第七章 点估计

【内容预览】

- 点估计
  - 具体内容:总体分布未知,寻求适当的统计量估计分布中的未知参数
  - 估计方法
    - 矩估计
      - 思想:样本矩替换总体矩
      - 掌握:一阶矩、二阶矩
    - 最大似然估计
      - 思想:样本似然函数取最大值时对应的未知参数
      - 掌握:最大似然估计步骤
- 估计量的评选标准
  - 无偏性:$E\left(\hat{\theta}\right)=\theta$
  - 有效性:若$\hat{\theta}_1,\hat{\theta}_2$为$\theta$的无偏估计量,且$D\left(\hat{\theta}_1\right)\leqslant D\left(\hat{\theta}_2\right)$,则$\hat{\theta}_1$比$\hat{\theta}_2$有效
  - 一致性(相合性):$\lim\limits_{n\to\infty}P\left\{\left|\hat{\theta}-\theta\right|<\varepsilon\right\}=1$,也可记为$\hat{\theta}\xrightarrow{p}\theta$
  - 均方误差标准:$\hat{\theta}$为θ的估计量,称$M\left(\hat{\theta}\right)=E\left(\hat{\theta}-\theta\right)^2$为$\hat{\theta}$关于$\theta$的均方误差,
    若$\hat{\theta}_1,\hat{\theta}_2$为$\theta$的无偏估计量,且$M\left(\hat{\theta}_1\right)\leqslant M\left(\hat{\theta}_2\right)$,
    则$\hat{\theta}_1$在均方误差下比$\hat{\theta}_2$有效

【知识清单】

7.1、点估计

设总体X的分布函数的形式已知,但它的一个或多个参数未知,利用总体X的一个样本来估计总体未知参数的问题称为参数点估计.

点估计的思路是构造一个适当的统计量$\hat{\theta}(X_1,X_2,\cdots,X_n)$,用它的观测值$\hat{\theta}(x_1,x_2,\cdots,x_n)$作为未知参数$\theta$的近似值.我们称$\hat{\theta}(X_1,X_2,\cdots,X_n)$为$\theta$的**估计量**,称$\hat{\theta}(x_1,x_2,\cdots,x_n)$为$\theta$的**估计值**.

1.方法一:矩估计

连续型随机变量:

总体$X\sim f(x,\theta_1,\theta_2,\cdots,\theta_k)$,但参数$\theta_1,\theta_2,\cdots,\theta_k$未知,需要对参数$\theta_1,\theta_2,\cdots,\theta_k$进行估计.

步骤:对$X\sim f(x,\theta_1,\theta_2,\cdots,\theta_k)$,

1)计算总体矩:$\mu_l=E(X^l)=\int_{-\infty}^{\infty}x^l f(x,\theta_1,\theta_2,\cdots,\theta_k)dx$

2)取样:$X_1,X_2,\cdots,X_n$;

3）计算样本矩：$A_l=\frac{1}{n}\sum_{i=1}^{n}X_i^{\ l}$，根据大数定律有$A_l=\frac{1}{n}\sum_{i=1}^{n}X_i^{\ l}\xrightarrow{P}\frac{1}{n}\sum_{i=1}^{n}E(X_i^{\ l})=E(X_i^{\ l})$；

4）令$A_l=E(X_i^{\ l})$，会得到下面的方程组：

$$\begin{cases}A_1=E(X_i\ )=\mu_1(\theta_1,\theta_2,\cdots,\theta_k)\\A_2=E(X_i^{\ 2})=\mu_2(\theta_1,\theta_2,\cdots,\theta_k)\\\qquad\vdots\\A_k=E(X_i^{\ k})=\mu_k(\theta_1,\theta_2,\cdots,\theta_k)\end{cases};$$

5）解得：$\hat{\theta}=\theta_i(A_1,A_2,\cdots,A_k),i=1,2,\cdots,k$.

以含有两个参数θ_1,θ_2为例，由大数定律

1）总体的一阶矩为EX，二阶矩为$E(X^2)$（若不使用二阶矩，使用$D(X)$也可）；

2）$A_1=\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，$A_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2$；

3）$\begin{cases}A_1=\mu_1\\A_2=\mu_2\end{cases}\Rightarrow\begin{cases}\frac{1}{n}\sum_{i=1}^{n}X_i=EX\\\frac{1}{n}\sum_{i=1}^{n}X_i^{\ 2}=EX^2\left(\text{或}\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=DX\right)\end{cases}$

4）解得θ_1,θ_2的估计$\widehat{\theta_1},\widehat{\theta_2}$.

离散型随机变量：

总体$P\{X=x\}=p(x;\theta_1,\theta_2,\cdots,\theta_k)$，但参数$\theta_1,\theta_2,\cdots,\theta_k$未知，需要对参数$\theta_1,\theta_2,\cdots,\theta_k$进行估计.

步骤：对$P\{X=x\}=p(x;\theta_1,\theta_2,\cdots,\theta_k)$，

1）计算总体矩：$\mu_l=E(X^l)=\sum_{x\in R_X}x^l p(x;\theta_1,\theta_2,\cdots,\theta_k)$，$R_X$为$X$可能的取值范围.

2）~5）步骤与连续型随机变量相同.

☞技巧：矩估计的思想是利用样本矩替换总体矩，从这个角度来理解与记忆矩估计的方法.

2. 方法二：极大似然估计

设连续型随机变量X的概率密度为$f(x;\theta_1,\theta_2,\cdots,\theta_k)$，离散型随机变量$X$的分布律为$P\{X=x\}=p(x;\theta_1,\theta_2,\cdots,\theta_k)$；其中$\theta_i(i=1,2,\cdots,k)$为未知参数，$X_1,X_2,\cdots,X_n$为来自总体$X$的样本，而$x_1,x_2,\cdots,x_n$为对应$X_1,X_2,\cdots,X_n$的样本值.

步骤：

1）对离散型：$L(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_k)=\prod_{i=1}^{n}P\{X=x_i\}$，

对连续型：$L(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_k)=\prod_{i=1}^{n}f(x_i;\theta_1,\theta_2,\cdots,\theta_k)$;

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

3）对k个参数，令$\frac{\partial}{\partial\theta_t}\ln L(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_k)=0,t=1,2,\cdots,k$.

4）解似然方程或方程组，求似然函数可能的最大值点.

注：若方程或方程组有且仅有唯一解，则该解就是参数的最大似然估计（目前考试若有解也是唯一解的情况，但注意，若得到的解有多组，则代入似然函数中，使函数最大的解就是最优解）；若似然方程或似然方程组无解，则参数的最大似然估计一般在参数的边界达到.

7.2、估计量的评选标准

1.无偏性——设$\hat{\theta}$为θ的一个估计量，若$E\hat{\theta}=\theta$，称$\hat{\theta}$为θ的无偏估计量；

2.有效性——设$\hat{\theta}_1,\hat{\theta}_2$都是$\theta$的无偏估计量，若$D\left(\hat{\theta}_1\right)\le D\left(\hat{\theta}_2\right)$，称$\hat{\theta}_1$比$\hat{\theta}_2$有效；

3.一致性——设$\hat{\theta}$为θ的一个估计量，若$\hat{\theta}\xrightarrow{P}\theta$，即若对于任意的$\theta\in\Theta$都满足：对于任意$\varepsilon>0$，有

$$\lim_{n\to\infty}P\left\{\left|\hat{\theta}-\theta\right|<\varepsilon\right\}=1$$

称$\hat{\theta}$为θ的一致估计，也称为相合估计量.

【重要题型】

题型 1：离散型概率分布的估计

例 7-1：假设总体 X 的概率分布为$P\{X=0\}=\theta^2,P\{X=1\}=2\theta(1-\theta)$，$P\{X=2\}=(1-\theta)^2$，其中$\theta\left(0<\theta<\frac{1}{2}\right)$是未知参数，利用总体的如下样本值：2,0,2,2,1,0,2,2，求θ的矩估计值和极大似然估计值.

解：$EX=0\times\theta^2+1\times2\theta(1-\theta)+2(1-\theta)^2=2(1-\theta)$

$\overline{X}=\frac{2+2+2+2+2+1}{8}=\frac{11}{8}$ ，$EX=\overline{X}\Rightarrow\hat{\theta}=\frac{5}{16}$

$L(\theta)=P(X=0)^2P(X=1)P(X=2)^5=2\theta^5(1-\theta)^{11}$

对方程两边取对数：$\ln L(\theta)=\ln2+5\ln\theta+11\ln(1-\theta)$

两边求导并令其导数为 0，$\frac{d\ln\theta}{d\theta}=\frac{5}{\theta}-\frac{11}{1-\theta}=0\Rightarrow\hat{\theta}=\frac{5}{16}$

题型 2：连续型概率分布的估计

连续型概率分布的极大似然估计通常有两种题型：单调型与非单调型. 这里的单调与非单调是指似然函数$L(\theta)$关于未知参数θ的单调性. 非单调型可以直接通过两边取对数，使一阶导数为 0 求解. 对于单调型的题目参数的极大似然估计量通常是位于未知参数范围的边界，首先确定似然函数$L(\theta)$关于未知参数θ的单调性，再来通过θ的取值范围来确定使似然函数$L(\theta)$取到最大值的极大似然估计量$\hat{\theta}$.

非单调型：

例 7-2：设$(X_1,X_2,\cdots,X_n)$为取自总体$N(0,\sigma^2)$的样本，$\overline{X}$为样本均值，S^2为样本方差，则σ^2的极大似然估计是(　　).

A、S^2　　B、$\frac{n-1}{n}S^2$　　C、$\overline{X}^2$　　D、$\frac{1}{n}\sum_{i=1}^{n}X_i^2$

解：$L(\mu,\sigma^2)=\prod_{i=0}^{n}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x_i^2}{2\sigma^2}}=\frac{1}{(\sqrt{2\pi}\sigma)^n}e^{-\frac{1}{2\sigma^2}\sum_{i=0}^{n}x_i^2}$

两边取对数，对σ^2求导，得：$(\ln L)'=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}x_i^2=0$

即极大似然估计值$\widehat{\sigma^2}=\frac{1}{n}\sum_{i=1}^{n}X_i^2$

例 7-3：设总体 X 具有密度函数$f_X(x,\theta)=\frac{\theta^x e^{-\theta}}{x!},x=0,1,2,\cdots$；且$0<\theta<+\infty$；总体 Y 具有密度函数$f_Y(y,\lambda)=\frac{6y}{\lambda^3}(\lambda-y),0<y<\lambda$; $X_1,X_2,\cdots,X_n$及$Y_1,Y_2,\cdots,Y_n$为分别来自总体X和总体Y的样本.

（1）θ的极大似然估计$\hat{\theta}$；　（2）λ的矩估计$\hat{\lambda}$；　（3）$\hat{\lambda}$的方差。

解：（1）$L(\theta)=\prod_{i=1}^{n}f(x)=\prod_{i=1}^{n}e^{-\theta}\frac{\theta^{x_i}}{x_i!}=e^{-n\theta}\theta^{\sum_{i=1}^{n}x_i}\prod_{i=1}^{n}\frac{1}{x_i!}$;

$$\ln L(\theta)=-n\theta+\sum_{i=1}^{n}x_i\ln\theta-\sum_{i=1}^{n}\ln(x_i!);$$

$$\frac{\partial\ln(L(\theta))}{\partial\theta}=0\Rightarrow -n+\frac{\sum_{i=1}^{n}x_i}{\theta}=0\Rightarrow\hat{\theta}=\overline{X}.$$

（2）$EY=\int_{-\infty}^{\infty}yf_Y(y)dy=\int_0^{\lambda}\frac{6y^2}{\lambda^3}(\lambda-y)dy=\frac{\lambda}{2}=\overline{Y}\Rightarrow\hat{\lambda}=2\overline{Y}$

（3）$D(\hat{\lambda})=D(2\overline{Y})=4D(\overline{Y})=4D\left(\frac{1}{n}\sum_{i=1}^{n}Y_i\right)=\frac{4}{n}DY$

$$EY^2=\int_{-\infty}^{\infty}y^2f_Y(y)dy=\int_0^{\lambda}\frac{6y^3}{\lambda^3}(\lambda-y)dy=\frac{3}{10}\lambda^2$$

$DY=EY^2-(EY)^2=\frac{1}{20}\lambda^2$，$D(\hat{\lambda})=\frac{\lambda^2}{5n}$;

单调型：

例 7-4：设总体 X 的概率密度函数为 $f(x,\theta)=\begin{cases}\dfrac{3x^2}{\theta^3}, & 0\le x\le\theta\\ 0, & 其他\end{cases}$

其中 $\theta>0$ 是未知参数，$X_1,X_2,\cdots,X_n$ 来自总体 X 的简单随机样本。求

（1）参数 θ 的矩估计量 $\widehat{\theta}_1$ 和最大似然估计量 $\widehat{\theta}_2$；

（2）判断估计量 $\widehat{\theta}_2$ 是否是参数 θ 的无偏估计量。

解：（1）$E(X)=\int_{-\infty}^{\infty}xf(x)dx=\int_0^{\theta}\frac{3x^3}{\theta^3}dx=\frac{3}{4}\theta$

$\overline{X}=\frac{3}{4}\theta\Rightarrow\widehat{\theta}_1=\frac{4}{3}\overline{X}$，极大似然方法：

$L(\theta)=\prod_{i=1}^{n}f(x_i;\theta)=\dfrac{3^n\prod_{i=1}^{n}x_i^2}{\theta^{3n}}$；可以看出 $L(\theta)$ 关于 θ 单调递减，故 $\widehat{\theta}_2=\max\{X_i\}$

（2）$P\left(\widehat{\theta}_2\le x\right)=P(\max\{X_i\}\le x)=P(X_1\le x)P(X_2\le x)\cdots P(X_n\le x)$

$$=\begin{cases}\left(\int_0^x\frac{3x^2}{\theta^3}dx\right)^n=\dfrac{x^{3n}}{\theta^{3n}} & 0\le x\le\theta\\ 0 & 其他\end{cases}$$

$f_{\widehat{\theta}_2}(x)=\begin{cases}\dfrac{3nx^{3n-1}}{\theta^{3n}} & 0\le x\le\theta\\ 0 & 其他\end{cases}$，$E\left(\widehat{\theta}_2\right)=\int_{-\infty}^{\infty}xf_{\widehat{\theta}_2}(x)dx=\int_0^{\theta}\frac{3nx^{3n}}{\theta^{3n}}dx=\frac{3n}{3n+1}\theta$

所以是有偏的估计量

题型 3：估计量的评选标准

例 7-5：设 X_1,X_2,X_3 是总体的样本，则下列统计量不是总体均值 μ 的无偏估计量的是（　　）.

A、$\frac{1}{3}(X_1+X_2+X_3)$

B、$\frac{1}{3}X_1+\frac{1}{6}X_2+\frac{1}{2}X_3$

C、$\frac{1}{6}X_1+\frac{1}{3}X_2+\frac{1}{2}X_3$

D、$\frac{1}{6}X_1+\frac{1}{6}X_2+\frac{1}{3}X_3$

解：A: $E\left(\frac{1}{3}(X_1+X_2+X_3)\right)=\frac{1}{3}\times3\mu=\mu$ ，B:$E\left(\frac{1}{3}X_1+\frac{1}{6}X_2+\frac{1}{2}X_3\right)=\left(\frac{1}{3}+\frac{1}{6}+\frac{1}{2}\right)\mu=\mu$

C:$E\left(\frac{1}{6}X_1+\frac{1}{3}X_2+\frac{1}{2}X_3\right)=\left(\frac{1}{3}+\frac{1}{6}+\frac{1}{2}\right)\mu=\mu$；

D:$E\left(\frac{1}{6}X_1+\frac{1}{6}X_2+\frac{1}{3}X_3\right)=\left(\frac{1}{3}+\frac{1}{6}+\frac{1}{6}\right)\mu=\frac{2}{3}\mu$

例 7-6：设 $X_1,X_2,\cdots,X_n$ 是来自分布 $f(x;\theta)=\begin{cases}\dfrac{3}{\theta^3}x^2, & 0\le x\le\theta \\ 0, & 其他\end{cases}$ 的简单随机样本，其中 θ 未知，若统计量 $g(X_1,X_2,\cdots,X_n)=c\overline{X}$ 为参数 θ 的无偏估计，则常数 $c=$__________.

解：$E(X)=\int_0^{\theta}\frac{3}{\theta^3}x^3dx=\frac{3}{4}\theta$，$E\left(c\overline{X}\right)=cE(X)=\frac{3c}{4}\theta=\theta\Rightarrow c=\frac{4}{3}$.

例 7-7：设总体 $X\sim N(0,1)$，X_1,X_2 是从总体抽取的一个样本，下面四个无偏估计量中最有效的是(　)

A、$\mu_1=X_1$　　B、$\mu_2=\frac{2}{3}X_1+\frac{1}{3}X_2$

C、$\mu_3=\frac{1}{4}X_1+\frac{3}{4}X_2$　　D、$\mu_4=\frac{1}{2}X_1+\frac{1}{2}X_2$

解：在 $\mu_i(i=1,2,3,4)$ 均为无偏的情况下，$D(\mu_1)=1$；$D(\mu_2)=\frac{4}{9}+\frac{1}{9}=\frac{5}{9}$；

$D(\mu_3)=\frac{1}{16}+\frac{9}{16}=\frac{5}{8}$；$D(\mu_4)=\frac{1}{4}+\frac{1}{4}=\frac{5}{10}$；$D(\mu_1)<D(\mu_3)<D(\mu_2)<D(\mu_4)$；故选项 D 正确.

【精选习题】

基础篇

1.设总体 X 的分布列如下：其中 θ 为未知参数，X 的一组样本观察值为 $X_1=1,X_2=2,X_3=1$，求 θ 的矩估计值和极大似然估计值.

X	1	2	3
P	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

2.设总体 X 的分布列为：

X	0	1	2	3
P	$\frac{\theta(1-\theta)}{2}$	θ^2	$\frac{\theta(1-\theta)}{2}$	$1-\theta$

其中 $\theta(0<\theta<1)$ 未知,现在从总体中抽取样本大小为 10 的简单随机样本，得到样本观测值为：0, 0, 1, 2, 2, 3, 3, 3, 0, 2.分别求参数 θ 的矩估计值与极大似然估计值.

3.为了检验某种自然水消毒设备的效果，现从消毒后的水中随机抽取 50 升，化验每升水中大肠杆菌的个数，设1升水中含有的大肠杆菌个数服从泊松分布，化验结果如下：

大肠杆菌数/升	0	1	2	3	4	5	6
升数	17	20	10	2	1	0	0

试问平均每升水中大肠杆菌的个数为多少时，才能使出现上述情况的概率为最大？

4.设总体X具有概率密度$f(x,\theta)=\begin{cases}\frac{1}{\theta}e^{-\frac{x}{\theta}}, & x>0\\ 0, & x\le 0\end{cases}\ (\theta>0)$

其中θ未知，$X_1,X_2,\cdots,X_n$是一组样本.求θ的极大似然估计$\hat{\theta}$，并证明$\hat{\theta}$是θ的无偏估计.

5.设总体X的分布函数为$F(x,\alpha,\beta)=\begin{cases}1-\left(\frac{\alpha}{x}\right)^{\beta}, & x>\alpha\\ 0, & x\le\alpha\end{cases}$，其中参数$\alpha>0$，$\beta>1$，$X_1,X_2,\cdots,X_n$为来自总体$X$的样本.

（1）当$\alpha=1$时，求未知参数β的矩估计和极大似然估计；

（2）当$\beta=2$时，求未知参数α的极大似然估计并讨论它的无偏性.

6.设总体X的概率密度为$f(x)=\begin{cases}e^{\theta-x}, & x>\theta\\ 0, & x\le\theta\end{cases}$，$\theta$是未知参数，$(X_1,X_2,\cdots,X_n)$为取自总体$X$的样本.（1）求$\theta$的矩估计量$\theta_M$；（2）求$\theta$的极大似然估计量$\theta_L$；　（3）比较$\theta_M$与$\theta_L$的优劣.

7.设总体$X\sim N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_n$和$Y_1,Y_2,\cdots,Y_m$分别是来自总体X和Y的简单随机样本，样本方差分别是$S_X{}^2,S_Y{}^2$，则σ^2的无偏估计是（　）.

A、$S_X{}^2+S_Y{}^2$

B、$(m-1)S_X{}^2+(n-1)S_Y{}^2$

C、$\dfrac{S_X{}^2+S_Y{}^2}{m+n-2}$

D、$\dfrac{(m-1)S_X{}^2+(n-1)S_Y{}^2}{m+n-2}$

8.设$X_1,X_2,\cdots,X_n$是取自总体$P(\lambda)$的一个样本，$\overline{X}$为其样本均值，则λ^2的一个无偏估计为（　）.

A、$\overline{X}^2-\overline{X}$　　B、$\overline{X}^2+\overline{X}$　　C、$\overline{X}^2-\frac{1}{n}\overline{X}$　　D、$\overline{X}^2+\frac{1}{n}\overline{X}$

9. 设样本X_1,X_2,X_3来自总体$U(0,\theta)$，则参数θ估计量$\widehat{\theta_1}=2\overline{X}$，$\widehat{\theta_2}=X_1+X_2$，$\widehat{\theta_3}=\frac{2}{5}X_1+\frac{4}{5}(X_2+X_3)$中，哪一个估计量较优________.

提高篇

10.设总体 X 的分布律为$P(X=x)=(1-p)^{1-x}p,(x=1,2,\cdots)$，其中$p(0<p<1)$为未知参数.$X_1,X_2,\cdots,X_n$为来自总体$X$的一个样本，试求参数$p$的矩估计量和最大似然估计量.

11.设$X_1,X_2,\cdots,X_m$是取自总体$N(1,\sigma^2)$的简单随机样本.$Y_1,Y_2,\cdots,Y_n$是取自总体$N(2,\sigma^2)$的简单随机样本.

（1）求样本$(X_1,X_2,\cdots,X_m,Y_1,Y_2,\cdots,Y_n)$的联合密度函数；

（2）求基于$(X_1,X_2,\cdots,X_m,Y_1,Y_2,\cdots,Y_n)$的参数$\sigma^2$的极大似然估计$\widehat{\sigma^2}$；

（3）问：（2）问中求得的参数σ^2的极大似然估计$\widehat{\sigma^2}$是否为σ^2的无偏估计？请说明理由.

12.设总体X的概率密度为$f(x)=\dfrac{1}{2\theta}e^{-\frac{|x|}{\theta}},-\infty<x<\infty$，其中$\theta>0$是未知参数．$X_1,X_2,\cdots,X_n$为来自总体$X$的一个样本，

试求：（1）θ的最大似然估计量$\hat{\theta}$；（2）证明$\hat{\theta}$是θ的无偏估计量；（3）证明$\hat{\theta}$是θ的一致（相合）估计量.

13.设总体X的概率密度为$f(x;\theta)=\begin{cases}\dfrac{1}{1-\theta}, & \theta\le x\le 1\\ 0, & \text{其它}\end{cases}$

其中θ为未知参数，$X_1,X_2,\cdots,X_n$为来自总体X的简单随机样本。求：

（1）θ的矩估计量$\widehat{\theta}_1$和最大似然估计量$\widehat{\theta}_2$；（2）讨论$\widehat{\theta}_1,\widehat{\theta}_2$无偏性.

14.设$X_1,X_2,\cdots,X_n$是取自总体X的简单随机样本.X 的密度函数为

$$f_X(x,\theta)=\begin{cases}4e^{-4(x-\theta)}, & x\ge\theta\\ 0, & \text{其他}\end{cases}$$

（1）求θ的极大似然估计量$\hat{\theta}$；

（2）问：（1）中求得的参数θ的极大似然估计量$\hat{\theta}$是否为θ的无偏估计？请说明理由；如果不是，修正为无偏估计.