第二章 参数估计

§2 估计量的好坏标准

理学院

预习提纲

1. 评价估计量的好坏有哪些标准?

2. 我们希望一个“好的”估计量具有什么特性?

3. 怎样确定一个估计量是否比另一个估计量“好”?

- 本次课教学目的:
  - 掌握如何评价估计量的好坏
- 重点难点:
  - 求得合理估计量的方法
  - 罗-克拉美不等式的推导及其应用

§2　估计量的好坏标准

通过上一节的学习，我们发现对母体的未知参数可以用几种不同的方法求得估计量，那应该怎样来衡量与比较估计量的好坏呢？

一个好的估计，应在多次试验中体现出优良性 .而且尽可能接近待估计参数值的真值，在真值左右摆动尽可能小.

本节主要介绍三条评价标准：

(1) 无偏性；

(2) 相合性；

(3) 优效估计；

2.1 无偏性

估计量是随机变量，对于不同的样本值会得到不同的估计值。我们希望估计值在未知参数真值附近摆动，而它的期望值等于未知参数的真值，这就产生了无偏性这个标准。

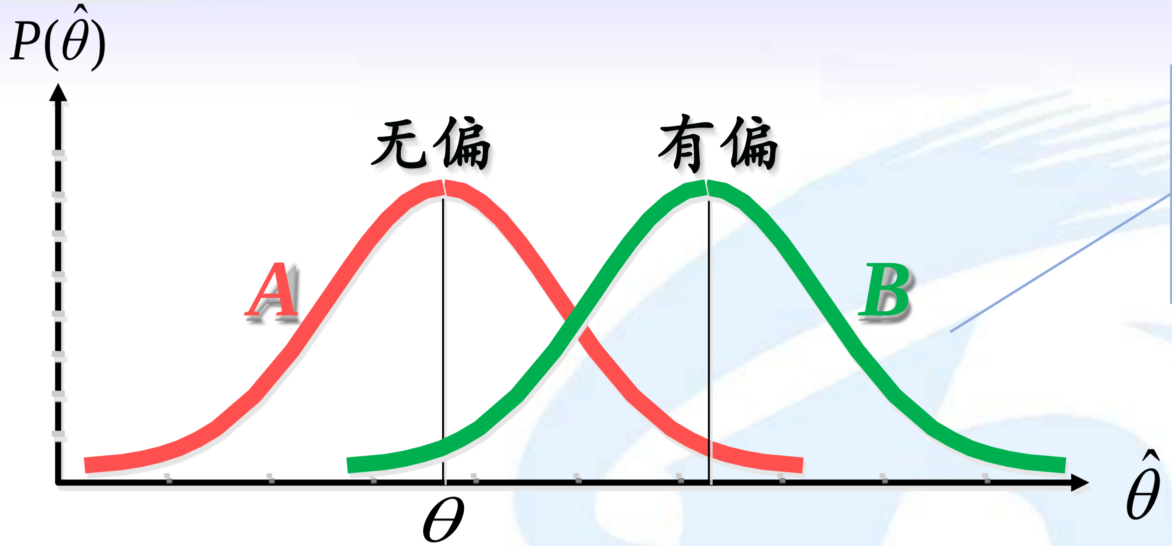
定义2.1：

设 $\hat{\theta}(X_1,\cdots,X_n)$是未知参数$\theta$ 的估计量

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

若 $E(\hat{\theta})\neq\theta$ 则称 $E(\hat{\theta})-\theta$ 为估计量 $\hat{\theta}$ 的偏差

若 $\lim\limits_{n\to\infty} E(\hat{\theta})=\theta$ 则称 $\hat{\theta}$ 为 θ 的渐近无偏估计



无偏性的实际意义是指没有系统性的偏差。

例1： 设$X_1, X_2, \ldots, X_n$是取自具有一阶矩、二阶矩存在的总体X一个样本，证明

(1). $\overline{X}$是$EX = \mu$的无偏估计；

(2). S_n^{*2}是$DX = \sigma^2$的无偏估计；

(3). S_n^2是$DX = \sigma^2$的渐近无偏估计 .

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

解：(1). 由于 $E(\overline{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} EX_i = EX$

(2). 由于 $E\left(S_n^{*2}\right) = \frac{n}{n-1} E\left(S_n^2\right) = \sigma^2$

(例2.2.2(P30))

而 $E\left(S_n^2\right) = \frac{n-1}{n}\sigma^2$

(3). $E\left(S_n^2\right) = \frac{n-1}{n}\sigma^2 \Rightarrow \lim\limits_{n\to\infty} E\left(S_n^2\right) = \sigma^2$

一般地，$\hat{\theta}$是θ的无偏估计，$f(\hat{\theta})$不一定是 $f(\theta)$的无偏估计。

思考：(1). $\overline{X}^2$是否是 μ^2的无偏估计 ？

事实上 $E(\overline{X}^2) = D(\overline{X}) + (E\overline{X})^2 = \frac{1}{n}\sigma^2 + \mu^2 \neq \mu^2.$

思考题

1、设 $X_1, X_2, \cdots, X_n$ 为母体$N(\mu, \sigma^2)$的一个子样。试选择适当常数C,使 $C\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$ 为 σ^2 的无偏估计。

解： 因为

$$E(\hat{\sigma}^2) = C\sum_{i=1}^{n-1} E(X_{i+1} - X_i)^2 = C\sum\left[D(X_{i+1} - X_i) + (E(X_{i+1} - X_i))^2\right]$$

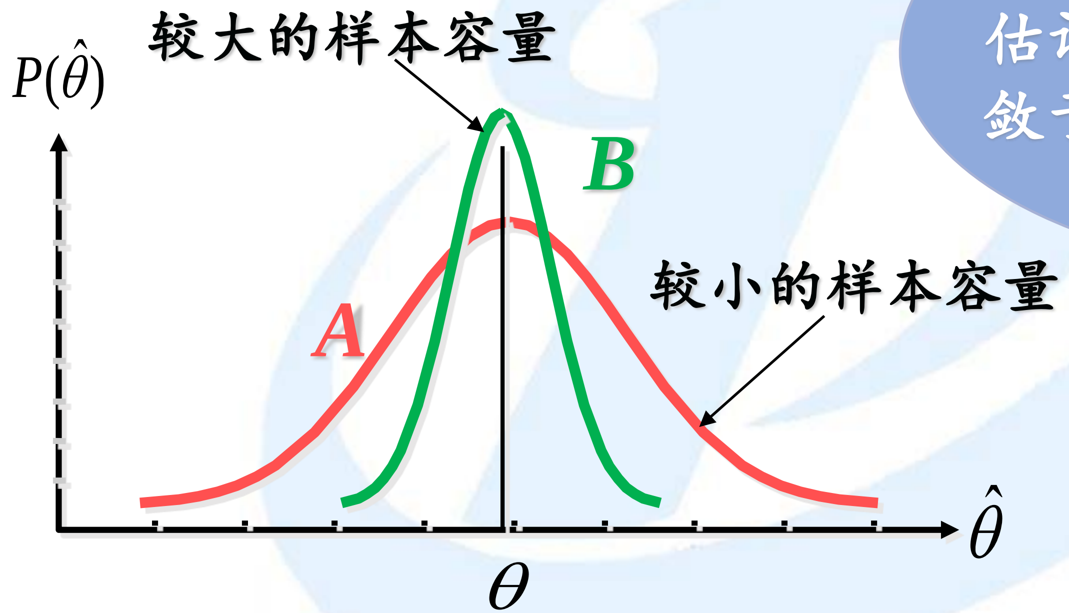
$$= C\sum_{i=1}^{n-1}\left[D(X_{i+1}) + D(X_i) + 0\right] = C\sum_{i=1}^{n-1} 2\sigma^2 = 2C(n-1)\sigma^2$$

要使 $E(\hat{\sigma}^2) = \sigma^2$，只需$C = \dfrac{1}{2(n-1)}$即可。

2.2 相合估计量

定义2.2： 若当 $n\to\infty$ 时 $\hat{\theta}$ 依概率收敛到 θ，即对任意 $\varepsilon>0$，有：$\lim\limits_{n\to\infty}P\{|\hat{\theta}-\theta|<\varepsilon\}=1$ 则称 $\hat{\theta}$ 是 θ 的相合估计(量)或一致估计(量)。

当样本容量n无限增大时，估计值能在某种意义下收敛于被估计参数的真值。



相合估计指在大样本条件下引进的，是对估计量的基本要求，并且由大数定理可知，这个要求也是容易满足的。

但是，这个定义是不容易验证的，且通过常用的方法构造的估计量可能不具备相合性的。为此有下列判别方法：

定理： 对任给的 $\varepsilon > 0$ 满足：

$$\begin{cases} \lim\limits_{n\to\infty} E\hat{\theta}_n = \theta \\ \lim\limits_{n\to\infty} D\hat{\theta}_n = 0 \end{cases}$$ 则 $\hat{\theta}_n$ 是 θ 的相合估计量。

例3： $X \sim B(1, P)$

$$E\overline{X} = p, D\overline{X} = \frac{p(1-p)}{n} \to 0, \quad (n \to \infty)$$

$\overline{X}$是p的相合估计量.

2.3 优效估计

一个参数往往有不止一个无偏估计，无偏性原则不够，还应要求估计量围绕参数的真值波动尽可能小，该如何刻画？

概率论用随机变量的均方误差（方差）刻画“离散”或“集中”的程度。

我们可以比较 $E(\hat{\theta}_1-\theta)^2$ 和 $E(\hat{\theta}_2-\theta)^2$ 的大小来判断二者谁更优，由于 $D(\hat{\theta}_1)=E(\hat{\theta}_1-\theta)^2, D(\hat{\theta}_2)=E(\hat{\theta}_2-\theta)^2$ 方差小者为佳，从而有最小方差无偏估计与有效性这一概念。

定义2.3： 若 $\hat{\theta}_1$和 $\hat{\theta}_2$都是θ的无偏估计量，且$D\hat{\theta}_1 < D\hat{\theta}_2$，则称$\hat{\theta}_1$比$\hat{\theta}_2$有效。

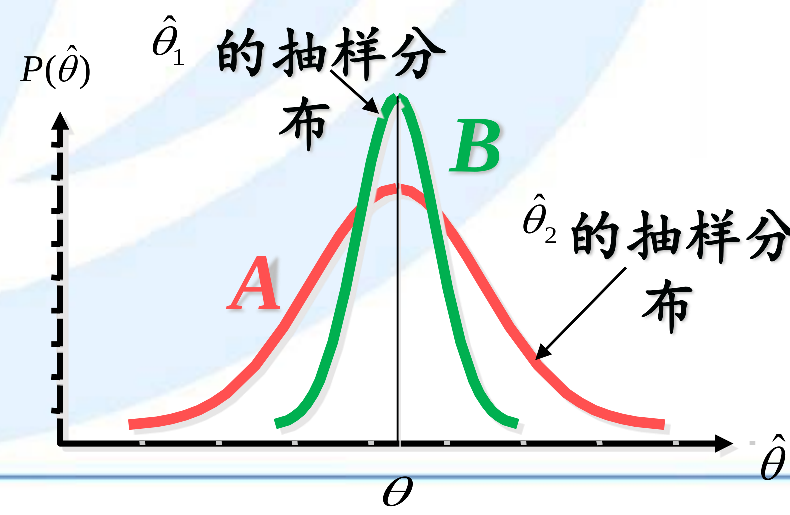
例2： 作为μ的估计量$\hat{\mu}_1 = \frac{1}{5}X_1 + \frac{3}{5}X_2 + \frac{1}{5}X_3$和$\hat{\mu}_2 = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$谁更有效？

解：

$$D\hat{\mu}_1 = \frac{1}{25}DX_1 + \frac{9}{25}DX_2 + \frac{1}{25}DX_3 = \frac{11}{25}\sigma^2$$

$$D\hat{\mu}_2 = \frac{1}{9}DX_1 + \frac{1}{9}DX_2 + \frac{1}{9}DX_3 = \frac{1}{3}\sigma^2 < \frac{11}{25}\sigma^2$$

$\hat{\mu}_2$比$\hat{\mu}_1$有效.



定义2.4：若θ的所有二阶矩存在的无偏估计量中存在一个估计量$\hat{\theta}_0$，使对任意无偏估计量$\hat{\theta}$有：$D\hat{\theta}_0 \le D\hat{\theta}$，则称$\hat{\theta}_0$是$\hat{\theta}$的最小方差无偏估计量。

对某个分布的一个待估参数,如果能找到这个待估参数的估计量下界,而某个无偏估计量又能达到这个下界,它就一定是最小方差无偏估计量。而该下界可以由罗-克拉美(C.R.Rap-H.Crainer)不等式给出。

罗-克拉美(C-R)不等式

定义2.5： 如果θ的无偏估计量$\hat{\theta}_0$的方差等于罗-克拉美下界，即$D(\hat{\theta}_0) = I_R$，则称$\hat{\theta}_0$是θ的优(有)效估计（量）。其中

$$I_R = \begin{cases} n\int_{-\infty}^{\infty}\left(\dfrac{\partial \ln f(x;\theta)}{\partial\theta}\right)^2 f(x;\theta)dx & \text{（连续母体）} \\ n\sum\limits_{x}\left(\dfrac{\partial \ln P(x;\theta)}{\partial\theta}\right)^2 P(x;\theta) & \text{（离散母体）} \end{cases}$$

若θ的无偏估计量为$\hat{\theta}$，则$e(\hat{\theta}) = \dfrac{I_R}{D\hat{\theta}}$称为估计量$\hat{\theta}$的(有)效率，从而有$0 \le e(\hat{\theta}) \le 1$，且$e(\hat{\theta}) = 1$时，$\hat{\theta}$为$\theta$的优效估计。

若估计量$\hat{\theta}$满足$\lim\limits_{n\to\infty} e(\hat{\theta}) = 1$，则称$\hat{\theta}$为$\theta$的渐近优效估计(量)。

例3：$X \sim B(1,\ p)$，求p 的优效估计量

解：由题意得：$P(x) = p^{x} * (1-p)^{1-x}$，$x = 0,1$，计算罗-克拉美下界。因为：

$$\sum_{x=0}^{1}\left(\frac{\partial \ln P(x)}{\partial p}\right)^2 P(x)=\left(\frac{\partial \ln(1-p)}{\partial p}\right)^2(1-p)+\left(\frac{\partial \ln p}{\partial p}\right)^2 p$$

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

$$I_R=\begin{cases}\left[n\int_{-\infty}^{\infty}\left(\frac{\partial \ln f(x;\theta)}{\partial \theta}\right)^2 f(x;\theta)dx\right]^{-1}\\ \left[n\sum_{x}\left(\frac{\partial \ln P(x;\theta)}{\partial \theta}\right)^2 P(x;\theta)\right]^{-1}\end{cases}$$

故得罗-克拉美下界：$I_R=\dfrac{p(1-p)}{n}$

又 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 是p的无偏估计，且$D\bar{X}=\frac{1}{n}p(1-p)$，所以$p$的优效估计为 $\bar{X}$

例4： 设母体 X 具有负指数分布 $f(x)=\begin{cases}\frac{1}{\mu}e^{-\frac{1}{\mu}x}, x\ge 0\\ 0, x<0\end{cases}$，其中 $\mu>0$

试问 $\overline{X}$ 是否为 μ 的优效估计？

解： 因为 $\int_0^{\infty}\left(\frac{\partial \ln f(x)}{\partial x}\right)^2 f(x)dx=\int_0^{\infty}\left[\frac{\partial(-\ln\mu-\frac{1}{\mu}x)}{\partial\mu}\right]^2\frac{1}{\mu}e^{-\frac{1}{\mu}x}dx$

$$=\int_0^{\infty}(-\frac{1}{\mu}+\frac{1}{\mu^2}x)^2\frac{1}{\mu}e^{-\frac{1}{\mu}x}dx=\frac{1}{\mu^4}\int_0^{\infty}(x-\mu)^2\frac{1}{\mu}e^{-\frac{1}{\mu}x}dx=\frac{1}{\mu^2}$$

故得罗-克拉美下界：$I_R=\frac{\mu^2}{n}$，由于 $E\overline{X}=EX=\mu$，故 $\overline{X}$ 是 μ 的无偏估计，又 $D\overline{X}=\frac{1}{n}DX=\frac{\mu^2}{n}$，所以 $\overline{X}$ 是 μ 的优效估计。

Thank you!