第二章　参数估计

2.1　点估计与优良性

2.2　点估计量的求法

2.3　最小方差无偏估计与有效估计

2.4　区间估计

点估计的概念

设总体X的分布函数形式已知, 但它的一个或多个参数未知, 借助于总体X的一个样本来估计总体未知参数的值的问题称为**点估计问题**.

例1 在某炸药制造厂,一天中发生着火现象的次数X是一个随机变量,假设它服从以$\lambda > 0$为参数的泊松分布,参数λ为未知,设有以下的样本值,试估计参数λ.

着火次数k	0	1	2	3	4	5	6	
发生k次着火的天数n_k	75	90	54	22	6	2	1	$\Sigma=250$

证 因为 $X \sim P(\lambda)$, 所以 $\lambda = E(X)$.

用样本均值来估计总体的均值 $E(X)$

$$\bar{x} = \frac{\sum_{k=0}^{6} k n_k}{\sum_{k=0}^{6} n_k} = \frac{1}{250}[0\times 75 + 1\times 90 + 2\times 54 + 3\times 22 + 4\times 6 + 5\times 2 + 6\times 1] = 1.22$$

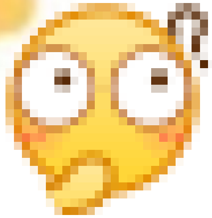
故 $E(X) = \lambda$ 的估计为 1.22.

点估计问题的一般提法

设总体 X 的分布函数 $F(x;\theta)$的形式为已知，θ是待估参数. $X_1, X_2, \cdots, X_n$是 X 的一个样本，$x_1, x_2, \cdots, x_n$为相应的一个样本值.

点估计问题就是要构造一个适当的统计量 $\hat{\theta}(X_1, X_2, \cdots, X_n)$,用它的观察值 $\hat{\theta}(x_1, x_2, \cdots, x_n)$来估计未知参数 θ.

$$\left.\begin{array}{l}\hat{\theta}(X_1, X_2, \cdots, X_n)\text{称为}\theta\text{的估计量.}\\ \hat{\theta}(x_1, x_2, \cdots, x_n)\text{称为}\theta\text{的估计值.}\end{array}\right\}\begin{array}{l}\text{通称估计,}\\ \text{简记为}\hat{\theta}.\end{array}$$



$\hat{\theta}(X_1, X_2, \cdots, X_n)$是好是坏

2.1 点估计优良性

一　无偏估计

二　均方误差准则

三　相合估计(一致估计)

四　渐近正态估计

一、无偏性

若 $X_1, X_2, \cdots, X_n$ 为总体 X 的一个样本，

$\theta \in \Theta$是包含在总体 X 的分布中的待估参数，

定义2.1 若估计量 $\hat{\theta} = \theta(X_1, X_2, \cdots, X_n)$的数学期望 $E(\hat{\theta})$存在，且对于任意 $\theta \in \Theta$ 有 $E(\hat{\theta}) = \theta$，则称 $\hat{\theta}$是θ 的无偏估计量.

无偏估计的实际意义：无系统误差.

若 $\lim\limits_{n \to \infty} E\hat{\theta} = \theta$，则称 $\hat{\theta}$是θ的渐近无偏估计.

例3 设总体 X 的k阶矩$\alpha_k = E(X^k)\ (k \geq 1)$存在，又设 $X_1, X_2, \cdots, X_n$ 是 X 的一个样本，试证明不论总体服从什么分布，k阶样本矩 $A_k = \dfrac{1}{n}\sum_{i=1}^{n} X_i^k$ 是 k阶总体矩α_k的无偏估计. 看定理1.2及其推论

证 因为$X_1, X_2, \cdots, X_n$与X同分布，

故有 $E(X_i^k) = E(X^k) = \alpha_k,\quad i = 1, 2, \cdots, n.$

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

定理1.2的推论中

$$ES_n^2 = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2\right) = \frac{n-1}{n}\sigma^2 \neq \sigma^2,$$

所以S_n^2是σ^2的有偏估计.

若以 $\dfrac{n}{n-1}$ 乘S_n^2，　即

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

因为 $\dfrac{n}{n-1}S_n^2 = S_n^{*2} = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}^2)$,

即 S_n^{*2}是σ^2 的无偏估计,故通常取S_n^{*2}作σ^2的估计量.

例4 设总体X的方差$D(X)$存在，且 $D(X) > 0$,(X_1, $X_2, \cdots, X_n$) 为来自总体X的样本，试选择适当的常数C，使得

$$C\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$$

为$D(X)$的无偏估计.

分析 需选择C，使

$$E[C\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2]=D(X)$$

证 $\because$ $E[C\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2]=C\sum_{i=1}^{n-1}E(X_{i+1}-X_i)^2$

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

而$X_1, X_2, \cdots, X_n$相互独立，且与X同分布

$\therefore$ $E(X_i)=E(X),\quad D(X_i)=D(X)\quad (i=1,2,\cdots,n)$

$$D(X_{i+1}-X_i)=D(X_{i+1})+D(X_i)=2D(X)$$

$$E(X_{i+1}-X_i)=E(X_{i+1})-E(X_i)=0$$

$$\therefore \quad E[C\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2]$$

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

$$=C\sum_{i=1}^{n-1}2D(X) \quad =C\cdot 2(n-1)D(X)$$

依题意，要求: $E[C\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2]=D(X)$

即 $\quad C\cdot 2(n-1)D(X)=D(X)$

$\because \quad D(X)>0 \quad \therefore \quad C=\dfrac{1}{2(n-1)}.$

注 一般地，一个参数θ的无偏估计量不唯一.

如：设样本$(X_1, X_2, \cdots, X_n)$来自总体X，$E(X)=\mu$,

则$\overline{X}$是μ的无偏估计. 此外，

$$\sum_{i=1}^{n} C_i X_i \qquad (\sum_{i=1}^{n} C_i = 1)$$

均是μ的无偏估计.

? 对于同一个参数的多个无偏估计量，

如何进一步评价它们的优劣？

二、均方误差准则

1. 均方误差

定义2.2 设θ为一个未知参数，$\hat{\theta}$是θ的一个估计量$\hat{\theta}$的均方误差定义为

$$MSE(\hat{\theta},\theta)=E(\hat{\theta}-\theta)^2$$

利用上述定义，能否找到最优估计呢？答案：否.

这是因为：

$E(\theta^*-\theta)^2$是θ的函数，不是一个确定的数.

由θ的任意性,没法确定一个函数的优劣性。

又因为 $MSE(\hat{\theta},\theta) = E(\hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta)^2$

$$= \underbrace{D(\hat{\theta})}_{\hat{\theta}\text{的方差}} + \underbrace{(E(\hat{\theta}) - \theta)^2}_{\hat{\theta}\text{的偏差的平方和}}$$

当$E\hat{\theta} = \theta$时，$MSE(\hat{\theta},\theta) = D(\hat{\theta})$.

此时可通过比较$D(\hat{\theta})$的大小来判定$\hat{\theta}$的优劣.

2. 有效性

定义2.3 设 $\hat{\theta}_1=\hat{\theta}_1(X_1,X_2,\cdots,X_n)$，$\hat{\theta}_2=\hat{\theta}_2(X_1,X_2,\cdots,X_n)$

均是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.

例5 设$E(X)=\mu, D(X)=\sigma^2>0$存在，$(X_1, X_2)$是来自总体X的样本，问：下列三个对$\mu$的无偏估计量哪一个最有效？$\hat{\mu}_1=\frac{3}{4}X_1+\frac{1}{4}X_2$,

$\hat{\mu}_2=\frac{1}{2}X_1+\frac{1}{2}X_2$, $\hat{\mu}_3=\frac{2}{3}X_1+\frac{1}{3}X_2$.

证 $D(\hat{\mu}_1)=(\frac{9}{16}+\frac{1}{16})\sigma^2=\frac{5}{8}\sigma^2$,

$D(\hat{\mu}_2)=\frac{1}{2}\sigma^2, D(\hat{\mu}_3)=\frac{5}{9}\sigma^2$,

$\because \quad D(\hat{\mu}_2)<D(\hat{\mu}_3)<D(\hat{\mu}_1) \quad \therefore \quad \hat{\mu}_2$最有效.

3. 最小方差无偏估计

定义2.4 如果存在θ的一个无偏估计量$\hat{\theta}_0$,使得对于θ的任一方差存在的无偏估计量$\hat{\theta}$,都有

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

则称$\hat{\theta}_0$是θ的最小方差无偏估计(量),缩写为$MVUE$.

注 最小方差无偏估计是一种最优估计. 基于完备统计量的无偏估计一定是*MVUE*. 后面2.3专门讲述.

三、相合估计（一致估计）

1. 相合估计

定义2.5 若$\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$为参数$\theta$的估计量，若对于任意$\theta\in\Theta$，当$n\to\infty$时，$\hat{\theta}(X_1,X_2,\cdots,X_n)$依概率收敛于$\theta$，则称$\hat{\theta}$为$\theta$的相合估计量.即

$$\lim_{n\to\infty}P\{|\hat{\theta}_n-\theta|<\varepsilon\}=1\text{或}\lim_{n\to\infty}P\{|\hat{\theta}_n-\theta|\geq\varepsilon\}=0.$$

记为 $\hat{\theta}_n\xrightarrow{P}\theta$

注 相合估计也是对估计量的一个基本要求

2. 相和估计的性质

定理2.2 如果$\hat{\theta}_n$是θ的相合估计，$g(\theta)$在$x=\theta$连续，则$g(\hat{\theta}_n)$也是$g(\theta)$的相合估计.

证 由连续性定义可知：对任意$\varepsilon>0$,存在$\delta>0$,当

$|x-\theta|<\delta$时,$|g(x)-g(\theta)|<\varepsilon$.由此可以得到

$$P\{|g(\hat{\theta}_n)-g(\theta)|>\varepsilon\}\le P\{|\hat{\theta}_n-\theta|>\delta\}$$

因此$0\le\lim\limits_{n\to\infty}P\{|g(\hat{\theta}_n)-g(\theta)|>\varepsilon\}\le\lim\limits_{n\to\infty}P\{|\hat{\theta}_n-\theta|>\delta\}=0$

定理得证. 对于多元函数也有类似结论。

例6 试证:样本均值 $\bar{X}$ 是总体均值 μ 的相合估计量，修正样本方差 $S_n^{*2} = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$ 及样本方差 $S_n^2 = \dfrac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$ 都是总体方差 σ^2 的相合估计量.

证 由大数定律知,

$$\forall \varepsilon > 0, \quad \text{有} \quad \lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \varepsilon\right\} = 1,$$

所以 $\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ 是 μ 的相合估计量.

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

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

(A_2是样本二阶原点矩)

由大数定律知,

$$A_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2\xrightarrow{P}E(X^2),$$

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

故 由定理2.2可知 $S_n^2 = A_2 - \bar{X}^2$

$$\xrightarrow{P} E(X^2) - [E(X)]^2 = \sigma^2,$$

所以S_n^2 是 σ^2 的相合估计量.

又 $\lim\limits_{n\to\infty} \dfrac{n}{n-1} = 1,$

所以 $S_n^{*2} = \dfrac{n}{n-1} B_2$ 也是 σ^2 的相合估计量.

由此例可以看出：用定义证明相当麻烦。

3. 相合估计的判别法则

定理2.1　设$\hat{\theta}_n$是θ的一个估计量,若

$$\underbrace{\lim_{n\to\infty} E(\hat{\theta}_n)=\theta,}_{(1)}\quad 且\ \underbrace{\lim_{n\to\infty} D(\hat{\theta}_n)=0}_{(2)}$$

则$\hat{\theta}_n$是θ的相合估计(或一致估计).

证　这是因为

$$0\le P\{|\hat{\theta}_n-\theta|\ge\varepsilon\}\le\frac{1}{\varepsilon^2}E(\hat{\theta}_n-\theta)^2=\frac{1}{\varepsilon^2}E(\hat{\theta}_n-E\hat{\theta}_n+E\hat{\theta}_n-\theta)^2$$

$$=\frac{1}{\varepsilon^2}\left[E(\hat{\theta}_n-E\hat{\theta}_n)^2+2E\left[(\hat{\theta}_n-E\hat{\theta}_n)(E\hat{\theta}_n-\theta)\right]+(E\hat{\theta}_n-\theta)^2\right]$$

$$=\frac{1}{\varepsilon^2}[D\hat{\theta}_n+(E\hat{\theta}_n-\theta)^2]\xrightarrow{n\to\infty}0$$

定理得证

例7 设总体$X \sim N(\mu,\sigma^2)$,$(X_1, X_2,\cdots, X_n)$来自于总体X，证明$\bar{X}$是μ的相合估计，S_n^2是σ^2的相合估计。

证：由于 $E(\bar{X}) = E(X) = \mu$

$$\lim_{n\to\infty} D(\bar{X}) = \lim_{n\to\infty} \frac{D(X)}{n} = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0$$

故,由定理2.1,$\bar{X}$是X的相合估计.

同样 $\lim\limits_{n\to\infty} E(S_n^2) = \lim\limits_{n\to\infty} \frac{n-1}{n}\sigma^2 = \sigma^2,$

$\because \quad \frac{n}{\sigma^2} S_n^2 \sim \chi^2(n-1), \therefore \quad D(\frac{n}{\sigma^2} S_n^2) = 2(n-1),$

从而 $\lim\limits_{n\to\infty} D(S_n^2) = \lim\limits_{n\to\infty} \frac{2(n-1)}{n^2}\sigma^4 = 0,$

故由定理2.1, S_n^2是σ^2的相合估计.

由相合估计的定义，相合估计不唯一。

比较参数θ的两个相合估计$\hat{\theta}_1$和$\hat{\theta}_2$, 在$n \to \infty$时，它们之间有何差异？

一般，这种差异可由估计量的渐近分布的渐近方差来反映，常用的渐近分布是渐近正态分布。

四、渐近正态估计

定义2.6 设$\hat{\theta}_n = \hat{\theta}_n(X_1, X_2, \cdots, X_n)$是$\theta$的估计量，如果存在一列$\sigma_n > 0$,满足$\lim_{n\to\infty}\sqrt{n}\sigma_n = \sigma$,其中$0 < \sigma < \infty$,使得当$n \to \infty$时，有

$$\lim_{n\to\infty} P\left\{\frac{\hat{\theta}_n - \theta}{\sigma_n} \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

记为
$$\frac{\hat{\theta}_n - \theta}{\sigma_n} \xrightarrow{L} N(0,1),$$

则称$\hat{\theta}_n$是θ的渐近正态估计，σ_n^2称为$\hat{\theta}_n$的渐近方差.

对于 $\dfrac{\hat{\theta}_n - \theta}{\sigma_n} \xrightarrow{L} N(0,1)$,

注 1 .只要n充分大，则估计量

$$\hat{\theta}_n \sim AN(\theta, \sigma_n^2)$$ 但没有要求θ是$\hat{\theta}_n$的均值。

2. 对于渐近正态估计，其优越性由σ_n^2大小决定.

$\lim\limits_{n\to\infty} \sqrt{n}\sigma_n = \sigma \Longrightarrow \sigma_n^2$越来越小，与$\dfrac{1}{\sqrt{n}}$同阶。

σ_n^2越小越好。

定理2.3 渐近正态估计一定是相合估计

证 设$\hat{\theta}_n$是θ的渐近正态估计，则

$$P\{|\hat{\theta}_n-\theta|>\varepsilon\}=P\{|\hat{\theta}_n-\theta|/\sigma_n>\varepsilon/\sigma_n\}$$

$$\le P\{|\hat{\theta}_n-\theta|/\sigma_n>K\}$$

由于$\lim\limits_{n\to\infty}\sqrt{n}\sigma_n=\sigma\Rightarrow\forall\varepsilon>0,$

$\exists K>0,n$足够大时，σ_n可以足够小，

则有$\dfrac{\varepsilon}{\sigma_n}>K$

$$\xrightarrow[n\to\infty]{} 2[1-\Phi(K)]$$

由于K的任意性，令$K\to\infty$,则$1-\Phi(K)\to 0$,因而

$$P\{|\hat{\theta}_n-\theta|>\varepsilon\}\to 2[1-\Phi(K)]\to 0,$$

所以$\hat{\theta}_n$是相合估计.

例8(p41例2.6) 设$X_1, X_2, \cdots, X_n$是来自总体$B(1, p)$的一个样本，p的一个估计量是$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，试证其为渐近正态估计.

证 由中心极限定理可知

$$\frac{\bar{X} - p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow{L} N(0,1)$$

因而

$\bar{X}$是p的渐近正态估计，渐近方差为$p(1-p)/n$.

Thank You!