哈爾濱工業大學

点估计

参数估计

参数估计问题是利用从总体抽样得到的信息来估计总体的某些参数或者参数的某些函数.

估计合格率 

估计降雪量

估计湖中的鱼数

参数估计问题的一般提法

设有一个统计总体，总体的分布函数为 $F(x, \theta)$，其中θ为未知参数. 现从该总体取样本 $X_1,X_2,\ldots,X_n$，要依据样本对参数θ作出估计，或估计θ的某个已知函数$g(\theta)$. 这类问题称为参数估计.

参数估计分类

参数估计
- 点估计
  - 矩估计
  - 最大似然估计
- 区间估计

例如，估计降雨量

预计今年的降雨量：550mm ——点估计

预计今年的降雨量：500-600mm —区间估计

点估计

点估计:就是由样本$X_1, X_2,\ldots X_n$确定一个统计量 $\hat{\theta}=\theta(X_1, X_2, \cdots, X_n)$

用它估计总体的未知参数θ，称为总体参数的估计量.当具体的样本抽出后，可求出样本统计量的值$\hat{\theta}=\theta(x_1, x_2, \cdots, x_n)$

用它作为总体参数的估计值，称作总体参数的点估计值.

下面学习两种点估计方法：矩估计法和最大似然估计法.

矩估计法

英国统计学家皮尔逊最早提出的.

统计思想：用样本矩估计总体矩，

用样本矩函数估计总体矩函数.

理论依据: 大数定律

$$\hat{\alpha}_k = A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow[n\to\infty]{P} \alpha_k = E(X^k), k = 1, 2, \cdots m.$$

$$\hat{\theta} = h(A_1, A_2, \cdots, A_m) \xrightarrow[n\to\infty]{P} \theta = h(\alpha_1, \alpha_2, \cdots, \alpha_m).$$

矩估计的步骤

设总体有m个未知参数$\theta_1,\theta_2,\cdots,\theta_m$,假定总体$X$的前$m$阶矩都存在. $X_1,X_2,\cdots X_n$ 是来自总体的样本.

(1) 求出总体X的前m阶原点矩,

$$\begin{cases} \alpha_1 = E(X) = q_1(\theta_1,\theta_2,\cdots,\theta_m), \\ \alpha_2 = E(X^2) = DX + (EX)^2 = q_2(\theta_1,\theta_2,\cdots,\theta_m), \\ \vdots \\ \alpha_m = E(X^m) = q_m(\theta_1,\theta_2,\cdots,\theta_m). \end{cases}$$

矩估计的步骤

(2) 解上面矩方程组，把未知参数用原点矩表示.

$$\begin{cases} \theta_1 = h_1(\alpha_1, \alpha_2, \cdots, \alpha_m), \\ \theta_2 = h_2(\alpha_1, \alpha_2, \cdots, \alpha_m), \\ \quad \vdots \\ \theta_m = h_m(\alpha_1, \alpha_2, \cdots, \alpha_m). \end{cases}$$

(3) 用样本各阶原点矩 $A_1, A_2, \cdots, A_m$ 代替总体各阶原点矩 $\alpha_1, \alpha_2, \cdots \alpha_m$，得各参数的矩估计

$$\hat{\theta}_k = h_k(A_1, A_2, \cdots A_m) \quad (k = 1, 2, \cdots, m),$$

其中 $A_k = \dfrac{1}{n}\sum_{i=1}^{n} X_i^k$

矩估计的步骤

在第(1)步中，也可以求总体的前m阶中心矩β_m来建立矩方程，最后用样本的前m阶中心矩B_m替换总体的前m阶中心矩β_m，得各参数的矩估计.

矩估计中采用的矩不同，所得矩估计也可能不同.

矩估计

例1 求总体均值$\mu=E(X)$与方差$\sigma^2=D(X)$的矩估计量.

解 (1) $\alpha_1 = E(X) = \mu,$

$$\alpha_2 = E(X^2) = D(X) + [E(X)]^2 = \sigma^2 + \mu^2.$$

(2) $\mu = \alpha_1,$

$$\sigma^2 = \alpha_2 - \alpha_1^2.$$

(3) $\hat{\mu} = A_1 = \bar{X},$

$$\hat{\sigma}^2 = A_2 - A_1^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2$$

$$= S^{*2}.$$

矩估计

例2 设总体X的概率密度为

$$f(x)=\begin{cases}(\alpha+1)x^{\alpha}, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

其中 $\alpha>-1$ 是未知参数，$X_1,X_2,\ldots,X_n$是取自X的样本，求参数α的矩估计量.

解

$$\alpha_1=E(X)=\int_0^1 x(\alpha+1)x^{\alpha}dx$$

$$=(\alpha+1)\int_0^1 x^{\alpha+1}dx=\frac{\alpha+1}{\alpha+2},\ \Rightarrow \alpha=\frac{2\alpha_1-1}{1-\alpha_1},$$

$$\Rightarrow \hat{\alpha}=\frac{2\bar{X}-1}{1-\bar{X}}$$ 为α的矩估计量.

矩估计

例3 设总体的分布列为

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中$\theta(0<\theta<1/2)$是未知参数,利用总体X的样本值3, 1, 3, 0, 3, 1, 2, 3，求θ的矩估计值.

解 $\alpha_1 = E(X) = 2\theta(1-\theta) + 2\theta^2 + 3(1-2\theta) = 3-4\theta,$

$$\theta = (3-\alpha_1)/4, \Rightarrow \hat{\theta} = (3-\bar{X})/4.$$

$$\bar{X} = (3+1+3+3+1+2+3)/8 = 2.$$

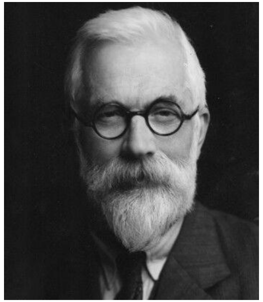
$$\Rightarrow \hat{\theta} = (3-2)/4 = 1/4.$$

最大似然估计(MLE)

最大似然估计是在总体分布类型已知条件下使用的一种参数估计方法.

Gauss

它首先是由德国数学家高斯在1821年提出的，英国统计学家费歇在1922年发现并首先研究了这种方法的一些性质，给这种方法命名为最大似然估计.



Fisher

最大似然估计法的基本思想

例1 罐中有黑、白围棋子若干，黑子与白子两者的数目之比是3:2, 但黑子与白子谁占3/5，谁占2/5不知道，今

从罐中有放回取4个棋子，观察结果为：白、黑、白、 白，请估计取到白子的概率p?

最大似然估计法的基本思想

解 设 $X=\begin{cases}1, & 取到白子,\\ 0, & 取到黑子,\end{cases}$ 则$X \sim B(1,p)$.

p是每次取到白子的概率，未知，可能取值为：

p=3/5 或 p=2/5.

今重复试验4次，得样本值 $x_1=1$，$x_2=0$, $x_3=1$，$x_4=1$.

$$P(x_1=1, x_2=0, x_3=1, x_4=1)=p^3(1-p).$$

大

p的取值	观察结果出现的概率
3/5	$(3/5)^3(2/5)=54/625$
2/5	$(2/5)^3(3/5)=24/625$

p的估计

最大似然估计法的基本思想

解 设 $X=\begin{cases}1, & 取到白子,\\ 0, & 取到黑子,\end{cases}$ 则 $X \sim B(1,p)$.

p是每次取到白子的概率，未知，可能取值为： p=3/5 或 p=2/5.

今重复试验4次，得样本值 $x_1=1$，$x_2=0$，$x_3=1$，$x_4=1$.

p的取值	观察结果出现的概率
3/5	$(3/5)^3(2/5)=54/625$
2/5	$(2/5)^3(3/5)=24/625$

大

p的估计

最大似然估计法的基本思想

p=3/5 时，试验结果发生的概率最大，因此取 $\hat{p}=3/5$ 更合理.

以上这种**选择一个参数使得实验结果具有最大概率**的思想就是最大似然法的基本思想 .

似然函数

$$P(x_1=1, x_2=0, x_3=1, x_4=1)=p^3(1-p).$$

最大似然估计的概念

设离散型总体X的分布列为

$P(X=x)=p(x;\theta_1,\cdots,\theta_m)$，$\theta_1,\cdots,\theta_m$为未知参数，

$X_1,\cdots,X_n$为样本，其观察值为$x_1,\cdots,x_n$，观察值

$(X_1=x_1,\cdots,X_n=x_n)$出现的概率为

$$L(\theta_1,\cdots,\theta_m)=P\{X_1=x_1,\cdots,X_n=x_n\}$$

$$=\prod_{i=1}^{n}p(x_i;\theta_1,\cdots,\theta_m),$$ 称为似然函数.

最大似然估计的概念

若统计量 $\hat{\theta}_1(X_1,\cdots,X_n),\cdots,\hat{\theta}_m(X_1,\cdots,X_n)$ 使得

$$L(\hat{\theta}_1,\cdots,\hat{\theta}_m)=\max_{\theta_1,\cdots,\theta_m} L(\theta_1,\cdots,\theta_m)$$

则称 $\hat{\theta}_1(X_1,\cdots,X_n),\cdots,\hat{\theta}_m(X_1,\cdots,X_n)$ 为 $\theta_1,\cdots,\theta_m$

的最大似然估计量 . $\hat{\theta}_i(x_1,\cdots,x_n), i=1,2,\cdots,m$

为 θ_i 的最大似然估计值.

最大似然估计的概念

设连续型总体X的概率密度为$f(x;\theta_1,\cdots,\theta_m)$，$\theta_1,\cdots,\theta_m$ 是未知参数，$X_1,\cdots,X_n$为样本，其观察值为$x_1,\cdots,x_n$，样本$X_1,\cdots,X_n$的联合概率密度

$$L(\theta_1,\cdots,\theta_m)=f(x_1,\cdots,x_n)=\prod_{i=1}^{n} f(x_i;\theta_1,\cdots,\theta_m),$$

为似然函数. 若存在 $\hat{\theta}_i(X_1,\cdots,X_n),i=1,\cdots,m.$

使得 $L(\hat{\theta}_1,\cdots,\hat{\theta}_m)=\max\limits_{\theta_1,\cdots,\theta_m} L(\theta_1,\cdots,\theta_m)$

则称$\hat{\theta}_i(X_1,\cdots,X_n)$为$\theta_i,i=1,\cdots,m$ 的最大似然估计量.

最大似然估计的概念

- $\ln x$是x的单调增函数，因此, 对数似然函数$\mathrm{Ln}\, L(\theta)$与似然函数$L(\theta)$的最大似然估计值是相同的.
- 若 $\hat{\theta}$ 是 θ 的最大似然估计，则 $g(\hat{\theta})$为 $g(\theta)$的最大似然估计.

也称此结论为最大似然估计的不变原理.

求最大似然估计的一般步骤

1. 写出似然函数

$$L = L(\theta_1,\theta_2,\cdots,\theta_m) = L(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_m)$$

$$= \begin{cases} \prod_{i=1}^{n} f(x_i;\theta_1,\theta_2,\cdots,\theta_m), & \text{连续总体;} \\ \prod_{i=1}^{n} P(X_i = x_i;\theta_1,\theta_2,\cdots,\theta_m). & \text{离散总体.} \end{cases}$$

求最大似然估计的一般步骤

2. 对似然函数取对数

$$\ln L=\begin{cases}\sum_{i=1}^{n}\ln f(x_i;\theta_1,\cdots,\theta_m), & \text{连续总体;}\\ \sum_{i=1}^{n}\ln P(X_i=x_i;\theta_1,\cdots,\theta_m), & \text{离散总体.}\end{cases}$$

3. 建立似然方程 $\dfrac{\partial \ln L(\theta_1,\cdots,\theta_m)}{\partial\theta_j}=0$ $(j=1,\cdots,m).$

4. 解似然方程，即可求出参数 θ_j 的最大似然估计. 若似然函数不可微，用定义求.

例2 设 $X_1,\cdots,X_n$ 是来自总体$X \sim B(1,p)$的一个样本，$0<p<1$是未知参数，求参数 p 的最大似然估计量.

解 总体X的分布列为

$$P(X=x;p)=p^x(1-p)^{1-x}, x=0,1$$

似然函数为

$$L(p)=P\{X_1=x_1,X_2=x_2,\cdots,X_n=x_n;p\}$$

$$=P\{X_1=x_1\}P\{X_2=x_2\}\cdots P\{X_n=x_n\}$$

$$L(p)=\prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

$$=p^{\sum\limits_{i=1}^{n}x_i}(1-p)^{n-\sum\limits_{i=1}^{n}x_i},(0<p<1)$$

对数似然函数为

$$\ln L(p)=(\sum_{i=1}^{n}x_i)\ln(p)+(n-\sum_{i=1}^{n}x_i)\ln(1-p)$$

对p求导并令其为0，得似然方程

$$\frac{d\ln L(p)}{dp}=\frac{1}{p}\sum_{i=1}^{n}x_i-\frac{1}{1-p}(n-\sum_{i=1}^{n}x_i)=0$$

解似然方程得 $\hat{p}=\dfrac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}$

参数p的最大似然估计量为 $\hat{p}=\bar{X}.$

若已获得n=8的样本值：0, 0, 1, 0, 1, 1, 1, 0,

参数p的最大似然估计值为

$$\hat{p}=\bar{x}=\frac{4}{8}=\frac{1}{2}.$$

例3 设总体X的概率密度为

$$f(x)=\begin{cases}(\alpha+1)x^{\alpha}, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

其中 $\alpha>-1$ 是未知参数, $X_1,X_2,\ldots,X_n$是取自X的样本,求参数α的最大似然估计量.

解

$$L(\alpha)=f(x_1,x_2,\cdots,x_n)=f(x_1)f(x_2)\cdots f(x_n)$$

$$=\prod_{i=1}^{n}f(x_i)=\begin{cases}\prod_{i=1}^{n}(\alpha+1)x_i^{\alpha}, & 0<x_i<1,\\ 0, & \text{其他},\end{cases}$$

$$L(\alpha)=\begin{cases}(\alpha+1)^n\prod\limits_{i=1}^{n}x_i^{\alpha}, & 0<x_i<1,\\ 0, & \text{其他},\end{cases}$$

$$\ln L(\alpha)=n\ln(\alpha+1)+\alpha\sum_{i=1}^{n}\ln x_i,\ (0<x_i<1)$$

$$\frac{d\ln L(\alpha)}{d\alpha}=\frac{n}{\alpha+1}+\sum_{i=1}^{n}\ln x_i=0,$$

矩估计

$$\boxed{\hat{\alpha}=\frac{2\bar{X}-1}{1-\bar{X}}}$$

$$\hat{\alpha}=-\frac{n}{\sum\limits_{i=1}^{n}\ln X_i}-1$$

为 α 的最大似然估计量.

例4 设总体 $X \sim N(\mu,\sigma^2)$, 参数μ, σ未知，$X_1, X_2, \cdots, X_n$为总体X的样本，求参数 μ,σ^2 最大似然估计量.

解 总体X的概率密度为

$$f(x;\mu,\sigma^2)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x\in R$$

$$L(\mu,\sigma^2)=\prod_{i=1}^{n} f(x_i)=\prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$=(2\pi\sigma^2)^{-\frac{n}{2}}e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}$$

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

$$\begin{cases} \dfrac{\partial \ln L(\mu,\sigma^2)}{\partial \mu} = \dfrac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \\ \dfrac{\partial \ln L(\mu,\sigma^2)}{\partial \sigma^2} = -\dfrac{n}{2\sigma^2} + \dfrac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \end{cases}$$

$$\hat{\mu} = \bar{X} \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

为 μ,σ^2 的最大似然估计量.

例5 设总体$X \sim U(0, \theta)$， $\theta>0$未知，$X_1,\cdots,X_n$ 为总体X的样本，求 θ 的最大似然估计量.

解 总体X的概率密度为

$$f(x;\theta)=\begin{cases}\theta^{-1}, 0\le x\le\theta,\\ 0, \quad 其他.\end{cases}$$

似然函数为 $L(\theta)=f(x_1)f(x_2)\cdots f(x_n)$

$$=\begin{cases}\theta^{-n}, 0\le x_1\cdots x_n\le\theta,\\ 0, \quad 其他.\end{cases}=\begin{cases}\theta^{-n}, 0\le x_{(1)}\le x_{(n)}\le\theta,\\ 0, \quad 其他.\end{cases}$$

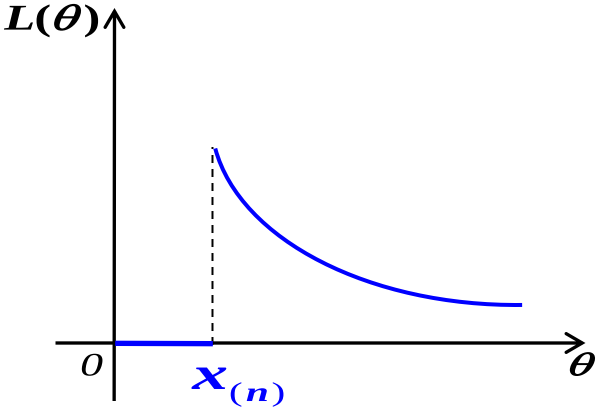
解 $$L(\theta)=\begin{cases}\theta^{-n}, 0\le x_{(1)}\le x_{(n)}\le\theta,\\ 0,\quad \text{其它}.\end{cases}$$

$L(\theta)$是θ的单调减函数，

当 $\theta = x_{(n)}$时，$L(\theta)=\theta^{-n}$取得最大值；

故θ的最大似然估计量为

$$\hat{\theta}=X_{(n)}=\max(X_1,\dots,X_n).$$



鉴定估计量的标准

对总体未知参数，用不同的估计方法得到的估计量不同，如何评价不同估计量的好坏呢？

常用的四条标准是

1．无偏性
2．有效性
3．相合性
4．有偏估计的均方误差准则

无偏性

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

$$E[\hat{\theta}(X_1,X_2,\cdots,X_n)]=\theta,$$

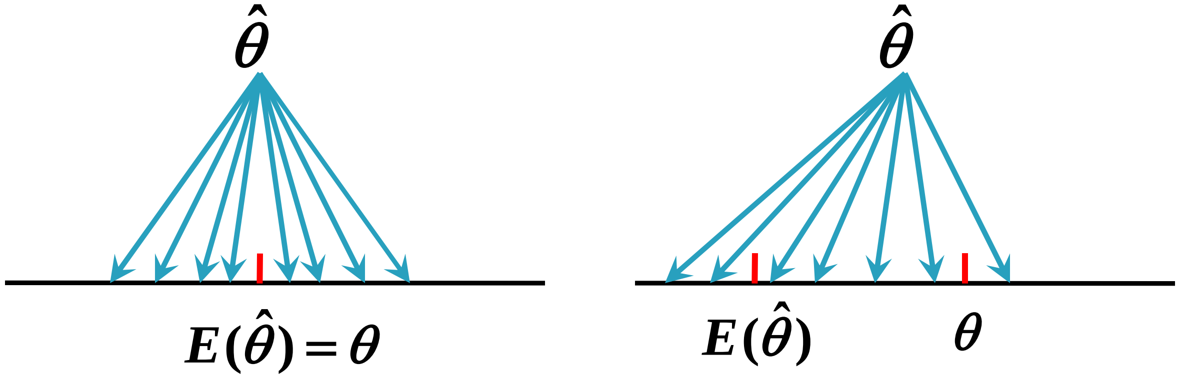
则称 $\hat{\theta}$ 是θ的无偏估计量.

若 $E(\hat{\theta})\neq\theta$, 称$|E(\hat{\theta})-\theta|$为估计量 $\hat{\theta}$的(绝对)偏差.

若$\lim\limits_{n\to\infty}E(\hat{\theta})=\theta$,称$\hat{\theta}$ 是θ的渐进无偏估计量.

无偏性的含义

- 无偏性 $E(\hat{\theta})=\theta \Leftrightarrow E(\hat{\theta}-\theta)=0.$



无偏性的实际意义是指在大量重复试验中，$\hat{\theta}(X_1,\ldots,X_n)$ 的平均估计为 θ ，即没有系统性的误差 .

例1 样本k阶原点矩 $A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$ 是总体k阶原点矩 $\alpha_k=E(X^k)$ 的无偏估计($k\geq 1$).

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

- 样本均值 $\bar{X}$ 是总体均值 $E(X)=\mu$ 的无偏估计.
- 样本方差 S^2 是总体方差 $D(X)=\sigma^2$ 的无偏估计.

  推导过程见第31讲例2.

◆ 样本二阶中心矩 $S^{*2}=B_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 是 $D(X)=\sigma^2$ 的渐进无偏估计.

因为
$$E(S^{*2})=E\left[\frac{(n-1)S^2}{n}\right]=\frac{(n-1)}{n}E(S^2)$$

$$=\frac{(n-1)}{n}\sigma^2\neq\sigma^2,$$ 不是 σ^2 无偏估计.

$$\lim_{n\to\infty}E(S^{*2})=\lim_{n\to\infty}\frac{n-1}{n}\sigma^2=\sigma^2.$$

修正方法 $\frac{n}{n-1}S^{*2}$ 是 σ^2 的无偏估计.

例2 设总体$X \sim U(0, \theta)$, $X_1, \ldots, X_n$ 为总体X的样本, 求 θ 的矩估计和最大似然估计量，并判断其是否为无偏估计?

解 矩估计

$$\alpha_1 = E(X) = \int_0^{\theta} \frac{1}{\theta} dx = \frac{\theta}{2} \Rightarrow \theta = 2\alpha_1 \Rightarrow \hat{\theta} = 2\bar{X}.$$

为θ的矩估计.

$$E(\hat{\theta}) = E(2\bar{X}) = 2E(\bar{X}) = 2E(X) = 2 \cdot \frac{\theta}{2} = \theta.$$

故 $\hat{\theta} = 2\bar{X}$是无偏估计.

θ的最大似然估计为 $\hat{\theta}_L = X_{(n)} = \max(X_1, \ldots, X_n)$.

(具体解法见第32讲例3.)

总体X的分布函数

$$F(x) = \begin{cases} 0, & x < 0, \\ x/\theta, & 0 \le x < \theta, \\ 1, & x \ge \theta. \end{cases}$$

由第22讲(4)最大值分布函数的公式有

$$F_{X_{(n)}}(x) = [F(x)]^n = \begin{cases} 0, & x < 0, \\ (x/\theta)^n, & 0 \le x < \theta, \\ 1, & x \ge \theta. \end{cases}$$

求导得$X(n)$的密度函数

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

$$E(\hat{\theta}_L)=E(X_{(n)})=\int_0^\theta x\cdot\frac{nx^{n-1}}{\theta^n}dx=\frac{n}{n+1}\theta\neq\theta.$$

故 $\hat{\theta}_L=X_{(n)}$ 不是θ的无偏估计量.

有效性

定义 设$\hat{\theta}_1=\hat{\theta}_1(X_1,\dots,X_n)$与$\hat{\theta}_2=\hat{\theta}_2(X_1,\dots,X_n)$都是$\theta$的无偏估计，若对任何可能的参数$\theta$都有

$$D(\hat{\theta}_1)\le D(\hat{\theta}_2),$$

且至少对某个参数值θ_0小于号成立，则称$\hat{\theta}_1$比$\hat{\theta}_2$有效.

例3 在例2中总体$X \sim U(0, \theta)$，未知参数θ的矩估计 $\hat{\theta} = 2\bar{X}$是无偏估计. 最大似然估计$\hat{\theta}_L = X_{(n)}$是有偏估计，修正为无偏估计为$\hat{\theta}_1 = \dfrac{n+1}{n} X_{(n)}$

问 $\hat{\theta}$与$\hat{\theta}_1$哪个更有效?$(n \geq 2)$

解

$$D(\hat{\theta}) = 4D(\bar{X}) = \frac{4D(X)}{n} = \frac{4}{n} \cdot \frac{\theta^2}{12} = \frac{\theta^2}{3n}.$$

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

$$E(X_{(n)})=\int_0^\theta x\cdot\frac{nx^{n-1}}{\theta^n}dx=\frac{n}{n+1}\theta.$$

$$E(X_{(n)}^2)=\int_0^\theta x^2\cdot\frac{nx^{n-1}}{\theta^n}dx=\frac{n}{n+2}\theta^2.$$

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

$$D(X_{(n)})=\frac{n\theta^2}{(n+2)(n+1)^2}.$$

$$D(\hat{\theta}_1)=D(\frac{n+1}{n}X_{(n)})=\frac{(n+1)^2}{n^2}D(X_{(n)})$$

$$=\frac{(n+1)^2}{n^2}\cdot\frac{n\theta^2}{(n+2)(n+1)^2}=\frac{\theta^2}{n(n+2)}.$$

当$n\geq 2$时，$D(\hat{\theta})=\dfrac{\theta^2}{3n}>D(\hat{\theta}_1)=\dfrac{\theta^2}{n(n+2)}.$

$\hat{\theta}_1$比$\hat{\theta}$更有效.

相合性

- **定义** 如果 $\hat{\theta}_n = \hat{\theta}_n(X_1, \ldots, X_n)$ 依概率收敛于θ, 即对任意的ε>0, 有

$$\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| \geq \varepsilon) = 0.$$

则称估计量 $\hat{\theta}_n = \hat{\theta}_n(X_1, \ldots, X_n)$ 是未知参数θ的相合(或一致)估计量.

例4 样本原点矩 $A_k=\dfrac{1}{n}\sum_{i=1}^{n}X_i^k$ 是总体原点矩

$\alpha_k=E(X^k)(k\geq 1)$的相合估计.

解 由样本 $X_1,\ldots,X_n$ 独立同分布，可得

$\forall k\geq 1, X_1^k,\ldots,X_n^k$也相互独立且与$X^k$同分布.

由大数定律，对任意ε>0, 有

$$\lim_{n\to\infty}P\left[\left|\frac{1}{n}\sum_{i=1}^{n}X_i^k-E(X^k)\right|\geq\varepsilon\right]=0.$$

A_k是α_k的相合估计. 特例$\bar{X}$是$E(X)$的相合估计.

例5 样本二阶中心矩 $S^{*2}=\dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2$ 和样本方差S^2都是总体方差 $D(X)=\sigma^2$ 的相合估计.

解

$$S^{*2}=\frac{1}{n}\sum_{i=1}^{n}X_i^2-(\bar{X})^2=A_2-(\bar{X})^2\xrightarrow[n\to\infty]{P}\alpha_2-\alpha_1^2=D(X)$$

故 S^{*2}是 $D(X)=\sigma^2$ 的相合估计.

又 $S^2=\dfrac{nS^{*2}}{n-1}$ ，从而S^2也是 $D(X)=\sigma^2$ 的相合估计.

有偏估计的均方误差准则

定义 设 $\hat{\theta}$ 是参数 θ 的估计量，称 $E(\hat{\theta}-\theta)^2$ 为 $\hat{\theta}$ 的**均方误差**. 记为 $MSE(\hat{\theta})$.

若 $\hat{\theta}$ 是 θ 的无偏估计量，则 $MSE(\hat{\theta})=D(\hat{\theta})$.

设 $\hat{\theta}_1$ 与 $\hat{\theta}_2$ 是参数 θ 的两个估计量，若

$$MSE(\hat{\theta}_1)\le MSE(\hat{\theta}_2)$$

且至少存在一个 θ 有严格不等号成立，则称在均方误差准则下 $\hat{\theta}_1$ 优于 $\hat{\theta}_2$.

均方误差分解式

$$
\begin{aligned}
MSE(\hat{\theta}) &= E(\hat{\theta}-\theta)^2 \\
&= E\{[\hat{\theta}-E(\hat{\theta})]+[E(\hat{\theta})-\theta]\}^2 \\
&= E[\hat{\theta}-E(\hat{\theta})]^2+[E(\hat{\theta})-\theta]^2 \\
&= D(\hat{\theta})+\delta^2
\end{aligned}
$$

其中 $\delta=|E(\hat{\theta})-\theta|$ 称为绝对偏差.

例6 设$X_1,\cdots,X_n$是正态总体$N(\mu,\sigma^2)$的一个样本，

$Q=\sum_{i=1}^{n}(x_i-\overline{x})^2$为样本的偏差平方和，求$c$使

cQ在均方误差准则下是σ^2的最优估计.

解 $MSE(cQ)=E(cQ-\sigma^2)^2$

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

$$=c^2D(Q)+[cE(Q)-\sigma^2]^2$$

在正态总体下 $\dfrac{(n-1)S^2}{\sigma^2}=\dfrac{Q}{\sigma^2}\sim\chi^2(n-1)$

$$E\left(\frac{Q}{\sigma^2}\right)=(n-1)\Rightarrow E(Q)=(n-1)\sigma^2,$$

$$D\left(\frac{Q}{\sigma^2}\right)=2(n-1)\Rightarrow D(Q)=2(n-1)\sigma^4.$$

$$MSE(cQ)=c^2D(Q)+[cE(Q)-\sigma^2]^2$$

$$=2c^2(n-1)\sigma^4+[c(n-1)-1]^2\sigma^4$$

$$=[(n^2-1)c^2-2(n-1)c+1]\sigma^4$$

$$MSE(cQ)=[(n^2-1)c^2-2(n-1)c+1]\sigma^4$$

对上式求导可得：当$c=\dfrac{1}{n+1}$，

$MSE(cQ)=\dfrac{2\sigma^4}{n+1}$最小，$\dfrac{Q}{n+1}$在均方误差准则下是$\sigma^2$的最优估计.

样本方差 S^2是σ^2的无偏估计，$D(S^2)=\dfrac{2\sigma^4}{n-1}$.

当$n>1$时，有$\dfrac{2\sigma^4}{n+1}<\dfrac{2\sigma^4}{n-1}$，说明在均方误差准则下，$\dfrac{Q}{n+1}$优于$S^2$.

谢　　谢！