概率论与数理统计

华中科技大学 概率统计系

吴娟

07

参数估计

评选原则

无偏性　若 $E(\hat{\theta})=\theta$，则称 $\hat{\theta}$ 为 θ 的无偏估计量.

例　证明样本均值 $\bar{X}$ 为总体期望 $\mu=E(X)$ 的无偏估计.

证

$$E(\bar{X})=\frac{1}{n}\sum_{i=1}^{n}E(X_i)\ =E(X)=\mu$$

一般

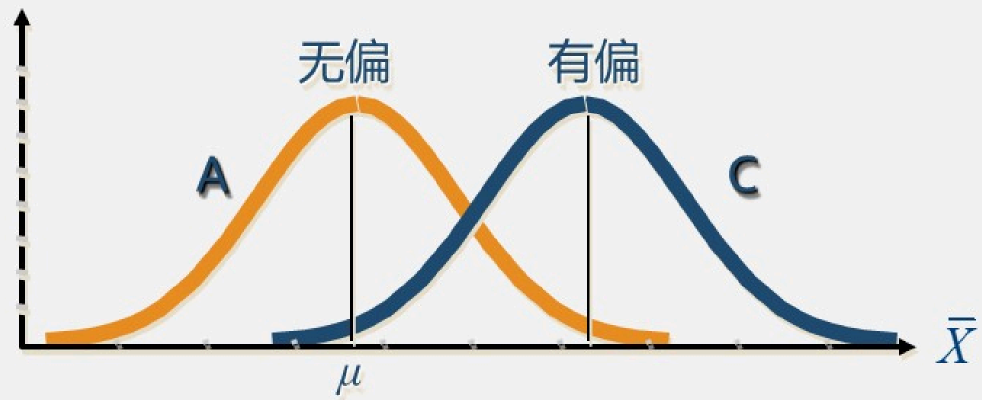
$$E(A_k)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^k)=E(X^k)=\alpha_k$$

单选题 100分

若抽样一次，无偏估计量的结果会更好吗？

A 会

B 不一定



评选原则

例 证明样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 是总体方差 $D(X)=\sigma^2$ 的无偏估计.

证

$$E(S^2)=E[\frac{1}{n-1}(\sum_{i=1}^{n}X_i^2-n\bar{X}^2)]$$

$$=\frac{1}{n-1}[nE(X^2)-nE(\bar{X}^2)]$$

$$\xrightarrow{D(X)=E(X^2)-(EX)^2} \quad =\frac{1}{n-1}[n(\sigma^2+\mu^2)-n(D(\bar{X})+\mu^2)]$$

$$=\frac{1}{n-1}(n\sigma^2+n\mu^2-\sigma^2-n\mu^2)$$

$$=\sigma^2$$

评选原则

注1 $\tilde{S}^2$ 不是 σ^2 的无偏估计.

$$E(\tilde{S}^2) = E(\frac{n-1}{n}S^2) = \frac{n-1}{n}\sigma^2 \xrightarrow[n\to\infty]{} \sigma^2$$ 渐近无偏估计

注2 $\hat{\theta}$ 为 θ 的无偏估计，但 $u(\hat{\theta})$不一定是 $u(\theta)$的无偏估计.

例 若 $D(X) > 0$，则 $E[(\hat{\mu})^2] = E(\bar{X}^2) = D(\bar{X}) + [E(\bar{X})]^2 > [E(\bar{X})]^2 = \mu^2$

注3 无偏估计不唯一，如X_1和 $\bar{X}$ 均为 $\mu=E(X)$ 的无偏估计.

事实上对任何 $c_1, c_2, ..., c_n$，当 $c_1+c_2+\cdots+c_n=1$ 时，

$$E(\sum_{i=1}^{n} c_i X_i) = \mu$$

评选原则

应用 对某种产品从1开始用整数连续编号，试估计产品的总量.

分析 设产品编号为X, 则

X	1	2	$\cdots$	N
p	$\frac{1}{N}$	$\frac{1}{N}$	$\cdots$	$\frac{1}{N}$

样本：20, 96, 190, 255

无偏估计：317.75

$$L(N)=\frac{1}{N^n},\quad 1\le x_1^*\le x_2^*\le\cdots\le x_n^*\le N,\qquad \hat{N}=x_n^*$$

$$E(X_n^*)=\frac{n}{n+1}(N+1),\qquad E(\frac{n+1}{n}X_n^*\;-1)=N$$

评选原则

证明 $E(X_n^*)=\dfrac{n}{n+1}(N+1)$

$1\le x_1^*\le x_2^*\le\cdots\le x_n^*\le N,\qquad P(X_n^*=k)=C_{k-1}^{n-1}\big/C_N^n,\ \ k=n,n+1,\ldots,N$

$$E(X_n^*)=\sum_{k=n}^{N}kP(X_n^*=k)=n\sum_{k=n}^{N}C_k^n\big/C_N^n=n\,C_{N+1}^{n+1}\big/C_N^n=\frac{n}{n+1}(N+1).$$

$$(C_k^n+C_k^{n+1}=C_{k+1}^{n+1}\quad\Rightarrow\quad\sum_{k=n}^{N}C_k^n=C_{N+1}^{n+1})$$

评选原则

有效性

定义 设 $\hat{\theta}_1, \hat{\theta}_2$ 都是参数θ的无偏估计量，若

$$D(\hat{\theta}_1) < D(\hat{\theta}_2)$$

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

设 $\hat{\theta}_0$ 是参数θ的无偏估计量，若对θ的任一无偏估计$\hat{\theta}$有

$$D(\hat{\theta}_0) \leq D(\hat{\theta})$$

则称 $\hat{\theta}_0$ 是θ的最小方差无偏估计.

评选原则

有效性

例 证明 $\bar{U}=\{\hat{\mu}=\sum_{i=1}^{n}k_iX_i \mid \sum_{i=1}^{n}k_i=1\}$ 是μ=$E(X)$的线性无偏估计类.

并在此无偏估计类里寻找μ的最小方差无偏估计.

证

$$E(\hat{\mu})=\sum_{i=1}^{n}k_iE(X_i)=\sum_{i=1}^{n}k_i\mu=\mu, \qquad \mu\in\bar{U}$$

评选原则

例 证明 $\bar{U}=\{\hat{\mu}=\sum_{i=1}^{n}k_iX_i \mid \sum_{i=1}^{n}k_i=1\}$ 是$\mu=E(X)$的线性无偏估计类.

在此无偏估计类里寻找μ的最小方差无偏估计.

分析 $(\sum_{i=1}^{n}a_ib_i)^2 \le (\sum_{i=1}^{n}a_i^2)(\sum_{i=1}^{n}b_i^2)$

$$D(\hat{\mu})=D(X)\sum_{i=1}^{n}k_i^2=\frac{1}{n}(\sum_{i=1}^{n}1^2)(\sum_{i=1}^{n}k_i^2)D(X)\ge\frac{1}{n}\left(\sum_{i=1}^{n}1\times k_i\right)^2D(X)$$

$$=\frac{1}{n}D(X)$$

而 $D(\bar{X})=\frac{1}{n}D(X)$ 故 $\hat{\mu}_0=\bar{X}$ 是μ的最小方差无偏估计.

评选原则

一致性

定义 设 $\hat{\theta}$ 是参数θ的估计量，若对任意$\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left(\left|\hat{\theta} - \theta\right| < \varepsilon\right) = 1$$

即 $\hat{\theta} \xrightarrow{P} \theta$

则称 $\hat{\theta}$为 θ 的 一致估计量.

区间估计

问题　如何让$\hat{\theta}$与θ的误差体现在估计中?

办法　对给定的置信水平(置信度)1-α

置信下限$\hat{\theta}_1(X_1, X_2, \ldots, X_n)$和置信上限$\hat{\theta}_2(X_1, X_2, \ldots, X_n)$

使

$$P(\hat{\theta}_1 < \theta < \hat{\theta}_2) \geq 1-\alpha$$

称$(\hat{\theta}_1, \hat{\theta}_2)$为未知参数$\theta$的置信度为1 - α的置信区间.

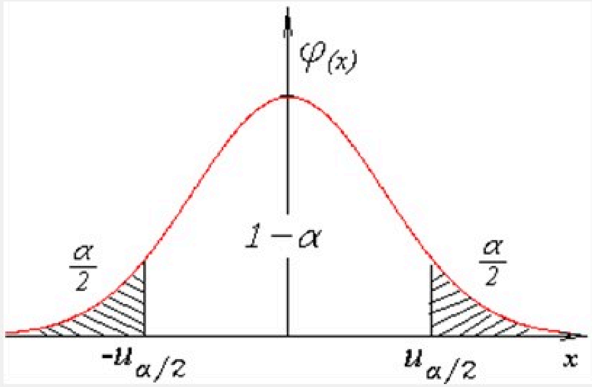
含义　若1 - α =0.95，抽样100次中约有95个$(\hat{\theta}_1, \hat{\theta}_2)$包含$\theta$.

区间估计

单个正态总体均值的区间估计

σ^2已知 $\quad \dfrac{\bar{X}-\mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$

$$P(\bar{X}-u_{\alpha/2}\frac{\sigma}{\sqrt{n}}<\mu<\bar{X}+u_{\alpha/2}\frac{\sigma}{\sqrt{n}})=1-\alpha$$

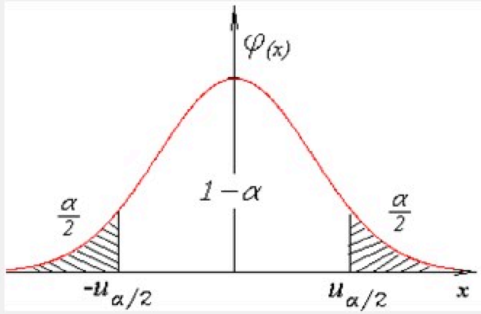


μ的置信度为1 - α的置信区间为$(\bar{X}-u_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{X}+u_{\alpha/2}\frac{\sigma}{\sqrt{n}})$; $\quad (\bar{X}\pm u_{\alpha/2}\frac{\sigma}{\sqrt{n}})$

区间估计

单个正态总体均值的区间估计

注 $(\bar{X}-u_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{X}+u_{\alpha/2}\frac{\sigma}{\sqrt{n}})$



- 置信区间长度

$$l=\frac{2\sigma}{\sqrt{n}}u_{\alpha/2}$$

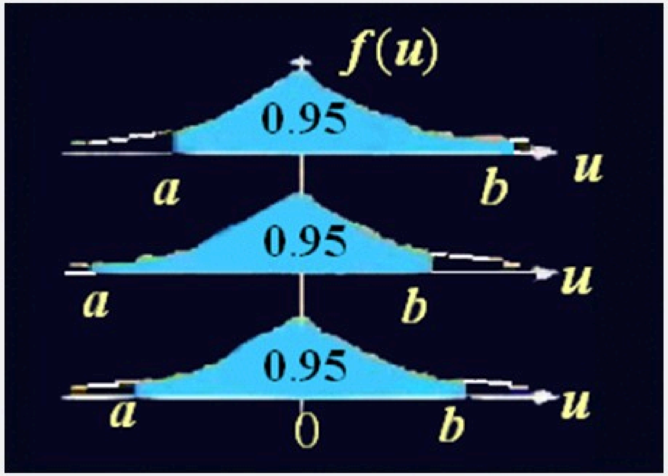
15

投票(匿名) 最多可选1项

未知参数的置信度为1-α的置信区间唯一吗？

A 唯一

B 不唯一



区间估计

单个正态总体均值的区间估计

例 滚珠直径$X \sim N(\mu, 0.0006)$，从某天生产的滚珠中随机抽取6个，测得直径为(单位：mm) 1.46 1.51 1.49 1.48 1.52 1.51

求 μ 的置信度为95%的置信区间.

解 $(\bar{X} - u_{\alpha/2}\dfrac{\sigma}{\sqrt{n}},\ \bar{X} + u_{\alpha/2}\dfrac{\sigma}{\sqrt{n}})$

$$\bar{x} = 1.495, \qquad \alpha = 0.05, \qquad u_{0.05/2} = u_{0.025} = 1.96$$

$$(1.495 \pm 1.96 \times \sqrt{\frac{0.0006}{6}}) = (1.4754,\ 1.5146)$$

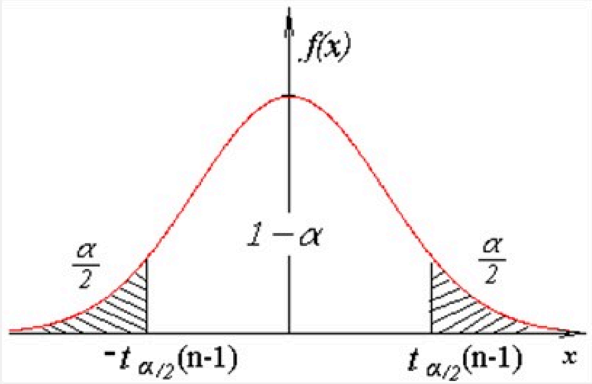
17

区间估计

单个正态总体均值的区间估计

σ^2未知

$$T=\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$$



$$P(|T|<t_{\alpha/2}(n-1))=1-\alpha$$

$$P(\bar{X}-t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}<\mu<\bar{X}+t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}})=1-\alpha$$

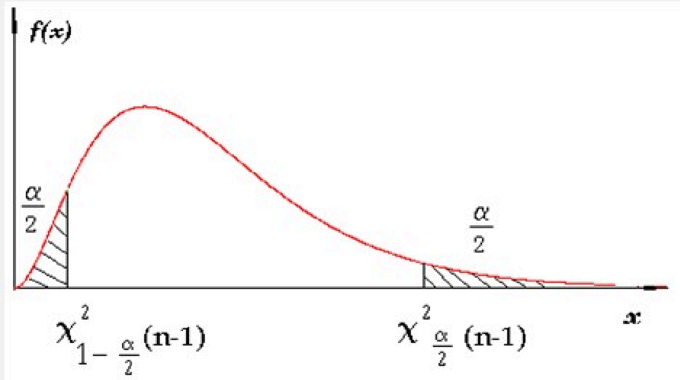
μ的置信度为1-α的置信区间为 $(\bar{X}\pm t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}})$

区间估计

单个正态总体方差的区间估计

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

$$P\left(\chi^2_{1-\frac{\alpha}{2}}(n-1) < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{1-\frac{\alpha}{2}}(n-1)\right) = 1-\alpha$$



$$P\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right) = 1-\alpha$$

σ^2 的置信度为1-α 的置信区间为 $\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$

区间估计

单个正态总体方差的区间估计

例 从自动车床加工的一批零件中随机的抽取16件，测得各零件长度为(单位：cm)

2.15	2.10	2.12	2.10	2.14	2.11	2.15	2.13
2.13	2.11	2.14	2.13	2.12	2.13	2.10	2.14

设零件长度服从正态分布，求零件长度标准差σ的置信度为95%的置信区间.

解 计算 $\bar{x}=2.125$, $S^2=0.000293$, $\alpha=0.05$, $\chi^2_{0.025}(15)=27.488$, $\chi^2_{0.975}(15)=6.262$

$$\left(\sqrt{\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}},\ \sqrt{\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}}\right)$$

$$=(0.01265，0.02651)$$