第六章　参 数 估 计

内容提要

（一）点估计

当母体 X 的分布形式已知，但其中含有未知参数 $\theta_1, \theta_2, \cdots \theta_m$ 时，试采用样本 $(X_1, X_2, \cdots X_n)$ 来估计参数 $\theta_1, \theta_2, \cdots \theta_m$ 的问题称为参数的点估计问题，

1. 矩法估计

设总体 ξ 具有已知类型的分布函数 $F(x;\theta_1, \theta_2, \cdots, \theta_k)$，其中 $(\theta_1, \theta_2, \cdots, \theta_k) \in \Theta$ 是未知参数或参数向量。$(X_1, X_2, \cdots, X_n)$ 为取自总体 ξ 的一个样本。假定总体的 k 阶原点矩 $\mu_k = E\xi^k$ 存在，显然 μ_j （只要 $j < k$ ）也都存在，并且它们都是 $\theta_1, \theta_2, \cdots, \theta_k$ 的函数。这时利用“矩法原则”：样本的 j 阶原点矩 $\overline{X^j}$ 等于总体的 j 阶原点矩 $E\xi^j$，从而构造以下的矩方程组

$$\overline{X^j} = E\xi^j, \quad j = 1,2,\cdots,k$$

求解方程组可以求得 $\theta_1, \theta_2, \cdots, \theta_k$ 的一组解：

$$\hat{\theta}_j = \hat{\theta}_j(X_1, X_2, \cdots, X_n), \quad j = 1,2,\cdots,k$$

这里用符号 $\hat{\theta}_j$ 表示 θ_j 的估计量，由于它仅依赖于样本 $(X_1, X_2, \cdots, X_n)$，且不包含未知参数，所以 $\hat{\theta}_j$ 为一个统计量。进一步，若要估计 $\theta_1, \theta_2, \cdots, \theta_k$ 的函数 $\tau = h(\theta_1, \theta_2, \cdots \theta_k)$，则可直接将得到的 $\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_k$ 代入 τ 的函数形式中，得到 τ 的矩法估计：

$$\hat{\tau} = h(\hat{\theta}_1, \hat{\theta}_2, \cdots \hat{\theta}_k)。$$

其中的矩常常是原点矩，而样本原点矩 $\overline{X^k} = \dfrac{1}{n}\sum_{i=1}^{n} X_i^{\,k}$，而母体矩则可用前面几章中所学方法计算。该法的理论基础是辛钦大数定律，在计算时矩法的困难主要来自母体矩的计算。

2. 极大似然法

在试验中如果出现结果A，则可近似认为A出现的概率最大。具体计算分两步：
（1）先构造似然函数L：

设总体$\xi \sim F(x;\theta),\theta \in \Theta$，其中$\theta$是一个未知参数（或几个未知参数组成的向量），$\Theta$是参数$\theta$可能取值的参数空间。又设$(X_1, X_2, \cdots, X_n)$为取自该总体的样本，$(x_1, x_2, \cdots, x_n)$为相应的一组观测值。

当ξ为离散型总体时，定义

$$L(\theta) = P(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i);$$

当ξ为连续型总体时（不妨假定其密度函数为$p(x)$），定义

$$L(\theta) = p(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} p(x_i)$$

称$L(\theta)$为样本的“似然函数”（Likelihood Function）。

（2）可在待估参数的容许范围内求使L达到最大的参数值，它就是未知参数的最大（极大）似然估计值。即寻找统计量$\hat{\theta}$，使其满足

$$L(\hat{\theta}) = \max_{\theta \in \Theta} L(\theta)$$

称满足上式的$\hat{\theta}$为θ的极大似然估计，简记为 MLE（Maximum Likelihood Estimate）
应注意的是本方法所代入的往往是样本观察值，它们是确定性的。

具体求解步骤：

1：　写出似然函数$L(\theta)$。若总体为离散型，似然函数$L(\theta)$就是样本的联合分布列；若总体为连续型，似然函数$L(\theta)$就是样本的联合密度函数。

2：将似然函数取对数，求得对数似然函数$\ln L(\theta)$。

3：$\ln L(\theta)$关于θ求导（若θ为多个参数，则分别关于每个参数求偏导），并令其为 0，得对数似然方程组。

4：解上述似然方程组，求得θ的极大似然估计$\hat{\theta}$。

（二）点估计的优良性判别

1. 无偏：

设$\hat{\theta}(\omega)$是θ的估计，当$E\hat{\theta}(\omega)=\theta$时，称$\hat{\theta}(\omega)$为$\theta$的无偏估计。

2. 有效：

设$\hat{\theta}_1(\omega)$，$\hat{\theta}_2(\omega)$都是θ的无偏估计，则当$D\hat{\theta}_1(\omega)\leqslant D\hat{\theta}_2(\omega)$时，称估计$\hat{\theta}_1(\omega)$比$\hat{\theta}_2(\omega)$更有效。

3. 相合（一致）：

设n为样本容量，$\hat{\theta}(\omega)$是θ的估计，（有时为强调$\hat{\theta}(\omega)$与n的关系，记为$\hat{\theta}_n(\omega)$。则当成立 $\hat{\theta}_n(\omega)\xrightarrow{P}\theta$ （$n\to\infty$）时称$\hat{\theta}_n(\omega)$为θ的相合（或一致）估计，

4. 几个重要结论

（1）一般的母体X，设其期望或方差分别存在，则有结果：

待估参数	估计量	矩法估计	是否无偏	是否相合
EX	$\overline{X}$	是	是	是
DX	S^2	否	是	是
DX	$\frac{n-1}{n}S^2$	是	否	是
$\sqrt{DX}$	S	否	否	是
$\sqrt{DX}$	$\sqrt{\frac{n-1}{n}}S$	是	否	是

（2）正态母体$X\sim N(\mu,\sigma^2)$，则有结果：

① $\overline{X}$，$\frac{n-1}{n}S^2$，$\frac{n-1}{n}S$ 分别是 μ ，σ^2，σ 的最大似然估计；

② S^2，S 都不是 σ^2，σ 的最大似然估计。

（三）区间估计

1．基本思想

从样本出发，找到的区间覆盖待估参数的可能性（概率）尽量大。

2．定义 假设 $X_1, X_2, \cdots, X_n$ 为来自总体 X 的样本，θ 为总体 X 的未知参数，对给定的数 α $(0<\alpha<1)$，若存在统计量 $\underline{\theta}=\underline{\theta}(X_1, X_2, \cdots, X_n), \overline{\theta}=\overline{\theta}(X_1, X_2, \cdots, X_n)$，

使 $P\{\underline{\theta}\le\theta\le\overline{\theta}\}=1-\alpha$，称 $[\underline{\theta}, \overline{\theta}]$ 为 θ 的置信水平（或置信度，或置信概率）为 $1-\alpha$ 的（双侧）置信区间，$\underline{\theta}$ 为置信下限，$\overline{\theta}$ 为置信上限。

3．正态总体参数的置信区间

进行区间估计时必需选择合适的统计量（枢轴量），而正态总体的统计量的分布是已知的。现已将各种情况下正态总体待估参数置信区间的公式(表一)

表一：正态总体参数的区间估计

总体个数	待估参数	其他参数	统计量（枢轴量）及其分布	置信水平 $1-\alpha$ 下待估参数的置信区间
单个总体	μ	σ 已知	$U=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$	$[\overline{x}-\frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}, \overline{x}+\frac{\sigma}{\sqrt{n}}u_{1-\frac{\alpha}{2}}]$
		σ 未知	$T=\frac{\overline{X}-\mu}{s/\sqrt{n}}\sim t(n-1)$	$[\overline{x}-\frac{s}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1), \overline{x}+\frac{s}{\sqrt{n}}t_{1-\frac{\alpha}{2}}(n-1)]$
	σ^2	μ 已知	$\chi^2=\frac{(n-1)s^2+n(\overline{X}-\mu)^2}{\sigma^2}\sim\chi^2(n)$	$[\frac{(n-1)s^2+n(\overline{X}-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}, \frac{(n-1)s^2+n(\overline{X}-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)}]$
		μ 未知	$\chi^2=\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma^2}=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$	$[\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}, \frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}]$
		σ_1，σ_2 已知	$U=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}\sim N(0,1)$	$[\overline{X}-\overline{Y}-u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}, \overline{X}-\overline{Y}+u_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}]$

两个总体	$\mu_1-\mu_2$	σ_1，σ_2 未知，但 $\sigma_1=\sigma_2$	$T=\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{s_w\sqrt{\frac{1}{m}+\frac{1}{n}}}\sim t(m+n-2)$	$[\overline{X}-\overline{Y}-t_{1-\frac{\alpha}{2}}(m+n-2)s_w\sqrt{\frac{1}{m}+\frac{1}{n}},\ \overline{X}-\overline{Y}+t_{1-\frac{\alpha}{2}}(m+n-2)s_w\sqrt{\frac{1}{m}+\frac{1}{n}}]$
	σ_1^2/σ_2^2	μ_1，μ_2 已知	$F=\dfrac{[(m-1)s_x^2/m+(\overline{X}-\mu_1)^2]/\sigma_1^2}{[(n-1)s_y^2/n+(\overline{Y}-\mu_2)^2]/\sigma_2^2}\sim F(m,n)$	$[\dfrac{[(m-1)s_x^2/m+(\overline{X}-\mu_1)^2]}{[(n-1)s_y^2/n+(\overline{Y}-\mu_2)^2]}\cdot\dfrac{1}{F_{1-\frac{\alpha}{2}}(m,n)},\ \dfrac{[(m-1)s_x^2/m+(\overline{X}-\mu_1)^2]}{[(n-1)s_y^2/n+(\overline{Y}-\mu_2)^2]}\cdot\dfrac{1}{F_{\frac{\alpha}{2}}(m,n)}]$
		μ_1，μ_2 未知	$F=\dfrac{s_x^2/\sigma_1^2}{s_y^2/\sigma_2^2}\sim F(m-1,n-1)$	$[\dfrac{s_x^2}{s_y^2}\cdot\dfrac{1}{F_{1-\frac{\alpha}{2}}(m-1,n-1)},\dfrac{s_x^2}{s_y^2}\cdot\dfrac{1}{F_{\frac{\alpha}{2}}(m-1,n-1)}]$

说明：$s^2=\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$，$s_x^2=\dfrac{1}{m-1}\sum_{i=1}^{m}(X_i-\overline{X})^2$，$s_y^2=\dfrac{1}{n-1}\sum_{i=1}^{n}(Y_i-\overline{Y})^2$，$s_w^2=\dfrac{\sum_{i=1}^{m}(X_i-\overline{X})^2+\sum_{i=1}^{n}(Y_i-\overline{Y})^2}{m+n-2}$