probability
probability

第七章

参 数 估 计

probability
probability

§ 7.1 参数的点估计

§ 7.2 估计量的优良性准则

§ 7.3 区间估计

参数估计是对已知分布类型的总体，利用样本对其未知参数作出估计.

本章参数估计有以下内容的介绍:

- 参数估计
  - 点估计
    - 矩估计
    - 极大似然估计
  - 区间估计

§7.1 参数的点估计

一、点估计

设总体X的分布函数$F(x;\theta)$类型已知, θ是未知参数，θ的取值范围Ω称为**参数空间**.

点估计思想 按一定的优化原则建立一个统计量，将其统计值作为参数θ的估计值.

本节介绍矩法估计和极大似然法估计.

定义7.1.1（点估计）设总体X的分布函数$F(x;\theta)$中的参数θ末知, $\theta\in\Omega$. 由样本X_1, X_2, …,X_n 建立统计量$T(X_1,X_2, \dots,X_n)$, 若将其统计值

$$t=T(x_1, x_2, \dots, x_n)$$

作为θ的估计值，称

$$\hat{\theta}=T(X_1, X_2 \cdots, X_n)$$

为θ的点估计量.

注 总体X的分布函数中可有多个不同未知参数.

二、矩估计法

矩是最简单的数字特征.

设总体X的k阶矩$E(X^k)$存在, $X_1,X_2,\ldots,X_n$是总体X的样本，有

$$E(A_k)=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i^k\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^k)=E(X^k)$$

另一方面,根据辛钦大数定律知

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow{P} E(X^k), \quad as\ n \to \infty$$

样本矩在一定程度反映了总体矩的特性.

矩估计法的思想简单直观--替换原则:

* 用样本矩去替换相应的总体矩;

* 用样本矩的函数替换相应的总体矩的同一函数.

定义7.1.2 设总体X的分布函数

$$F(x;\theta_1,\theta_2,\dots,\theta_l)$$

中含有l个未知参数，假定X的l阶原点矩存在，记

$$\gamma_k(\theta_1,\theta_2,\cdots,\theta_l)=E(X^k),\quad (k=1,2,\cdots,l)$$

由方程组

$$\frac{1}{n}\sum_{i=1}^{n}X_i^k=\hat{\gamma}_k(\theta_1,\theta_2,\cdots,\theta_l),\quad k=1,2,\cdots,l.$$

解得 $\hat{\theta}_k=\hat{\theta}_k(X_1,X_2,\cdots,X_n),\quad k=1,2,\cdots,l,$

称$\hat{\theta}_k$为θ_k的**矩法估计量**.

若 $\hat{\theta}$为 θ的矩法估计量，$g(\theta)$是关于θ的连续函数, 称$g(\hat{\theta})$为$g(\theta)$的矩法估计量.

注　意

样本矩是随机变量,而总体矩是数值

例7.1.1

TIPS

指数分布参数矩估计

二、极大似然估计法（M.L.E.)

问题 若随机试验有若干个可能结果: A_1, A_2, … ,A_m，在一次试验中事件A_1出现了，关于A_1的概率你怎样认为?

根据**小概率事件原理**, 应认为A_1发生的可能性大.

TIPS

不合格品率的M.L.E.估计

极大似然估计法基本思想：

按照最大可能性准则进行推断.

定义7.1.3 若总体 X 的概率密度函数为 $f(x,\theta)$ (θ可以是向量), $X_1, X_2, \cdots, X_n$为来自X的一个样本, n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 的联合概率密度函数记为

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

称为参数θ的**似然函数**.

对于离散型样本,其似然函数为其联合分布律.

极大似然估计法

求参数的估计值，使似然函数达到极大值

定义7.1.4 若

$$L(x_1, x_2, \ldots, x_n; \hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_l) = \max_{(\theta_1, \theta_2, \ldots, \theta_l) \in \Omega} L(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_l)$$

称$(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_l)$为$(\theta_1, \theta_2, \ldots, \theta_l)$的**极大似然估计值**.

相应的估计量 $\hat{\theta}_k = \hat{\theta}_k(X_1, X_2, \ldots, X_n), k = 1,2,\cdots,l$

称为参数θ_k的**极大似然估计量**.

注：$\ln x$ 是 x 的严格单增函数，$\ln L$ 与L有相同的极大值点，**一般**只需求$\ln L$ 的极大值点.

定义7.1.5 称

$$\frac{\partial \ln L(\theta_1, \theta_2, \cdots, \theta_l)}{\partial \theta_k} = 0, \quad k = 1,2,\cdots,l$$

为**似然方程组**.

求极大似然估计量的一般步骤:

1. 写出似然函数

$$L(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_l) = \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2, \ldots, \theta_l)$$

2. 对似然函数取对数

$$\ln L = \sum_{i=1}^{n} \ln f(x_i; \theta_1, \theta_2, \ldots, \theta_l)$$

3. 对θ_j $(j=1,\ldots,l)$分别求偏导,建立似然方程(组)

$$\frac{\partial \ln L}{\partial \theta_j} = 0, \qquad (j = 1, 2, \ldots, l)$$

解得 $\hat{\theta}_1,\ldots,\hat{\theta}_l$ 分别为 $\theta_1,\ldots,\theta_l$ 的极大似然估计值.

4. 写出 $\theta_1,\ldots,\theta_l$ 的极大似然估计量.

TIPS

指数分布的点估计

矩估计与似然估计不等的例子

均匀分布的极大似然估计

小　结

1. 矩法估计量与极大似然估计量不一定相同；

2. 用矩法估计参数比较简单,但有信息量损失；

3. 用矩法估计参数时，估计量不唯一；

4.极大似然估计法精度较高,但运算较复杂；

5.极大似然估计法估计量不唯一；

6.不是所有M.L.E都需要建立似然方程求解.

例7.1.1 设总体X的数学期望和方差都存在，但均未知，求期望和方差矩估计

解 设$X_1,X_2,\ldots,X_n$是总体X的样本，由

$$\begin{cases}\gamma_1 = E(X) = \mu \\ \gamma_2 = E(X^2) = D(X) + E^2(X) = \sigma^2 + \gamma_1^2\end{cases}$$

令 $\gamma_1 = A_1, \gamma_2 = A_2$

$$\begin{cases}\mu = \dfrac{1}{n}\sum_{i=1}^{n} X_i \\ \sigma^2 + (\dfrac{1}{n}\sum_{i=1}^{n} X_i)^2 = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2\end{cases} \Rightarrow \begin{cases}\hat{\mu} = \dfrac{1}{n}\sum_{i=1}^{n} X_i \\ \hat{\sigma}^2 = \dfrac{1}{n}\sum_{i=1}^{n} (X_i - \overline{X})^2\end{cases}$$

例7.1.2 设总体X的概率密度为

$$f(x)=\begin{cases}\frac{1}{\theta}e^{-(x-\mu)/\theta}, & x\ge\mu;\\ 0, & x<\mu.\end{cases}$$

其中$\theta>0$, μ与θ是未知参数, $X_1,X_2,\ldots,X_n$是总体X的一组样本, 求μ与θ的矩法估计量.

解 $$E(X)=\int_{\mu}^{+\infty}\frac{x}{\theta}e^{-\frac{x-\mu}{\theta}}dx$$

$$=\int_0^{+\infty}\frac{1}{\theta}(y+\mu)e^{-\frac{y}{\theta}}dy=\theta+\mu,$$

$$E(X^2)=\int_{\mu}^{+\infty}\frac{x^2}{\theta}e^{-\frac{x-\mu}{\theta}}dx$$

$$=\int_{0}^{+\infty}\frac{1}{\theta}(y+\mu)^2e^{-\frac{y}{\theta}}dy$$

$$=2\theta^2+2\mu\theta+\mu^2=\theta^2+(\theta+\mu)^2$$

由方程组

$$\begin{cases}\hat{\gamma}_1=\hat{\theta}+\hat{\mu}=\overline{X},\\ \hat{\gamma}_2=\hat{\theta}^2+(\hat{\theta}+\hat{\mu})^2=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2\end{cases}$$

解得 $\hat{\theta} = \sqrt{M_2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2},$

$$\hat{\mu} = \overline{X} - \sqrt{M_2}.$$

另一种写法:

令 $\begin{cases} \theta + \mu = \overline{X}, \\ \theta^2 + (\theta + \mu)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \end{cases}$

解得相同结果.

#

例7.1.3 不合格品率p 的M.L.E.估计

设 总体X是抽一件产品的不合格品数，记

$p= P\{X=1\}=P\{$产品不合格$\}$

则 X 的分布律可表示为

$$P(x;p)=\begin{cases} p^{x}(1-p)^{1-x}, & x=0,1; \\ 0, & \text{其它}. \end{cases}$$

现得到X的一组样本X_1，X_2，…,X_n的实际观察值为 $x_1, x_2, \ldots, x_n$，则事件

$\{X_1=x_1，X_2=x_2，\ldots, X_n=x_n\}$

出现的可能性应最大, 其概率为

$$L(x_1, x_2, \dots, x_n; p)$$

$$= P\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\}$$

$$= \prod_{i=1}^{n} P\{X_i = x_i\} = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

$$= p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}, \ (x_i = 0,1; 0 < p < 1)$$

应选取使 $L(p)$ 达到最大的值作为参数 p 的估计.

注意到

$$L(\hat{p}) = \underbrace{\max}_{0<p<1} L(p) \Leftrightarrow \ln L(\hat{p}) = \underbrace{\max}_{0<p<1} \ln L(p),$$

$$\ln L(p) = \sum_{i=1}^{n} x_i \ln p + (n - \sum_{i=1}^{n} x_i)\ln(1-p),$$

令

$$\frac{d\ln L(p)}{dp} = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n - \sum_{i=1}^{n} x_i}{1-p} = 0,$$

解得

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x} = \frac{m}{n}.$$ **(频率值)**

\#

例7.1.4 指数分布的点估计

某电子管的使用寿命 X （单位：小时) 服从指数分布

$$X \sim f(x;\theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}}, & x>0\\ 0, & else\end{cases} \quad (\theta>0)$$

今取得一组样本数据如下，问如何估计θ?

16	29	50	68	100	130	140	270	280
340	410	450	520	620	190	210	800	1100

分析 可用两种方法：矩法估计和极大似然估计.

1）矩法估计 设$X_1,X_2,\ldots,X_n$为总体X的一组样本，其样本值为$x_1, x_2, \ldots, x_n$，则

$$\because \quad E(X)=\int_0^{+\infty} x\cdot\frac{1}{\theta}e^{-\frac{x}{\theta}}dx=\theta$$

令 $\overline{X}=\theta$ 则可得θ的矩法估计量为：$\hat{\theta}=\overline{X}$.

代入具体数值可得θ的估计值为：

$$\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{1}{18}\cdot 5723\approx 318(\text{小时}).$$

2）极大似然估计

1. 构造似然函数

当$x_i>0$，$(i=1,2,\ldots,n)$ 时，似然函数为

$$L(x_1,\ldots,x_n;\theta)=\prod_{i=1}^{n}\frac{1}{\theta}e^{-\frac{x_i}{\theta}}=\theta^{-n}\cdot e^{-\frac{1}{\theta}\sum_{i=1}^{n}x_i}$$

2. 取对数

$$\ln L=-n\ln\theta-\frac{1}{\theta}\sum_{i=1}^{n}x_i$$

3. 建立似然方程

$$\frac{d\ln L}{d\theta}=-\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n}x_i=0.$$

4. 求解得M.L.E值 $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x},$

代入具体数值可得θ的估计值为：

$$\frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{18}\cdot 5723 \approx 318(\text{小时}).$$

5. 得M.L.E量:

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}.$$

#

例7.1.5 矩估计与似然估计不等的例子

设总体概率密度为

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

求参数θ的极大似然估计, 并用矩法估计θ.

解 1)极大似然估计法

1. 构造似然函数

设$X_1,X_2,\ldots,X_n$为总体X的一组样本，其样本值为$x_1,x_2,\ldots,x_n$，则

$$L(x_1,\dots,x_n;\theta)=\begin{cases}(\theta+1)^n\prod_{i=1}^{n}x_i^{\theta}, & 0<x_i<1;\\ 0, & \text{其它}\end{cases}$$

2. **取对数：** **当** $0<x_i<1$, $(i=1,2,\dots,n)$ **时**

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

3. **建立似然方程**

$$\frac{d\ln L}{d\theta}=\frac{n}{\theta+1}+\sum_{i=1}^{n}\ln x_i=0,$$

4. 求解得M.L.E. 值为 $\hat{\theta} = \dfrac{-n}{\sum_{i=1}^{n} \ln x_i} - 1,$

5. M.L.E. 量为 $\hat{\theta} = \dfrac{-n}{\sum_{i=1}^{n} \ln X_i} - 1.$

2) 矩估计法

$$\because E(X) = \int_0^1 x(\theta+1)x^{\theta}dx = \frac{x^{\theta+2}}{\theta+2}(\theta+1)\Big|_0^1 = \frac{\theta+1}{\theta+2},$$

令 $\dfrac{\theta+1}{\theta+2}=\overline{X}$，可得 θ 的矩法估计量为

$$\hat{\theta}=\frac{1-2\overline{X}}{\overline{X}-1}=\frac{1}{1-\overline{X}}-2.$$

#

例 7.1.6 均匀分布的极大似然估计

设样本$X_1,X_2,\ldots,X_n$来自在区间$[\theta_1,\theta_2]$上均匀分布的总体X，求θ_1和θ_2的极大似然估计量.

解 设$x_1,x_2,\ldots,x_n$是$X_1,X_2,\ldots,X_n$的样本值，

总体X的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\theta_2-\theta_1}, & \theta_1\le x\le\theta_2\\ 0, & \text{其他}\end{cases}$$

似然函数为

$$L(x_1,\ldots,x_n;\theta_1,\theta_2)=\begin{cases}\dfrac{1}{(\theta_2-\theta_1)^n}, \theta_1\le x_i\le\theta_2, \quad (i=1,2,\cdots,n);\\ 0 \quad ,\text{其他}.\end{cases}$$

注　　意:

该似然函数不能通过求导构造似然方程.

尝试用其他方法求解!

分析 θ_1和θ_2的估计应满足:

1. $\theta_2-\theta_1>0$，且取值尽可能小;

2. $\theta_1\le x_i\le\theta_2$对所有$i$成立.

因 $L(\theta_1,\theta_2)=\begin{cases}\dfrac{1}{(\theta_2-\theta_1)^n}, \theta_1 \le x_1^*, \quad x_n^* \le \theta_2; \\ 0 \quad ,\text{其他}.\end{cases}$

且 $\dfrac{1}{(\theta_2-\theta_1)^n} \le \dfrac{1}{(x_n^*-x_1^*)^n}, \quad (\theta_1,\theta_2)\in\Omega$

故θ_1与θ_2的极大似然估计值分别为

$$\hat{\theta}_1 = x_1^* = \min_{1\le i\le n}\{x_i\},$$

$$\hat{\theta}_2 = x_n^* = \max_{1\le i\le n}\{x_i\}$$

极大似然估计量分别为

$$\hat{\theta}_1 = X_1^* = \min_{1\le i\le n}\{X_i\},$$

$$\hat{\theta}_2 = X_n^* = \max_{1\le i\le n}\{X_i\}.$$

#