均值小于1.45千克的概率。

解 $E(\overline{X})=\mu=1.56$，$D(\overline{X})=\sigma^2/n=0.22^2/50\approx(0.031)^2$

所以

$$P\{\overline{X}<1.45\}=\Phi(\frac{1.45-1.56}{0.031})=\Phi(-3.55)$$

$$=1-\Phi(3.55)=0.00019$$

3.设总体 X 与 Y 相互独立且均服从正态分布 $N\left(0,3^2\right)$，$X_1,X_2,\cdots,X_9$ 和 $Y_1,Y_2,\cdots,Y_9$ 分别为来自 X 与 Y 的样本，则统计量 $U=\dfrac{X_1+X_2+\cdots+X_9}{\sqrt{Y_1^2+Y_2^2+\cdots+Y_9^2}}$ 服从什么分布？

解：因为 $X\sim N(0,3^2)$，$Y\sim N(0,3^2)$，

所以 $X_i\sim N(0,3^2)$ $\quad Y_i\sim N(0,3^2)$ $\quad (i=1,2,\cdots,9)$. 于是有

$$EX_i=0 \qquad EY_i=0 \qquad S_X^2=DX=3^2=9=S_Y^2 \qquad (i=1,2,\cdots,9)$$

推得 $$U=\frac{X_1+X_2+\cdots+X_9}{\sqrt{Y_1^2+Y_2^2+\cdots+Y_9^2}}=\frac{\frac{1}{9}\sum_{i=1}^{9}X_i}{\frac{1}{9}\sqrt{\sum_{i=1}^{9}Y_i^2}}=\frac{\overline{X}}{\sqrt{\frac{1}{9}\cdot\frac{1}{9}\sum_{i=1}^{9}Y_i^2}}=\frac{\overline{X}}{\sqrt{\frac{1}{9}S_Y^2}}=\frac{\overline{X}-E\overline{X}}{\frac{S_Y}{\sqrt{9}}}$$

$$=\frac{\overline{X}-E\overline{X}}{S_X\big/\sqrt{9}}\sim t(9-1)=t(8)$$

即 $U\sim t(8)$ 分布.

例4　设总体$X \sim N(0,\sigma^2)$，X_1, X_2是X的一个样本，试求$Y=(X_1+X_2)^2/(X_1-X_2)^2$服从什么分布。

解　因为X_1+X_2和X_1-X_2均服从$N(0,2\sigma^2)$，所以

$[(X_1+X_2)/\sqrt{2}\sigma]^2 \sim \chi^2(1)$，$[(X_1-X_2)/\sqrt{2}\sigma]^2 \sim \chi^2(1)$

由F分布定义，知

$$\frac{[(X_1+X_2)/\sqrt{2}\sigma]^2}{[(X_1-X_2)/\sqrt{2}\sigma]^2} = \frac{(X_1+X_2)^2}{(X_1-X_2)^2} \sim F(1,1)$$

5、设总体 $X\sim N(50,6^2)$，总体 $Y\sim N(46,4^2)$，从总体 X 中抽取容量为10的样本，从总体 Y 中抽取容量为8的样本，求下列概率：

（1）；$P(0<\overline{X}-\overline{Y}<8)$；（2）$P(\frac{S_1^2}{S_2^2}<8.28)$.

解：（1）因为 $X\sim N(50,6^2)$，$Y\sim N(46,4^2)$，所以由

$$u=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1).$$

得

$$P(0<\overline{X}-\overline{Y}<8)=P(-\frac{50-46}{\sqrt{\frac{6^2}{10}+\frac{4^2}{8}}}<\frac{(\overline{X}-\overline{Y})-(50-46)}{\sqrt{\frac{6^2}{10}+\frac{4^2}{8}}}<\frac{8-(50-46)}{\sqrt{\frac{6^2}{10}+\frac{4^2}{8}}})$$

$$=P(-\frac{4}{\sqrt{5.6}}<\frac{(\overline{X}-\overline{Y})-(50-46)}{\sqrt{5.6}}<\frac{4}{\sqrt{5.6}})=2\Phi(1.69)-1\approx 0.909.$$

（2）因为 $X \sim N(50, 6^2)$，$Y \sim N(46, 4^2)$，所以由

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(9, 7).$$

得

$$P\left(\frac{S_1^2}{S_2^2} < 8.28\right) = P\left(\frac{S_1^2/6^2}{S_2^2/4^2} < 8.28 \times \frac{4^2}{6^2}\right) = P\left(\frac{S_1^2/6^2}{S_2^2/4^2} < 3.68\right) \approx 0.95.$$

1. 已知总体X的概率密度为:

$$f(x)=\begin{cases}\sqrt{\theta}x^{\sqrt{\theta}-1}, & 0\le x\le 1,\\ 0, & 其它,\end{cases}$$

其中未知参数θ>0, 求θ的矩估计量.

解. 单参数，连续型.

因为总体一阶矩

$$\mu_1=E(X)=\int_{-\infty}^{+\infty}xf(x)dx=\sqrt{\theta}\int_0^1 x^{\sqrt{\theta}}dx$$

$$=\frac{\sqrt{\theta}}{\sqrt{\theta}+1}x^{\sqrt{\theta}+1}\Big|_0^1=\frac{\sqrt{\theta}}{\sqrt{\theta}+1}$$

由

$$\mu_1=A_1$$

即

$$\frac{\sqrt{\theta}}{\sqrt{\theta}+1}=\overline{X}$$

解得: $\sqrt{\theta}=\overline{X}(\sqrt{\theta}+1)$ $\quad\quad$ $\sqrt{\theta}(1-\overline{X})=\overline{X}$

$$\sqrt{\theta}=\frac{\overline{X}}{1-\overline{X}}$$

故所求矩估计量为:

$$\boxed{\hat{\theta}=\left(\frac{\overline{X}}{1-\overline{X}}\right)^2}$$

2. 已知总体X的概率密度为:

$$f(x)=\frac{1}{2\theta}e^{-\frac{|x|}{\theta}}\ (-\infty<x<+\infty)$$

其中未知参数θ>0,求θ的矩估计量.

解:因为总体一阶矩

$$\mu_1=E(X)=\int_{-\infty}^{+\infty}x\cdot\frac{1}{2\theta}e^{-\frac{|x|}{\theta}}dx=0$$

不含θ,故不能由"样本一阶矩=总体一阶矩"解得所求
矩估计,需要继续求二阶矩:

$$\mu_2=E(X^2)=\frac{1}{2\theta}\int_{-\infty}^{+\infty}x^2\cdot e^{-\frac{|x|}{\theta}}dx$$

$$=\frac{1}{\theta}\int_0^{+\infty} x^2\cdot e^{-\frac{x}{\theta}}dx=\theta^2\int_0^{+\infty}\left(\frac{x}{\theta}\right)^2 e^{-\frac{x}{\theta}}d\left(\frac{x}{\theta}\right)$$

$$=\theta^2\Gamma(3)=2\theta^2,$$

Γ函数定义

由“样本二阶矩=总体二阶矩”得:

$$\frac{1}{n}\sum_{i=1}^{n}X_i^2=2\theta^2,$$

于是, 所求矩估计量为: $\boxed{\hat{\theta}=\sqrt{\frac{1}{2n}\sum_{i=1}^{n}X_i^2}}$

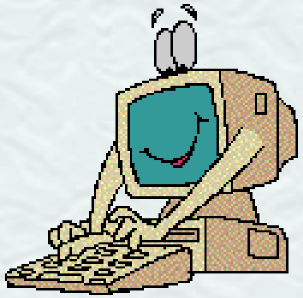
例3. 设总体 X 在 $[a,b]$ 上服从均匀分布,其中 a, b 未知, $x_1,x_2,\cdots,x_n$ 是来自总体 X 的一个样本值,求 a,b 的最大似然估计量.

解 记 $x_{(l)}=\min(x_1,x_2,\cdots,x_n)$,

$$x_{(h)}=\max(x_1,x_2,\cdots,x_n),$$

X 的概率密度为

$$f(x;a,b)=\begin{cases}\dfrac{1}{b-a}, & a\le x\le b,\\ 0, & \text{其他}.\end{cases}$$



因为$a \le x_1, x_2, \cdots, x_n \le b$等价于$a \le x_{(l)}, x_{(h)} \le b$,

作为a, b的函数的似然函数为

$$L(a,b)=\begin{cases}\dfrac{1}{(b-a)^n}, & a \le x_{(l)},\ b \ge x_{(h)}, \\ 0, & \text{其他}\end{cases}$$

于是对于满足条件 $a \le x_{(l)},\ b \ge x_{(h)}$的任意$a,b$有

$$L(a,b)=\frac{1}{(b-a)^n} \le \frac{1}{(x_{(h)}-x_{(l)})^n},$$

即似然函数$L(a,b)$在$a = x_{(l)}$, $b = x_{(h)}$时

取到最大值$(x_{(h)} - x_{(l)})^{-n}$,

a, b的最大似然估计值

$$\hat{a} = x_{(l)} = \min_{1\le i\le n} x_i, \qquad \hat{b} = x_{(h)} = \max_{1\le i\le n} x_i,$$

a, b的最大似然估计量

$$\hat{a} = \min_{1\le i\le n} X_i, \qquad \hat{b} = \max_{1\le i\le n} X_i .$$

4. 设从存在均值μ与方差$\sigma^2>0$的总体中，分别抽取容量为n_1, n_2的两个独立样本，其样本均值分别为 $\overline{X}_1, \overline{X}_2$. 证明：对任意常数a, b,

$$Y = a\overline{X}_1 + b\overline{X}_2 (a+b=1)$$

是总体均值μ的无偏估计；并确定常数a, b使D(Y)达到最小.

解：因为 $$E(\overline{X}_k) = \mu, D(\overline{X}_k) = \frac{\sigma^2}{n_k} (k=1,2)$$

由期望性质得：

$$E(Y)=E(a\overline{X}_1+b\overline{X}_2)$$
$$=aE(\overline{X}_1)+bE(\overline{X}_2)$$
$$=(a+b)\mu\ =\mu$$

由无偏性知:Y是μ的无偏估计量.

由方差性质得:

$$D(Y)=D(a\overline{X}_1+b\overline{X}_2)=a^2D(\overline{X}_1)+b^2D(\overline{X}_2)$$
$$=a^2\cdot\frac{\sigma^2}{n_1}+b^2\cdot\frac{\sigma^2}{n_2}=[\frac{a^2}{n_1}+\frac{(1-a)^2}{n_2}]\sigma^2$$

由导数应用知:

$$\frac{d}{da}D(Y)=[\frac{2a}{n_1}-\frac{2(1-a)}{n_2}]\sigma^2 \overset{令}{=} 0$$

即:

$$\frac{a}{n_1}=\frac{(1-a)}{n_2}$$

解得当

$$a=\frac{n_1}{n_1+n_2},b=\frac{n_2}{n_1+n_2}$$

时D(Y)最小.

5. 设总体X服从参数为1/θ的指数分布

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}}, & x>0,\\ 0, & \text{其它}\end{cases}$$

其中θ>0为未知参数, 试证:

1) $\overline{X}$和 $nZ=n\min(X_1,X_2,\cdots,X_n)$都是θ的无偏估计;

2) 评定 $\overline{X},nZ$的有效性.

解: 因为

$$E(\overline{X})=\theta,D(\overline{X})=\frac{\theta^2}{n},$$

所以, $\overline{X}$ 是θ的无偏估计量.

易知 $Z=\min\limits_{1\le i\le n}\{X_i\}$ 服从参数为**n/θ**的指数分布,故

$$E(Z)=\frac{\theta}{n},$$

于是,

$$E(nZ)=\theta,$$

所以, nZ也是 θ 的无偏估计量.

由于,

$$D(nZ)=n^2D(Z)=n^2\left(\frac{\theta}{n}\right)^2=\theta^2,$$

注意到当n>1时:

$$D(\overline{X})<D(nZ),$$

故当n>1时, $\overline{X}$ 较 nZ 为有效.