第七章

参数估计

统计推断DE基本问题

- 参数估计问题
  - 点估计
  - 区间估计
- 假设检验问题

什么是参数估计？

参数是刻画总体某方面的概率特性的数量.

当这个数量是未知的时候，从总体抽出一个样本，用某种方法对这个未知参数进行估计就是参数估计.

例如，$X \sim N(\mu, \sigma^2)$,

若μ, σ^2未知，通过构造样本的函数, 给出它们的估计值或取值范围就是参数估计的内容.

点估计　区间估计

参数估计的类型

点估计 —— 估计未知参数的值

区间估计—— 估计未知参数的取值范围，使得这个范围包含未知参数真值的概率为给定的值.

§ 7.1　点估计及评选标准

一、点估计的思想方法

设总体X 的分布函数的形式已知，但它含有一个或多个未知参数：$\theta_1, \theta_2, \ldots, \theta_k$

设 $X_1, X_2, \ldots, X_n$为总体的一个样本

构造 k 个统计量：

$$\left.\begin{array}{l} \theta_1(X_1, X_2, \cdots, X_n) \\ \theta_2(X_1, X_2, \cdots, X_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \theta_k(X_1, X_2, \cdots, X_n) \end{array}\right\} \boxed{\text{随机变量}}$$

当测得一组样本值($x_1, x_2, \ldots, x_n$)时，代入上述统计量，即可得到 k 个数：

$$\left.\begin{array}{l}\hat{\theta}_1(x_1, x_2, \cdots, x_n) \\ \hat{\theta}_2(x_1, x_2, \cdots, x_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \hat{\theta}_k(x_1, x_2, \cdots, x_n)\end{array}\right\} \boxed{\text{数值}}$$

称数 $\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_k$ 为未知参数 $\theta_1, \theta_2, \cdots, \theta_k$ 的估计值

对应的统计量为未知参数 $\theta_1, \theta_2, \cdots, \theta_k$ 的估计量

问题　如何构造统计量？

如何评价估计量的好坏？

二、矩估计法

方法

用样本的 k 阶矩作为总体的 k 阶矩的估计量, 建立含有待估计参数的方程，从而可解出待估计参数.

一般地，不论总体服从什么分布，总体期望 μ 与方差 σ^2 存在，则它们的矩估计量分别为

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

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

事实上，按矩法原理，令

$$\begin{cases} \mu_1 = \mu = \overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i \\ \mu_2 = E(X^2) = A_2 = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2 \end{cases}$$

$$\Rightarrow \begin{cases} \hat{\mu} = \overline{X} \\ \hat{\sigma}^2 = E(\hat{X}^2) - E^2(\hat{X}) = A_2 - \hat{\mu}^2 \end{cases}$$

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

设待估计的参数为 $\theta_1,\theta_2,\cdots,\theta_k$

设总体的 r 阶矩存在，记为

$$E(X^r)=\mu_r(\theta_1,\theta_2,\cdots,\theta_k)$$

设 $X_1, X_2,\ldots, X_n$为一样本，样本的 r 阶矩为

$$A_r=\frac{1}{n}\sum_{i=1}^{n}X_i^r$$

令 $\mu_r(\theta_1,\theta_2,\cdots,\theta_k)=A_r=\frac{1}{n}\sum_{i=1}^{n}X_i^r \quad r=1,2,\cdots,k$

—— 含未知参数 $\theta_1,\theta_2,\ldots,\theta_k$ 的方程组

解方程组，得 k 个统计量：

$$\left.\begin{array}{l}\hat{\theta}_1(X_1, X_2, \cdots, X_n)\\ \hat{\theta}_2(X_1, X_2, \cdots, X_n)\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \hat{\theta}_k(X_1, X_2, \cdots, X_n)\end{array}\right\}$$

——未知参数 $\theta_1, \theta_2, \ldots, \theta_k$ 的矩估计量

代入一组样本值得 k 个数:

$$\left.\begin{array}{l}\hat{\theta}_1 = \hat{\theta}_1(x_1, x_2, \cdots, x_n)\\ \hat{\theta}_2 = \hat{\theta}_2(x_1, x_2, \cdots, x_n)\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \hat{\theta}_k = \hat{\theta}_k(x_1, x_2, \cdots, x_n)\end{array}\right\}$$

——未知参数 $\theta_1, \theta_2, \ldots, \theta_k$ 的矩估计值

例1 设总体 $X \sim N(\mu, \sigma^2)$, $X_1, X_2, \ldots, X_n$为总体的样本，求 μ, σ^2 的矩法估计量。

解 $\hat{\mu}_{矩} = \overline{X}$

$$\hat{\sigma}^2_{矩} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2$$

例2 设总体 $X \sim E(\lambda)$, $X_1, X_2, \ldots, X_n$为总体的样本，求λ的矩法估计量。

解 $\mu_1 = E(X) = \int_0^{+\infty} x\lambda e^{-\lambda x} dx = \frac{1}{\lambda}$ 　　令 $\overline{X} = \frac{1}{\lambda}$

故 $\hat{\lambda}_{矩} = \frac{1}{\overline{X}}$

例3　设总体X的概率密度为

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

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.

解　$\mu_1=E(X)=\int_0^1 x(\alpha+1)x^{\alpha}dx$

数学期望是一阶原点矩

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

由矩法,

$$\overline{X}=\frac{\alpha+1}{\alpha+2}$$

样本矩　总体矩

从中解得 $\hat{\alpha}=\dfrac{2\overline{X}-1}{1-\overline{X}}$, 即为$\alpha$的矩估计.

例4 设总体 $X \sim U(a, b)$, a, b 未知，求 a, b 的矩法估计量.

解 由于 $E(X) = \dfrac{a+b}{2}$, $D(X) = \dfrac{(b-a)^2}{12}$

$$E(X^2) = D(X) + E^2(X) = \frac{(b-a)^2}{12} + \left(\frac{a+b}{2}\right)^2$$

令

$$\begin{cases} \dfrac{\hat{a}+\hat{b}}{2} = \overline{X} \\ \dfrac{(\hat{b}-\hat{a})^2}{12} + \left(\dfrac{\hat{a}+\hat{b}}{2}\right)^2 = A_2 = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2 \end{cases}$$

解得

$$\hat{a}_{矩} = \overline{X} - \sqrt{3(A_2 - \overline{X}^2)}$$

$$= \overline{X} - \sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

$$\hat{b}_{矩} = \overline{X} + \sqrt{3(A_2 - \overline{X}^2)}$$

$$= \overline{X} + \sqrt{\frac{3}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

三、极大似然估计法

思想方法：一次试验就出现的
事件有较大的概率

例如: 有两个外形相同的箱子,都装有100个球

一箱　99个白球，　1个红球

一箱　1个白球，　99个红球

现从两箱中任取一箱，并从箱中任取一球，结果所取得的球是白球。

问　所取的球来自哪一箱？

答　第一箱.

例5 设总体 X 服从0-1分布，且$P\{X = 1\} = p$，用极大似然法求 p 的估计值。

解 X 的概率分布可以写成

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

设 $X_1, X_2, \ldots, X_n$ 为总体 X 的样本，

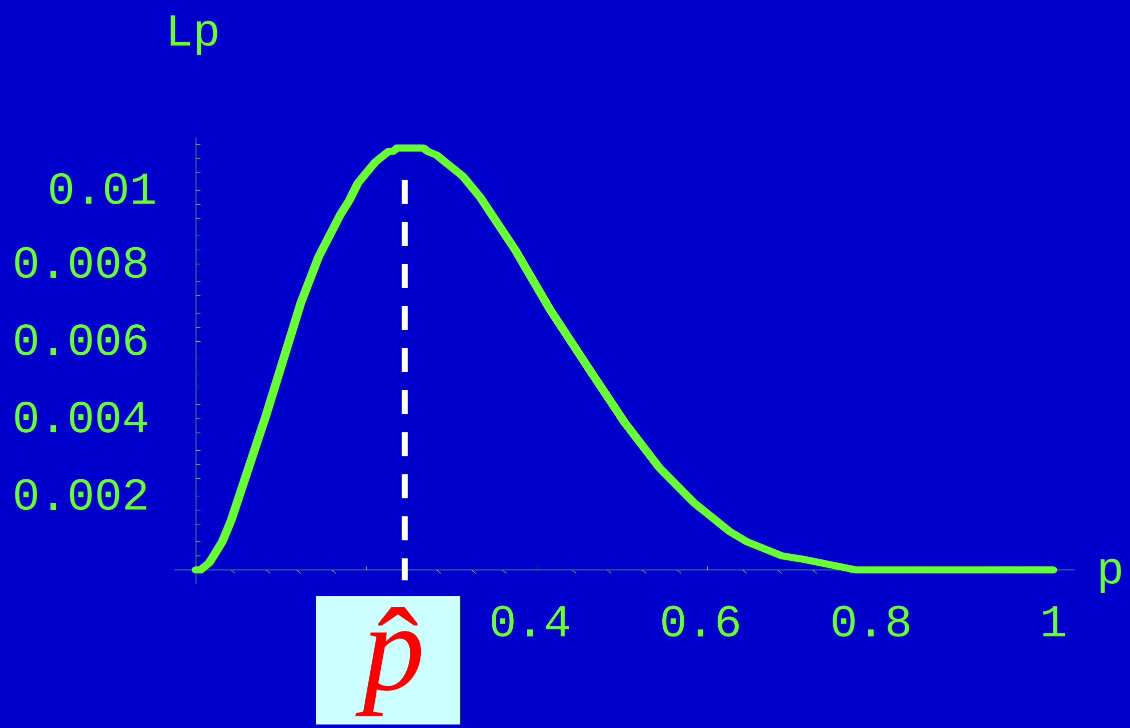
设 $x_1, x_2, \ldots, x_n$ 为总体 X 的样本值，

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

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

$$x_i = 0,1, i = 1,2,\cdots,n$$

对于不同的 p ,$L(p)$不同，见右下图



现经过一次试验， 事件

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

发生了， 则 p 的取值应使这个事件发生的概率最大。

在容许的范围内选择 **p**，使$L(p)$最大

注意到，$\ln L(p)$是 L 的单调增函数，故 若某个p 使$\ln L(p)$最大，则这个p 必使$L(p)$最大。

$$\frac{\mathrm{d}\ln L}{\mathrm{d}p}=\frac{\sum_{i=1}^{n}x_i}{p}-\frac{n-\sum_{i=1}^{n}x_i}{1-p}\overset{令}{=}0 \quad \Rightarrow \hat{p}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}$$

$$\left(\frac{\mathrm{d}^2\ln L}{\mathrm{d}p^2}=-\frac{\sum_{i=1}^{n}x_i}{p^2}-\frac{n-\sum_{i=1}^{n}x_i}{(1-p)^2}<0\right)$$

所以 $\hat{p}=\bar{x}$ 为所求 p 的估计值.

一般地，设X为离散型随机变量，其分布律为

$$P\{X=x\}=f(x,\theta),\quad x=u_1,u_2,\cdots,\theta\in\Theta$$

$X_1, X_2,\ldots, X_n$为总体 X 的样本,

$x_1, x_2,\ldots, x_n$为总体 X 的样本值,

则$X_1, X_2,\ldots, X_n$的分布律为

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

$$=f(x_1,\theta)f(x_2,\theta)\cdots f(x_n,\theta)$$

$$\overset{\text{记为}}{=} L(x_1,x_2,\cdots,x_n,\theta) \overset{\text{或}}{=} L(\theta)\quad x_i=u_1,u_2,\cdots,$$

$$i=1,2,\cdots,n,\theta\in\Theta$$

称$L(\theta)$为样本的似然函数

当给定一组样本值时，$L(\theta)$就是参数θ的函数,
极大似然估计法的思想就是：

选择适当的$\theta=\hat{\theta}$,使$L(\theta)$取最大值，即

$$L(x_1,x_2,\cdots,x_n,\hat{\theta})$$
$$=\max_{\theta\in\Theta}\{f(x_1,\theta)f(x_2,\theta)\cdots f(x_n,\theta)\}$$

则称这样得到的$\hat{\theta}$

$$\hat{\theta}=g(x_1,x_2,\cdots,x_n)$$

为参数 θ 的极大似然估计值 简记 $\hat{\theta}_{mle}$

称统计量

$$\hat{\theta}=g(X_1,X_2,\cdots,X_n)$$

为参数 θ的极大似然估计量 简记 $\hat{\theta}_{MLE}$

注1 若随机变量X 连续, 取 $f(x_i,\theta)$为X_i 的密度函数

似然函数为 $L(\theta)=\prod_{i=1}^{n} f(x_i,\theta)$

注2 未知参数个数可以不止一个, 如$\theta_1,\ \theta_2,\ldots,\ \theta_k$

设X 的概率密度(或分布律)为 $f(x,\theta_1,\theta_2,\cdots,\theta_k)$

则定义似然函数为

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

$$=L(\theta_1,\theta_2,\cdots,\theta_k)=\prod_{i=1}^{n} f(x_i,\theta_1,\theta_2,\cdots,\theta_k)$$

$$-\infty<x_i<+\infty, i=1,2,\cdots,n \quad (\theta_1,\theta_2,\cdots,\theta_k)\in\Theta$$

若 $L(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_k)$ 关于 $\theta_1, \theta_2,\ldots, \theta_k$ 可微，则称

$$\frac{\partial}{\partial\theta_r}L(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_k)=0 \quad r=1,2,\cdots,k$$

为似然方程组

若对于某组给定的样本值 $x_1, x_2,\ldots, x_n$，参数 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$ 使得似然函数取得最大值，即

$$L(x_1,x_2,\cdots,x_n;\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k)$$

$$=\max_{(\theta_1,\theta_2,\cdots,\theta_k)\in\Theta}\{L(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_k)\}$$

则称 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$ 为 $\theta_1, \theta_2,\ldots, \theta_k$ 的极大似然估计值

显然，

$$\hat{\theta}_r = g(x_1, x_2, \cdots, x_n) \qquad r = 1,2,\cdots,k$$

称统计量

$$\hat{\theta}_r = g(X_1, X_2, \cdots, X_n) \qquad r = 1,2,\cdots,k$$

为 $\theta_1, \theta_2, \ldots, \theta_k$ 的极大似然估计量

例6 设总体 $X \sim N(\mu, \sigma^2)$, $x_1, x_2, \ldots, x_n$ 是 X 的一组样本值，求 μ, σ^2 的极大似然估计.

解 $L(x_1, x_2, \cdots, x_n; \mu, \sigma^2)$

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

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

似然方程组为

$$\begin{cases} \left(\dfrac{\partial}{\partial\mu}\ln L\right) = \dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(x_i-\mu)=0 \\ \left(\dfrac{\partial}{\partial(\sigma^2)}\ln L\right) = \dfrac{1}{2(\sigma^2)^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2 - \dfrac{n}{2(\sigma^2)}=0 \end{cases}$$

$$\Rightarrow \begin{cases} \hat{\mu}_{mle} = \dfrac{1}{n}\sum\limits_{i=1}^{n}x_i = \bar{x} \\ \hat{\sigma}^2_{mle} = \dfrac{1}{n}\sum\limits_{i=1}^{n}(x_i-\bar{x})^2 \end{cases}$$

μ, σ^2 的极大似然估计量分别为

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

求未知参数的极大似然估计值(量)的方法

1） 写出似然函数L

2）求出 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$ ，使得

$$L(x_1,x_2,\cdots,x_n;\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k)$$

$$=\max_{(\theta_1,\theta_2,\cdots,\theta_k)\in\Theta}\{L(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_k)\}$$

若 L 是 $\theta_1,\theta_2,\cdots,\theta_k$ 的可微函数, 解似然方程组

$$\frac{\partial}{\partial\theta_r}\ln L(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_k)=0$$

$$r=1,2,\cdots,k$$

可求得未知参数的极大似然估计值 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$

然后, 再求得极大似然估计量.

若 L 不是$\theta_1,\theta_2,\cdots,\theta_k$ 的可微函数, 需用其它方法求极大似然估计值. 请看下例：

例7 设 $X \sim U(a,b)$, $x_1, x_2, \ldots, x_n$ 是 X 的一个样本, 求 a, b 的极大似然估计值与极大似然估计量.

解 X 的密度函数为

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

似然函数为

$$L(x_1,x_2,\cdots,x_n;a,b)=\begin{cases}\dfrac{1}{(b-a)^n}, & a<x_i<b,\ i=1,2,\cdots,n\\ 0, & \text{其它}\end{cases}$$

似然函数只有当 $a < x_i < b,\ i = 1,2,\ldots,n$ 时才能获得最大值, 且 a 越大, b 越小, L 越大.

令 $$x_{\min} = \min\{x_1, x_2,\ldots, x_n\}$$
$$x_{\max} = \max\{x_1, x_2,\ldots, x_n\}$$

取
$$\hat{a} = x_{\min},\quad \hat{b} = x_{\max}$$

则对满足 $a \le x_{\min} \le x_{\max} \le b$ 的一切 $a < b$, 都有

$$\frac{1}{(b-a)^n} \le \frac{1}{(x_{\max} - x_{\min})^n}$$

故 $\hat{a} = x_{\min}, \ \hat{b} = x_{\max}$

是 ***a*** , ***b*** 的极大似然估计值.

$$X_{\min} = \min\{X_1, X_2, \cdots, X_n\}$$

$$X_{\max} = \max\{X_1, X_2, \cdots, X_n\}$$

分别是 ***a*** , ***b*** 的极大似然估计量.

四、点估计的评价标准

对于同一个未知参数，不同的方法得到的估计量可能不同,于是提出问题

- 应该选用哪一种估计量?
- 用什么标准来评价一个估计量的好坏?

常用标准

(1) 一致性

(2) 无偏性

(3) 有效性

1、一致性

定义 设 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$ 是总体参数θ 的估计量. 若对于任意的$\theta\in\Theta$, 当$n\to\infty$时, $\hat{\theta}$ 依概率收敛于θ, 即 $\forall\varepsilon>0,$

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

则称 $\hat{\theta}$ 是总体参数θ的一致(或相合)估计量.

一致性估计量仅在样本容量 n 足够大时,才显示其优越性.

关于一致性的两个常用结论

1. 样本 k 阶矩是总体 k 阶矩的一致性估计量. } 由大数定律证明

2. 设 $\hat{\theta}$ 是 θ 的无偏估计量, 且 $\lim\limits_{n\to\infty} D(\hat{\theta}) = 0$, 则 $\hat{\theta}$ 是 θ 的一致估计量. } 用切贝雪夫不等式证明

矩法得到的估计量一般为一致估计量

在一定条件下, 极大似然估计具有一致性

2、无偏性

定义 设 $(X_1, X_2, \cdots, X_n)$ 是总体X 的样本

$\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$ 是总体参数θ 的估计量

$E(\hat{\theta})$ 存在, 且对于任意$\theta \in \Theta$ 都有

$$E(\hat{\theta}) = \theta$$

则 $\hat{\theta}$ 是θ 的无偏估计量.

例8 设总体X的 k 阶矩 $\mu_k = E(X^k)$ 存在

$(X_1, X_2, \cdots, X_n)$ 是总体X的样本,

证明：不论 X 服从什么分布，$A_k = \dfrac{1}{n}\sum_{i=1}^{n} X_i^k$

是 μ_k 的无偏估计量.

证 由于 $E(X_i^k) = \mu_k \quad i = 1,2,\cdots,n$ 因而

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

$$= \frac{1}{n}\cdot n \cdot \mu_k = \mu_k$$

特别地,

样本均值 $\overline{X}$ 是总体期望 $E(X)$ 的无偏估计量

样本二阶原点矩 $A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ 是总体二阶原点矩 $\mu_2 = E(X^2)$ 的无偏估计量

例9 设总体 X 的期望 $E(X)$与方差 $D(X)$存在, $(X_1, X_2, \cdots, X_n)$ 是 X 的一个样本, $n > 1$. 证明

(1) $S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ 不是 $D(X)$ 的无偏估计量;

(2) $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ 是 $D(X)$ 的无偏估计量.

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

$$E(X_i) = E(X) = \mu, \; D(X_i) = D(X) = \sigma^2$$

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

因而 $E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^2)-E(\bar{X}^2)$

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

$$=\frac{n-1}{n}\sigma^2\neq\sigma^2$$

故 $E\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)=\sigma^2$ 证毕.

例10 设 $(X_1, X_2, \cdots, X_m)$ 是总体 X 的一个样本，$X \sim B(n,p)$ $n>1$， 求 p^2 的无偏估计量.

解 由于样本矩是总体矩的无偏估计量以及数学期望的线性性质，只要将未知参数表示成总体矩的线性函数，然后用样本矩作为总体矩的估计量，这样得到的未知参数的估计量即为无偏估计量.

令 $$\overline{X} = E(X) = np$$

$$\frac{1}{m}\sum_{i=1}^{m} X_i^2 = E(X^2) = (np)^2 + np(1-p)$$

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

因此, p^2 的无偏估计量为

$$\hat{p}^2=\frac{1}{n^2-n}\left(\frac{1}{m}\sum_{i=1}^{m}X_i^2-\overline{X}\right)$$

$$=\frac{\frac{1}{m}\sum_{i=1}^{m}X_i(X_i-1)}{n(n-1)}$$

例12 设总体 X 的密度函数为

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

$(X_1,X_2,\cdots,X_n)$ 为 X 的一个样本

证明 $\overline{X}$ 与 $n\min\{X_1,X_2,\cdots,X_n\}$ 都是 θ 的无偏估计量

证 $X\sim E\left(\frac{1}{\theta}\right)$ $\qquad E(X)=\theta$

故 $E(\overline{X})=E(X)=\theta$

$\overline{X}$ 是 θ 的无偏估计量.

令 $Z = \min\{X_1, X_2, \cdots, X_n\}$

$$F_Z(z) = 1 - P(X_1 > z, X_2 > z, \cdots, X_n > z)$$
$$= 1 - P(X_1 > z)P(X_2 > z)\cdots P(X_n > z)$$
$$= 1 - \prod_{i=1}^{n}(1 - P(X_i \le z)) = \begin{cases} 0 & z < 0 \\ 1 - e^{-\frac{nz}{\theta}} & z \ge 0 \end{cases}$$

$$f_Z(z) = \begin{cases} 0 & z < 0 \\ \frac{n}{\theta} e^{-\frac{nz}{\theta}} & z \ge 0 \end{cases}$$

即 $Z \sim E\left(\frac{n}{\theta}\right)$ $\qquad E(Z) = \frac{\theta}{n}$ $\qquad E(nZ) = \theta$

故 nZ 是 θ 的无偏估计量.

3、有效性

定义 设 $\hat{\theta}_1 = \theta_1(X_1, X_2, \cdots, X_n)$

$$\hat{\theta}_2 = \theta_2(X_1, X_2, \cdots, X_n)$$

都是总体参数θ的无偏估计量, 且

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

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

例13 设$X_1, X_2, \cdots, X_n$)为 X 的一个样本, 密度函数为

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

由前面例4 可知, $\overline{X}$ 与 $n\min\{X_1, X_2, \cdots, X_n\}$ 都

是 θ 的无偏估计量,问哪个估计量更有效？

解 $D(\overline{X})=\dfrac{\theta^2}{n}$ $D(n\min\{X_1, X_2, \cdots, X_n\})=\theta^2$

所以, $\overline{X}$ 比$n\min\{X_1, X_2, \cdots, X_n\}$ 更有效.

例14 设总体期望为 $E(X)=\mu$，方差 $D(X)=\sigma^2$

$(X_1, X_2, \cdots, X_n)$ 为总体 X 的一个样本

(1) 设常数 $c_i \neq \frac{1}{n}$ $i=1,2,\cdots,n.$ $\sum_{i=1}^{n} c_i = 1.$

证明 $\hat{\mu}_1 = \sum_{i=1}^{n} c_i X_i$ 是 μ 的无偏估计量

(2) 证明 $\hat{\mu} = \overline{X}$ 比 $\hat{\mu}_1 = \sum_{i=1}^{n} c_i X_i$ 更有效

证: (1) $E(\hat{\mu}_1) = \sum_{i=1}^{n} c_i E(X_i) = \sum_{i=1}^{n} c_i \mu = \mu$

(2) $$D(\hat{\mu}_1)=\sum_{i=1}^{n}c_i^2 D(X_i)=\sigma^2\sum_{i=1}^{n}c_i^2$$

而 $$1=\left(\sum_{i=1}^{n}c_i\right)^2=\sum_{i=1}^{n}c_i^2+2\sum_{1\le i<j\le n}c_ic_j$$

$$<\sum_{i=1}^{n}c_i^2+\sum_{1\le i<j\le n}(c_i^2+c_j^2)=n\sum_{i=1}^{n}c_i^2$$

$$\Rightarrow \sum_{i=1}^{n}c_i^2>\frac{1}{n} \Rightarrow D(\hat{\mu})=\frac{1}{n}\sigma^2<D(\hat{\mu}_1)$$

结论 算术均值比加权均值更有效.

例如 $X \sim N(\mu, \sigma^2)$, (X_1, X_2) 是一样本.

$$\left.\begin{aligned} \hat{\mu}_1 &= \frac{2}{3}X_1 + \frac{1}{3}X_2 \\ \hat{\mu}_2 &= \frac{1}{4}X_1 + \frac{3}{4}X_2 \\ \hat{\mu}_3 &= \frac{1}{2}X_1 + \frac{1}{2}X_2 \end{aligned}\right\} \text{都是}\mu\text{的无偏估计量}$$

由例14(2) 知 $\hat{\mu}_3$ 最有效.

罗—克拉美（*Rao – Cramer*） 不等式

若 $\hat{\theta}$ 是参数 θ 的无偏估计量, 则

$$D(\hat{\theta}) \geq \frac{1}{nE\left[\frac{\partial}{\partial\theta}\ln p(X,\theta)\right]^2} = D_0(\theta)$$

其中 $\boldsymbol{p(x,\theta)}$ 是总体 $\boldsymbol{X}$ 的概率分布或密度函数；$D_0(\theta)$ 称为方差的下界

当 $D(\hat{\theta}) = D_0(\theta)$ 时; 称 $\hat{\theta}$ 为达到方差下界的无偏估计量, 此时称 $\hat{\theta}$ 为最有效的估计量，简称有效估计量.

例8 设总体 X 的密度函数为

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

$(x_1,x_2,\cdots,x_n)$ 为 X 的一个样本值.

求 θ 的极大似然估计量，并判断它是否是达到方差下界的无偏估计量.

解 由似然函数

$$L(\theta)=\frac{1}{\theta^n}e^{-\frac{\sum_{i=1}^{n}x_i}{\theta}} \longrightarrow \ln L(\theta)=-n\ln\theta-\frac{\sum_{i=1}^{n}x_i}{\theta}$$

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta) = -\frac{n}{\theta} + \frac{\sum_{i=1}^{n} x_i}{\theta^2} \overset{令}{=} 0$$

$$\longrightarrow \quad \hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$$

θ的极大似然估计量为 $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$

它是θ的无偏估计量.

$$D(\hat{\theta}) = D\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{\theta^2}{n}$$

而 $$\ln f(x,\theta) = -\ln\theta - \frac{x}{\theta}$$

$$\left[\frac{\partial}{\partial\theta}\ln f(x,\theta)\right]^2 = \left[-\frac{1}{\theta} + \frac{x}{\theta^2}\right]^2$$

$$E\left[\frac{\partial}{\partial\theta}\ln f(X,\theta)\right]^2 = E\left[-\frac{1}{\theta} + \frac{X}{\theta^2}\right]^2 = \frac{1}{\theta^2}$$

$$\Longrightarrow \quad \frac{1}{nE\left[\frac{\partial}{\partial\theta}\ln f(X,\theta)\right]^2} = \frac{\theta^2}{n} = D(\overline{X})$$

故 $\overline{X}$ 是达到方差下界的无偏估计量.