第7章 参数估计

二、有效性

对于一个参数，可能有多个无偏估计量，采用哪一个好?

包含两点：①无偏性；

②在所有无偏估计中找一个“好”的估计。

（估计量的波动性越小越好，即方差越小越好）

设$\hat{\theta}_1=\hat{\theta}_1(X_1,X_2,\dots,X_n)$与$\hat{\theta}_2=\hat{\theta}_2(X_1,X_2,\dots,X_n)$都是$\theta$的无偏估计，若$D(\hat{\theta}_1)\le D(\hat{\theta}_2)$，则称$\hat{\theta}_1$优于$\hat{\theta}_2$，或$\hat{\theta}_1$比$\hat{\theta}_2$有效。若对任一 θ的无偏估计 $\hat{\theta}$ 都有$D(\hat{\theta}_1)\le D(\hat{\theta})$，则称$\hat{\theta}_1$是$\hat{\theta}$的最小方差无偏估计（MVUE）。

西安电子科技大学 XIDIAN UNIVERSITY

二、有效性

例1： $\hat{\mu}=\bar{X}$ $\quad \hat{\mu}_1=\sum_{i=1}^{n}W_iX_i$ 哪个有效？

解

$$D(\hat{\mu})=\frac{\sigma^2}{n} \qquad D(\hat{\mu}_1)=D(\sum_{i=1}^{n}W_iX_i)=\sum_{i=1}^{n}W_i^2 \overset{\sigma^2}{D(X_i)}$$

比较 $\sum_{i=1}^{n}W_i^2$ 和 $\frac{1}{n}$ 的大小：

插入数学知识：对于 $x_1,x_2,\cdots,x_n$ ，n个正数，

$$\sqrt{\frac{x_1^2+x_2^2+\cdots+x_n^2}{n}} \geq \frac{x_1+x_2+\cdots+x_n}{n} \geq \sqrt[n]{x_1x_2\cdots x_n} \geq \frac{n}{\frac{1}{x_1}+\cdots+\frac{1}{x_n}}$$

平方平均 算术平均 几何平均 调和平均

等号成立，全相等

$$\sqrt{\frac{\sum W_i^2}{n}} \geq \frac{\sum W_i}{n}=\frac{1}{n} \Rightarrow \frac{\sum W_i^2}{n} \geq \frac{1}{n^2} \Rightarrow \sum W_i^2 \geq \frac{1}{n}$$

等号当W_i全相等时成立

$\therefore \hat{\mu}_1$ 有效性不如 $\hat{\mu}=\bar{X}$

$$f_{X_{(n)}}(x)=nF^{n-1}(x)f(x)$$
$$=n\frac{x^{n-1}}{\theta^{n-1}}\frac{1}{\theta}I_{(0,\theta)}(x)$$
$$=\frac{nx^{n-1}}{\theta^{n}}I_{(0,\theta)}(x)$$

例2： $X\sim U(0,\theta)$ $\quad \hat{\theta}_M=2\overline{X}$ $\quad \hat{\theta}_L=\frac{n+1}{n}X_{(n)}$

解 $D(\hat{\theta}_M)=4\cdot\frac{\sigma^2}{n}=4\cdot\frac{\theta^2}{12n}=\frac{\theta^2}{3n}$ $\qquad D(\hat{\theta}_L)=?$

$$E(\hat{\theta}_L{}^2)=\left(\frac{n+1}{n}\right)^2E(X_{(n)})^2=\left(\frac{n+1}{n}\right)^2\int_{-\infty}^{\infty}x^2f_{X_{(n)}}(x)dx$$

$$=\left(\frac{n+1}{n}\right)^2\int_0^{\theta}x^2\frac{n}{\theta^n}x^{n-1}dx=\left(\frac{n+1}{n}\right)^2\frac{n}{n+2}\theta^2$$

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

比较 $\frac{1}{3n}$ 大还是 $\frac{1}{n(n+2)}$ 大。 $\quad n+2\geq 3$

$\frac{1}{3n}\geq\frac{1}{n(n+2)}$ 等号在 $n=1$ 成立。

$\therefore \hat{\theta}_L$ 优于 $\hat{\theta}_M$ 最大似然估计比矩估计好。

三、一致性（相合性）

设 $\theta(X_1,\ldots,X_n)$ 是参数 θ 的估计量，若当 $n\to\infty$ 时，$\theta(X_1,\ldots,X_n)$ 依概率收敛于 θ，则称 $\theta(X_1,\ldots,X_n)$ 为 θ 的一致估计或相合估计。

无偏性和有效性都是针对固定样本大小 n 而言的。

一致性是考虑样本大小趋于无穷的性质。

例1： $X \sim U(0,\theta)$ $\hat{\theta}_M = 2\overline{X}$ $\hat{\theta}_L = \dfrac{n+1}{n} X_{(n)}$ 是否为 θ 的一致估计？

解

依概率收敛：

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} P\{|\theta - \theta| \ge \varepsilon\} = 0 \qquad \theta \xrightarrow{P} \theta$$

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} P\{|\theta - \theta| \ge \varepsilon\} = P(|\theta - E(\theta)| \ge \varepsilon) \le \frac{D(\theta)}{\varepsilon^2}$$

$$P(|\hat{\theta}_M - E(\hat{\theta}_M)| \ge \varepsilon) \le \frac{\theta^2}{3n\varepsilon^2} \quad \to 0 \qquad n \to \infty$$

$$P(|\hat{\theta}_L - E(\hat{\theta}_L)| \ge \varepsilon) \le \frac{\theta^2}{n(n+2)\varepsilon^2} \to 0 \qquad n \to \infty$$

所以 $\hat{\theta}_M$ ，$\hat{\theta}_L$ 均为 θ 的一致估计量。

西安电子科技大学 XIDIAN UNIVERSITY

例2： $X \sim N(\mu, \sigma^2)$ S^2 是否为 σ^2 的一致估计量？

解 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

$$E\left[\frac{(n-1)S^2}{\sigma^2}\right] = n-1 \quad \rightarrow \quad E(S^2) = \sigma^2$$

$$D\left[\frac{(n-1)S^2}{\sigma^2}\right] = 2(n-1) \quad \rightarrow \quad D(S^2) = \frac{2\sigma^4}{n-1}$$

$$\forall \varepsilon > 0 \qquad P(\left|S^2 - \sigma^2\right| \geq \varepsilon) \leq \frac{2\sigma^4}{(n-1)\varepsilon^2} \rightarrow 0 \qquad n \rightarrow \infty$$

$\therefore S^2 \xrightarrow{P} \sigma^2$ S^2 是 σ^2 的一致估计量。

7.3 区间估计

一、精度和可靠度

例如估计一个人的年龄：

20岁 ——点估计
（19, 21）之间 ——区间估计
（18, 22）之间 ——区间估计

这两个区间的区别是区间长度不一样

精度：区间长度的一半。 区间长度越短，精度越高。

区间越长，精度越差，可靠度越高。

可靠度和精度是一对矛盾体，我们在做区间估计时要考虑两个因素，缺一不可。谁优先？

Neyman提出的准则：先保证可靠度，在此前提下尽可能提高精度。

置信区间（ Neyman，上世纪30年代）：

对该区间能包含未知数 θ 可置信到何种程度。

区间（*a,b*）不敢保证可以将 θ 罩住，但可以以95%的可靠度保证区间（*a,b*）可以将 θ 罩住。

$\alpha = 0.05$ $\quad\quad 1-\alpha = 95\%$ **置信水平**

$$P\left\{\theta \in \left(a(X_1,\ldots,X_n),\ b(X_1,\ldots,X_n)\right)\right\} \geq 1-\alpha$$

（*a,b*）称为置信水平为 $1-\alpha$ 的置信区间。

区间估计就是在给定的置信水平下，去寻找精度最好的区间。

区间估计都是围绕点估计展开的，前几节可给的点估计都是性质优良的点估计，区间估计要把这些点包含在内。

$\hat{\mu} = \overline{X} \qquad \hat{\sigma}^2 = S^2$

$\overline{X}$ 是正态分布，分布函数对称，区间形式可以写为 $\mu \in (\overline{X} - d, \overline{X} + d)$

S^2 分布是 χ^2 分布　　不取负值，不对称，于是不能以 S^2 为中心±d的形式。

$\hat{\sigma}^2 \in (\frac{S^2}{B}, \frac{S^2}{A}) \qquad 0 < A < 1 < B$

二、样本正态总体均值μ和方差σ²的区间估计

总体 $X \sim N(\mu,\sigma^2)$，取一个样本 $(X_1,\ldots,X_n)$， μ 是未知参数， $\hat{\mu} = \overline{X}$

$\mu \in (\overline{X}-d, \overline{X}+d)$ ，求d。保证可靠的前提下，精度越高越好。

可靠度如何求?

$P\{\mu \in (\overline{X}-d, \overline{X}+d)\} \geq 1-\alpha$ 区间把μ罩住的概率超过 $1-\alpha$

概率P是对哪个量求的?

μ是未知常数θ，无随机性，因此概率不是对μ取的而是对 $\overline{X}$ 取的。

95%的置信区间的理解:

μ有95%的可能性落在区间内? 错误

正确：反复抽样多次，每个样本确定一个区间，在这么多区间内有95%的可能性把μ包含住。

$$P(\overline{X}-d\le\mu\le\overline{X}+d) \quad =P(-d\le\overline{X}-\mu\le d) \;=P(\left|\overline{X}-\mu\right|\le d)$$

$$X\sim N(\mu,\frac{\sigma^2}{n}) \qquad =P\left\{\left|\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right|\le\frac{d}{\sigma/\sqrt{n}}\right\}\ge 1-\alpha$$

$$=P\left\{\left|\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right|>\frac{d}{\sigma/\sqrt{n}}\right\}\le\alpha$$

$$\Rightarrow\frac{d}{\sigma/\sqrt{n}}\ge Z_{\frac{\alpha}{2}}\rightarrow d\ge\frac{\sigma}{\sqrt{n}}Z_{\frac{\alpha}{2}}$$

保证可靠度的前提下精度d越小越好，故等号成立时为最优精度。
（以后为了方便，直接按 $P[\theta\in(a,b)]=1-\alpha$ 求置信区间）

$\overline{X}-d$，$\overline{X}+d$都是统计量，不含未知常数，所以对σ^2要分情况讨论：

（1）σ^2已知　$d=\dfrac{\sigma}{\sqrt{n}}Z_{\frac{\alpha}{2}}$　　$\alpha=0.05$　　$Z_{\frac{\alpha}{2}}=Z_{0.025}=1.96$　　5%的误差

$\alpha=0.1$　　$Z_{\frac{\alpha}{2}}=Z_{0.05}=1.645$

μ的置信水平为1-α的置信区间　$(\overline{X}\pm\dfrac{\sigma}{\sqrt{n}}Z_{\frac{\alpha}{2}})$

（2）σ^2未知

σ 未知，不能构造统计量，那么把σ换成样本标准差S。

$$P\left\{\left|\frac{\overline{X}-\mu}{S/\sqrt{n}}\right|\le\frac{d}{S/\sqrt{n}}\right\}=1-\alpha$$ 但是新构造出的统计量不是正态分布 $$\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$$

$$\frac{d}{S/\sqrt{n}}=t_{\frac{\alpha}{2}}(n-1)\rightarrow d=\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)$$ 置信区间 $$(\overline{X}\pm\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1))$$

记住：σ与Z配套，S与t配套。

通过上面例子中的两种情况总结区间估计步骤：

1）找出一个待估参数θ的良好的点估计 θ（多数是通过最大似然估计）。

2）构造一个 $\hat{\theta}(X_1,\ldots,X_n)$ 与θ的函数 $W\left(\hat{\theta}(X_1,\ldots,X_n),\theta\right)$ 其分布与θ无关，不含未知参数。

W称为枢轴量。

可见枢轴量的三个特点：①包含样本；②包含待估参数；
③分布与θ无关，不含未知参数。

（其实就是构造一个已知分布的统计量，使之包含θ 和 θ）

3）根据置信水平，利用 $P(a \le W \le b) = 1-\alpha$ 再由分布函数的上 $\frac{\alpha}{2}$ 分位点求得

置信区间，这样的估计方法叫做枢轴变量法。

μ未知　求 σ^2 的置信区间

1） σ^2 的无偏估计量为S^2。

2） 枢轴量　含σ^2 、S^2　　$\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

3） $$P\left(\chi_{1-\frac{\alpha}{2}}^2(n-1) \le \frac{(n-1)S^2}{\sigma^2} \le \chi_{\frac{\alpha}{2}}^2(n-1)\right) = 1-\alpha$$

$$\therefore P\left(\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}^2(n-1)} \le \sigma^2 \le \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)}\right) = 1-\alpha$$

σ^2 的置信水平为1-α的置信区间为： $\left(\dfrac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}, \dfrac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)}\right)$

三、两样本正态总体均值差的区间估计

两个正态总体 $X\sim N(\mu_1,\sigma_1^{\,2})$，　样本 $\left(X_1,\dots,X_{n_1}\right)$　$\hat{\mu}_1=\overline{X}$，$\hat{\mu}_2=\overline{Y}$

$Y\sim N(\mu_2,\sigma_2^{\,2})$　$\left(Y_1,\dots,Y_{n_2}\right)$

易得 $\widehat{\mu_1-\mu_2}=\overline{X}-\overline{Y}$　估 $\mu_1-\mu_2\in(a,b)$

$$a=a(X_1,\dots,X_{n_1};Y_1,\dots,Y_{n_2})$$

$$b=b(X_1,\dots,X_{n_1};Y_1,\dots,Y_{n_2})$$

与合样本相关

X、Y独立，$\overline{X}-\overline{Y}$是两个正态分布$\overline{X},\overline{Y}$ 的线性组合，还是正态分布。

$$\therefore\ \overline{X}-\overline{Y}\sim N\left(\mu_1-\mu_2,\frac{\sigma_1^{\,2}}{n_1}+\frac{\sigma_2^{\,2}}{n_2}\right)$$

$$\mu_1-\mu_2\in(\overline{X}-\overline{Y}\pm d)$$

$$P\left\{\mu_1-\mu_2\in(\overline{X}-\overline{Y}\pm d)\right\}=1-\alpha\quad\rightarrow P\left(\left|\underbrace{\overline{X}-\overline{Y}}_{\text{看成}\overline{X}}-\underbrace{(\mu_1-\mu_2)}_{\text{看成}\mu}\right|\le d\right)=1-\alpha$$

$$P\left(\left|\overline{X}-\overline{Y}-(\mu_1-\mu_2)\right|\le d)\right)=1-\alpha$$

$$\to P\left(\left|\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}}\right|\le\frac{d}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}})\right)=1-\alpha$$

$$\to\frac{d}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}}=Z_{\frac{\alpha}{2}}$$

两样本是单样本的简单推广

再分 $\sigma_1^{\ 2}$ $\sigma_2^{\ 2}$ 已知和未知的情况：

1） $\sigma_1^{\ 2}$ $\sigma_2^{\ 2}$ 已知 $d=\sqrt{\frac{\sigma_1^{\ 2}}{n_1}+\frac{\sigma_2^{\ 2}}{n_2}}Z_{\frac{\alpha}{2}}$

置信水平为 $1-\alpha$ 的置信区间：$\left(\overline{X}-\overline{Y}\pm\sqrt{\frac{\sigma_1^{\ 2}}{n_1}+\frac{\sigma_2^{\ 2}}{n_2}}Z_{\frac{\alpha}{2}}\right)$

2） $\sigma_1^{\ 2}$ $\sigma_2^{\ 2}$ 未知 只能讨论 $\sigma_1^{\ 2}=\sigma_2^{\ 2}=\sigma^2$ 这种特殊情况。

$$P\left(\left|\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\sigma}\right|\le\frac{d}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\sigma}\right)=1-\alpha$$

$$P\left(\left|\underbrace{\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}S_w}}_{\text{服从什么分布？}}\right| \le \frac{d}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}S_w}\right)=1-\alpha$$

$$S_w^2=\frac{1}{n_1+n_2-2}\left[\sum_{n_1}(X_i-\overline{X})^2+\sum_{n_2}(Y_i-\overline{Y})^2\right]$$

$$=\frac{1}{n_1+n_2-2}\left[(n_1-1)S_1^{\ 2}+(n_2-1)S_2^{\ 2}\right]$$

$$\frac{(n_1-1)S_1^{\ 2}}{\sigma_1^{\ 2}}\sim\chi^2(n_1-1) \qquad \frac{(n_2-1)S_2^{\ 2}}{\sigma_2^{\ 2}}\sim\chi^2(n_2-1)$$

$$\frac{(n_1-1)S_1^{\ 2}}{\sigma_1^{\ 2}}+\frac{(n_2-1)S_2^{\ 2}}{\sigma_2^{\ 2}}\sim\chi^2(n_1+n_2-2)$$

χ^2 分布的可加性

即 $\dfrac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim t(n_1+n_2-2)$ $\quad S_w^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-1}$

$$\Rightarrow d=\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}S_w t_{\frac{\alpha}{2}}(n_1+n_2-2)$$

$n_1+n_2-2=(n_1-1)+(n_2-1)>n_1-1$ 或 n_2-1

与单样本相比，自由度n增加，$t_{\frac{\alpha}{2}}$ 下降，d下降，精度提高。

合样本的好处——提高精度。