第二章 参数估计

§1 点估计和估计量的求法

理学院

预习提纲

1. 什么叫做参数估计？

2. 矩法是如何进行参数估计的？

3. 什么是最大似然估计法？

- 本次课教学目的：
  - 掌握点估计的几种方法并会进行参数估计
- 重点难点：
  - 利用矩法进行参数估计
  - 最大似然估计法的应用

§1 点估计和估计量的求法

2.1.1 什么叫参数估计？

例如若想了解某学校全体学生的月平均消费金额这一总体的**数量特征**，有哪些方法？

方法1：对全体师生做普查——描述统计，操作难度大

方法2：抽取部分师生做样本，进行抽样调查，根据样本的数量特征推断总体的数量特征——推断统计（**参数估计**+假设检验）

参数估计：用样本的统计量去估计总体的(未知)参数。

如：用样本均值$\overline{X}$估计总体均值μ

用样本方差s^2估计总体方差σ^2

用样本的估计量直接去估计总体参数的值，这一过程称为“点估计”。例如，通过样本$X_1, X_2, \ldots, X_n$构造某个统计量$\hat{\theta} = \hat{\theta}(X_1, X_2, \ldots, X_n)$作为参数$\theta$的估计，称这样的估计为参数的点估计。

1.估计量 用于估计总体参数的随机变量；

- 例如样本均值就是总体均值μ的一个估计量。

2. 估计值 估计参数时计算出来的统计量的具体值；

3. 一般情况下，参数用θ表示，估计量用$\hat{\theta}$表示

4. 点估计的方法
- 矩估计法
- 最大似然估计法

2.1.2 矩法

1. 什么是矩法?

由大数定律知，当总体的k阶原点矩存在时，样本的k阶原点矩依概率收敛于总体的k阶原点矩。

用样本的k阶原点矩作为总体的k阶原点矩的估计量，建立含待估计参数的方程，从而可解出待估计参数。

2. 如何进行矩估计?

当n充分大时,令

$$\frac{1}{n}\sum_{i=1}^{n} X_i^k = EX^k = \mu_k(\theta_1,\theta_2,\cdots,\theta_m) \qquad (k=1,2,\cdots,m)$$

解此m个方程得到 $\theta_1,\theta_2,\cdots,\theta_m$的估计值 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_m$称为矩估计值。

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \qquad B_k = \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^k$$

例1：设总体X的概率密度为

$$f(x;\theta)=\begin{cases}\theta x^{\theta-1}, & 0<x<1\\ 0, & 其它\end{cases}$$

$X_1, X_2, \ldots, X_n$为来自于总体X的样本, $x_1, x_2, \ldots, x_n$为样本值, 求参数θ的矩估计。

解：先求总体矩

$$E(X)=\int_0^1 x\cdot\theta x^{\theta-1}dx=\theta\int_0^1 x^{\theta}dx=\frac{\theta}{\theta+1}x^{\theta+1}\Big|_0^1=\frac{\theta}{\theta+1}$$

解之： $\theta=\dfrac{E(X)}{1-E(X)}$ 令 $A_1=\dfrac{1}{n}\sum_{i=1}^{n}X_i=\overline{X}$

$\hat{\theta}=\dfrac{\overline{X}}{1-\overline{X}}$ 为θ的矩估计量， $\hat{\theta}=\dfrac{\overline{x}}{1-\overline{x}}$ 为θ的矩估计值。

例2：设总体X的概率密度为：

$$f(x,\theta)=\frac{1}{2\theta}e^{-\frac{|x|}{\theta}},\quad -\infty<x<+\infty,\quad \theta>0$$

求θ的矩估计量。

解法1： $$EX=\int_{-\infty}^{+\infty}x\cdot\frac{1}{2\theta}e^{-\frac{|x|}{\theta}}dx=0$$

一阶矩中并不含θ，无法求解，因此需继续求总体的二阶原点矩。

$$EX^2=\int_{-\infty}^{+\infty}x^2\cdot\frac{1}{2\theta}e^{-\frac{|x|}{\theta}}dx=\frac{1}{\theta}\int_{0}^{+\infty}x^2e^{-\frac{x}{\theta}}dx=\theta^2\Gamma(3)=2\theta^2$$

即 $A_2=\frac{1}{n}\sum_{i=1}^{n}X_i^2=2\theta^2$，由此可得 $\hat{\theta}=\sqrt{\frac{1}{2}\cdot\frac{1}{n}\sum_{i=1}^{n}X_i^2}=\sqrt{A_2/2},\quad \theta>0$

$$\Gamma(s)=\int_0^{+\infty}x^{s-1}e^{-x}dx\qquad(s>0)$$

$$\Gamma(n+1)=n!$$

解法2:

$$E|X| = \int_{-\infty}^{+\infty} |x| \cdot \frac{1}{2\theta} e^{-\frac{|x|}{\theta}} dx = \frac{1}{\theta} \int_{0}^{+\infty} |x| e^{-\frac{|x|}{\theta}} dx = \theta\Gamma(2) = \theta$$

即 $\theta = E|X|$ 用 $\frac{1}{n}\sum_{i=1}^{n}|X_i|$ 替换 $E|X|$ 即可得:

θ的另一矩估计量为 $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}|X_i|$

由此可见，总体不一定存在适当阶的矩，且矩估计并不唯一。

$$\Gamma(s) = \int_{0}^{+\infty} x^{s-1} e^{-x} dx \qquad (s > 0)$$

矩法的优缺点

★ 优点：

1、不依赖总体的分布，简便易行；

2、只要n充分大，精确度也能达到较高水平。

缺点：

1、要求总体的某个k阶矩存在；

2、要求未知参数能写为总体的原点矩的函数形式。

2.1.3 最大似然估计法

首先我们来看生活中一个有趣的现象：

某位大学生与一位猎人一起外出打猎，一只野兔从前方窜过．只听到一声枪响，野兔应声倒下．如果要你推测，是谁打中的呢？你会如何思考呢？

一般地，我们都会这样想，只打一枪便打中，猎人由于长期打猎，他命中兔子的概率一般大于这位同学命中的概率．看来这一枪很大可能都是猎人射中的．

如能使事件发生的概率最大化，则事件最有可能发生，这也便是最大似然估计法的基本思想了。

Southwest Petroleum University
西南石油大学

定义1.1：（1）设随机变量X的概率密度函数为f(x,θ)，其中θ为未知参数(f为已知函数)．$x_1,x_2,\cdots,x_n$ 为样本 $X_1,X_2,\cdots,X_n$ 的观察值，则称

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

为X关于样本观察值 $x_1,x_2,\cdots,x_n$ 的似然函数。

（2）若X是离散型随机变量，似然函数定义为

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

定义1.2：如果似然函数 $L(\theta)=L(x_1,x_2,\cdots,x_n;\theta)$ 在 $\theta=\hat{\theta}$ 时达到最大值，则称 $\hat{\theta}$ 是参数θ的最(极)大似然估计(量）。

求解方法

- 为求最大似然估计量,应解以下似然方程组:

$$\frac{\partial}{\partial\theta_i} L(\theta; x_1, x_2, \cdots, x_n) = 0 \qquad i = 1,2,\cdots,k \qquad \theta = (\theta_1, \cdots, \theta_k)$$

- 因函数的对数和函数本身同时取得最大值，为计算方便,常先对L**取对数**再**求导**，改写似然方程组为:

$$\frac{\partial}{\partial\theta_i} \ln L(\theta; x_1, x_2, \cdots, x_n) = 0 \qquad i = 1,2,\cdots,k \qquad \theta = (\theta_1, \cdots, \theta_k)$$

- 求解步骤: (1)构造$L(\theta; x_1, x_2, \cdots, x_n)$;

  (2)计算$\ln L(\theta; x_1, x_2, \cdots, x_n)$;

  (3)求导$\frac{\partial}{\partial\theta_i} \ln L(\theta; x_1, x_2, \cdots, x_n) = 0$;

  (4)解方程组.

离散母体分布情形

例3：设总体X服从参数为λ的泊松分布，即X有分布列：

$$p(k;\lambda)=P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda},k=0,1,2,\cdots$$

λ是未知参数，λ∈(0, +∞)，试求λ的极大似然估计。

解：(1) 样本的似然函数为

$$L(\lambda)=L(x_1,x_2,\cdots,x_n;\lambda)=p(x_1;\lambda)p(x_2;\lambda)\cdots p(x_n;\lambda)$$

$$=\frac{\lambda^{x_1}}{x_1!}e^{-\lambda}\frac{\lambda^{x_2}}{x_2!}e^{-\lambda}\cdots\frac{\lambda^{x_n}}{x_n!}e^{-\lambda}=\frac{\lambda^{\sum_{i=1}^{n}x_i}}{x_1!x_2!\cdots x_n!}e^{-n\lambda}$$

$$x_i\in\{0,1,2,\cdots\},i=1,2,\cdots,n$$

(2) $\ln L(\lambda) = \ln L(x_1, x_2, \cdots, x_n; \lambda)$

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

(3) $\dfrac{\partial}{\partial \lambda} \ln L(x_1, x_2, \cdots, x_n; \lambda) = -n + (\sum_{i=1}^{n} x_i) \dfrac{1}{\lambda}$

令 $\dfrac{\partial \ln L}{\partial \lambda} = 0$

(4) 得 $\lambda = \dfrac{1}{n} \sum_{i=1}^{n} x_i = \overline{x}$

因此 $\hat{\lambda}(x_1, x_2, \cdots, x_n) = \dfrac{1}{n} \sum_{i=1}^{n} x_i$

连续母体分布情形

例4：设总体X 服从参数为λ的指数分布，即有概率密度

$$f(x,\lambda)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases},(\lambda>0)$$

试求λ的极大似然估计.

解：(1) 似然函数为：

$$L=L(x_1,x_2,\cdots,x_n;\lambda)=\lambda^n\prod_{i=1}^{n}e^{-\lambda x_i}=\lambda^n\exp(-\lambda\sum_{i=1}^{n}x_i)$$

(2) 取对数 $\ln L=n\ln\lambda-\lambda\sum_{i=1}^{n}x_i$

(3) 由 $$\frac{d\ln L}{d\lambda}=\frac{d}{d\lambda}(n\ln\lambda-\lambda\sum_{i=1}^{n}x_i)=\frac{n}{\lambda}-\sum_{i=1}^{n}x_i$$

令 $$\frac{d\ln L}{d\lambda}=\frac{n}{\lambda}-\sum_{i=1}^{n}x_i=0$$

(4) $$\Rightarrow\hat{\lambda}=\frac{n}{\sum_{i=1}^{n}x_i}=\frac{1}{\overline{x}}$$

改写为 $$\lambda=\hat{\lambda}=\frac{1}{\overline{x}}$$

所以 $\hat{\lambda}$是λ的最大似然估计。

思考题

1. 设$x_1, x_2, \cdots, x_n$为正态总体 $N(\mu, \sigma^2)$的一个样本值，求：μ和 σ^2的极大似然估计。（例2.1.5）

求解思路：若总体X的概率密度为 $f(x;\theta_1,\theta_2,\cdots,\theta_k)$，其中$\theta_1,\theta_2,\cdots,\theta_k$为未知参数，$x_1, x_2, \cdots, x_n$ 为样本观察值，此时似然函数为：

$$L(x_1,x_2,\cdots x_n;\theta_1,\theta_2,\cdots,\theta_\mathrm{k})=\prod_{i=1}^{n} f(x_i;\theta_1,\theta_2,\cdots,\theta_\mathrm{k})$$

求解方程组 $\dfrac{\partial \ln L(\theta_1,\theta_2,\cdots,\theta_k)}{\partial\theta_i}=0\,, i=1,2,\cdots,k$

即可得到极大似然估计 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$

解： 似然函数为

$$L(x_1,\cdots,x_n;\mu,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\exp[-\frac{1}{2\sigma^2}(x_i-\mu)^2]$$

$$=\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}}\exp[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2]$$

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

解方程组
$$\begin{cases}\dfrac{\partial\ln L}{\partial\mu}=\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)=0\\ \dfrac{\partial\ln L}{\partial\sigma^2}=-\dfrac{n}{2\sigma^2}+\dfrac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2=0\end{cases}$$

思考题答案

得 $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$

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

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$= s^2$$

思考：例2.1.7

小结

- 最大似然估计法的优点：

 1. 充分利用总体分布的函数信息，克服了矩法的某些不足

 2. 得到的估计量的精度一般较高

- 最大似然估计法的缺点：

 要求必须知道总体的分布函数形式

2.1.4 用子样中位数估计正态母体的参数

Southwest Petroleum University 西南石油大学

定理： 设 $X_1, X_2, \cdots, X_n$ 是独立同分布随机变量，且每一随机变量服从正态分布 $N(\mu, \sigma^2)$。若 me 是 $X_1, X_2, \cdots, X_n$ 的中位数，则对任意的 x，有

$$\lim_{n\to\infty} P\left\{\sqrt{\frac{2n}{\pi}}(me-\mu) \le x\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$$

该定理表明：当 n 很大时，$\sqrt{\frac{2n}{\pi}}(me-\mu)$ 近似服从标准正态分布，因而 me 近似服从正态分布 $N(\mu, \frac{\pi}{2n})$。所以，n很大时可取 $\boxed{\hat{\mu} = me}$ 。

Southwest Petroleum University
西南石油大學

Thank you!