Ch 9 参数估计

回顾前一次课

参数估计问题： 如何依据样本$X_1, X_2, \cdots, X_n$估计参数θ，或θ的函数$g(\theta)$

点估计

- 矩估计法：样本矩去估计总体矩求参数θ
- 最大似然估计: $\hat{\theta} = \mathrm{argmax}_\theta L(x_1, x_2, \cdots, x_m; \theta)$
- 最大似然估计的不变性

例题

设$X_1, X_2, \cdots, X_n$是来自总体X的样本, 以及总体X的密度函数为

$$f(x) = \begin{cases} \dfrac{1}{\theta} e^{-(x-\mu)/\theta} & x \ge \mu \\ 0 & 其它 \end{cases}$$

其中$\theta > 0$, 求μ和θ的最大似然估计.

估计量的评价标准

不同的估计方法可能得到不同的估计值,

问题: 采用哪一种估计量更好, 或更好的标准是什么呢?

常用标准:

- 无偏性
- 有效性
- 一致性

无偏性

设$X_1, X_2, \cdots, X_n$来自总体X的样本, 令$\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$是$\theta$的一个估计量, 若

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

则称**$\hat{\theta}$为θ的无偏估计**

无偏估计不要求估计值$\hat{\theta}$在任意情况下都等于θ，但在期望的情形下有$E(\hat{\theta}) = \theta$成立．其意义在于无系统性偏差，无偏性是一种对估计量常见而且重要的标准.

例题

样本k阶原点矩为总体k阶原点矩的无偏估计

设$X_1, X_2, \cdots, X_n$是总体X的样本，若$E[X^k]$存在，则$A_k = \sum_{i=1}^{n} X_i^k / n$是总体$a_k = E[X^k]$的无偏估计

设$X_1, X_2, \cdots, X_n$来自总体X的样本, 期望μ, 方差σ^2, 则:

1) $S_0^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2 / n$是$\sigma^2$的有偏估计;

2) $S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2 / (n-1)$是$\sigma^2$是$\sigma^2$的无偏估计.

无偏估计不具有函数不变性

注意$\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$是$\theta$的无偏估计，但并不一定有$g(\hat{\theta})$是$g(\theta)$的无偏估计, 因为

$$E[\hat{\theta}] = \theta\text{并不能推导出}E\left[g\left(\hat{\theta}\right)\right] = g(\theta)$$

例如： $E[\bar{X}] = E[X] = \mu$但$E[\bar{X}^2] \neq \mu^2$

例题

设$X_1, X_2, \cdots, X_n$是总体X的样本, 及总体X的概率密度为

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & x \geq 0 \\ 0 & \text{其他} \end{cases}$$

证明: $\bar{X} = \sum_{i=1}^{n} X_i / n$和$n\, min\{X_1, X_2, \cdots, X_n\}$均是$\theta$的无偏估计

有效性

参数可能存在多个无偏估计，若$\hat{\theta}_1$和$\hat{\theta}_2$都是θ的无偏估计, 则可以比较方差

$$\mathrm{Var}(\hat{\theta}_1) = E\left[(\hat{\theta}_1 - \theta)^2\right]$$

$$\mathrm{Var}(\hat{\theta}_2) = E\left[(\hat{\theta}_2 - \theta)^2\right]$$

一般而言: 方差越小, 无偏估计越好

设$\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \cdots, X_n)$ 和$\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \cdots, X_n)$是$\theta$的两个无偏估计, 若

$$\mathrm{Var}(\hat{\theta}_1) \leq \mathrm{Var}(\hat{\theta}_2)$$

则称$\boldsymbol{\theta_1}$**比**$\boldsymbol{\theta_2}$**有效**

例题

设$X_1, X_2, \cdots, X_n$是来自总体X的样本, 且X的概率密度为

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

证: 当$n>1$时$\bar{X}=\sum_{i=1}^{n}X_i/n$比$n\min\{X_1, X_2, \cdots, X_n\}$有效

例题

设 $X_1, X_2, \cdots, X_n$ 是总体 X 的样本，且 $E(X) = \mu$ 和 $\operatorname{Var}(X) = \sigma^2$．设常数 $c_1, c_2, \cdots, c_n \geq 0$ 满足 $\sum_{i=1}^n c_i = 1, c_i \neq 1/n$, 求证: $\bar{X}$ 比 $\sum_{i=1}^n c_i X_i$ 有效

有效统计量

随机变量X的概率密度为$f(x;\theta)$或分布函数为$F(x;\theta)$, 令

$$Var_0(\theta) = \left[nE\left(\frac{\partial \ln f(X;\theta)}{\partial \theta}\right)^2\right]^{-1}$$

$$Var_0(\theta) = \left[nE\left(\frac{\partial \ln F(X;\theta)}{\partial \theta}\right)^2\right]^{-1}$$

对任意的无偏估计量$\hat{\theta}$有$\mathrm{Var}(\hat{\theta}) \geq \mathrm{Var}_0(\theta)$,称$Var_0(\theta)$为估计量$\hat{\theta}$方差的下界.

当$Var(\hat{\theta}) = Var_0(\theta)$时称$\hat{\theta}$为达到方差下界的无偏估计量, 此时$\hat{\theta}$为最有效估计量, 简称有效估计量

例题

设$X_1, X_2, \cdots, X_n$是来自总体X的样本, 且X的概率密度为

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & x \geq 0 \\ 0 & \text{其他} \end{cases}$$

证明: θ的最大似然估计为有效估计量.

一致性

设$\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \cdots, X_n)$是$\theta$的一个估计量，若当$n \to \infty$时有$\hat{\theta}_n \xrightarrow{P} \theta$成立, 即对任意$\epsilon > 0$有

$$\lim_{n\to\infty} P[|\hat{\theta}_n - \theta| > \epsilon] = 0$$

则称$\hat{\theta}_n$为θ的一致估计量

估计量的一致性刻画了在足够多样本情形下估计量$\hat{\theta}$能有效逼近真实值θ，一致性是对估计的基本要求，不满足一致性的估计量一般不予考虑.

一致性的充分条件

定理：设$\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \cdots, X_n)$是$\theta$的一个估计量，若满足以下两个条件:

$$\lim_{n\to\infty} E\left[\hat{\theta}_n\right] = \theta \qquad \lim_{n\to\infty} \mathrm{Var}\left(\hat{\theta}_n\right) = 0$$

则$\hat{\theta}_n$为θ的一致估计量.

定理

设$\hat{\theta}_{n_1},\hat{\theta}_{n_2},\cdots,\hat{\theta}_{n_k}$分别为$\theta_1,\theta_2,\cdots,\theta_k$满足一致性的估计量，对连续函数$g:\mathrm{R}^n\to\mathrm{R}$，有函数$\hat{\eta}_n=g(\hat{\theta}_{n_1},\hat{\theta}_{n_2},\cdots,\hat{\theta}_{n_k})$是$\eta=g(\theta_1,\theta_2,\cdots,\theta_k)$满足一致性的估计量

估计量的评价标准

设$X_1, X_2, \cdots, X_n$是来自总体X的样本，且X的概率密度为

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & x \geq 0 \\ 0 & \text{其他} \end{cases}$$

则样本均值$\bar{X}_n = \sum_{i=1}^{n} X_i / n$为$\theta$的无偏、有效、一致估计量.

例题

设$X_1, X_2, \cdots, X_n$是来自总体$X \sim U(0,\theta)$的样本，证明：θ的最大似然估计量是一致估计量.