概率论与数理统计

最大似然估计

主讲人：曾华琳

信息科学与技术学院

一、最大似然估计原理

1 最大似然法

它是在总体类型已知条件下使用的一种参数估计方法。

Gauss

1821年

德国数学家

高斯提出

Fisher

1922年

英国统计学家

费歇重新发现，

并首先研究一

些性质。

一、最大似然估计原理

最大似然估计原理：

设 $X_1,X_2,\ldots X_n$ 是取自总体X的一个样本，样本的联合密度（连续型）联合分布律（离散型）为 $f(x_1,x_2,\ldots,x_n;\theta)$。

当给定样本 $X_1,X_2,\ldots X_n$ 时，定义**似然函数**为：

$$\boxed{L(\theta)=f(x_1,x_2,\ldots,x_n;\theta)}$$

一、最大似然估计原理

似然函数： $L(\theta)=f(x_1,x_2,\ldots,x_n;\theta)$

$L(\theta)$看作参数θ的函数，它可作为θ将以多大可能产生样本值$x_1,x_2,\ldots,x_n$的一种度量。

最大似然估计法就是用使$L(\theta)$达到最大值的$\hat{\theta}$去估计θ。

$$L(\hat{\theta})=\max_{\theta}L(\theta)$$ 称$\hat{\theta}$为θ的**最大似然估计值**。

而相应的**统计量** $\hat{\theta}(X_1,\ldots,X_n)$ 称为θ的**最大似然估计量**。

一、最大似然估计原理

例1 设 $X_1,X_2,\ldots X_n$ 是取自总体 $X \sim B(1,p)$ 的一个样本，求参数 p 的最大似然估计量。

解：似然函数为：

$$L(p) = f(x_1, x_2, \ldots, x_n; p) \quad X_i \sim \begin{Bmatrix} 0 & 1 \\ 1-p & p \end{Bmatrix}$$

$$= \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \quad = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

一、最大似然估计原理

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

对数似然函数为：

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

一、最大似然估计原理

对 p 求导并令其为 0，

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

得 $\hat{p}=\frac{1}{n}\sum_{i=1}^{n}x_i=\overline{x}$ 即为 p 的最大似然估计值。

从而 p 的最大似然估计量为

$$\hat{p}(X_1,\ldots,X_n)=\frac{1}{n}\sum_{i=1}^{n}X_i=\overline{X}$$

二、最大似然估计量的步骤

求最大似然估计量的步骤：

（一）写出似然函数

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

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

（二）取对数

$$\ln L(\theta)=\sum_{i=1}^{n}\ln p(x_i;\theta) \quad 或 \quad \ln L(\theta)=\sum_{i=1}^{n}\ln f(x_i;\theta);$$

二、最大似然估计量的步骤

(三) 对 θ 求导 $\dfrac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}$， 并令 $\boxed{\dfrac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}=0,}$ 对数似然方程

解方程即得未知参数 θ 的最大似然估计值 $\hat{\theta}$.

最大似然估计法也适用于分布中含有多个未知参数的情况。

此时只需令 $\dfrac{\partial}{\partial\theta_i}\ln L=0,\quad i=1,2,\cdots,k.$ 对数似然方程组

解出由 k 个方程组成的方程组, 即可得各未知参数

$\theta_i\ (i=1,2,\cdots,k)$ 的最大似然估计值 $\hat{\theta}_i$.

三、例题解析

例2 设总体 $X \sim N(\mu, \sigma^2)$，μ，σ^2 未知。$x_1, \ldots, x_n$ 是来自 X 的样本值，试求，μ，σ^2的最大似然估计量。

解：X 的概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

似然函数为 $$L(\mu,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

三、例题解析

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

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

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

令 $$\frac{\partial}{\partial\mu}LnL=\frac{1}{\sigma^2}(\sum_{i=1}^{n}x_i-n\mu)=0$$

$$\frac{\partial}{\partial\sigma^2}LnL=-\frac{n}{2\sigma^2}+\frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2=0$$

三、例题解析

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

$\mu\ ,\sigma^2$ 的**最大似然估计量**为

$$\hat{\mu}=\bar{X},\widehat{\sigma^2}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

三、例题解析

例3 设$X_1,X_2,\ldots X_n$是取自总体 X 的一个样本

$$X \sim f(x)=\begin{cases}\dfrac{1}{\theta}e^{-(x-\mu)/\theta}, & x\ge\mu \\ 0, & \text{其它}\end{cases} \qquad \theta,\mu\text{为未知参数}$$

其中$\theta>0$，求θ,μ 的最大似然估计。

三、例题解析

解：似然函数为

$$L(\theta,\mu)=\begin{cases}\prod_{i=1}^{n}\frac{1}{\theta}e^{-(x_i-\mu)/\theta}, & x_i\geq\mu \\ 0, & \text{其它}\end{cases}\qquad i=1,2,\ldots,n$$

$$=\begin{cases}\frac{1}{\theta^n}e^{-\frac{1}{\theta}\sum_{i=1}^{n}(x_i-\mu)}, & \min x_i\geq\mu \\ 0, & \text{其它}\end{cases}$$

四、例题解析

对数似然函数为 $\ln L(\theta,\mu)=-n\ln\theta-\frac{1}{\theta}\sum_{i=1}^{n}(x_i-\mu)$

对 θ,μ 分别求偏导并令其为 0，

$$\frac{\partial \ln L(\theta,\mu)}{\partial\theta}=-\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n}(x_i-\mu)=0 \qquad (1)$$

$$\frac{\partial \ln L(\theta,\mu)}{\partial\mu}=\frac{n}{\theta}=0 \qquad (2)$$

三、例题解析

由(1)得

$$\theta=\frac{1}{n}\sum_{i=1}^{n}x_i-\mu \qquad L(\theta,\mu)=\begin{cases}\dfrac{1}{\theta^n}e^{-\frac{1}{\theta}\sum_{i=1}^{n}(x_i-\mu)}, & \min x_i\ge\mu\\ 0, & \text{其它}\end{cases}$$

对 $\mu\le\min x_i, L(\theta,\mu)>0$, **且是** μ **的增函数**

μ **取其它值时，** $L(\theta,\mu)=0$

故使 $L(\theta,\mu)$ **达到最大的** μ **即** μ **的MLE是** $\mu^*=\min\limits_{1\le i\le n}x_i$

于是 $\theta^*=\frac{1}{n}\sum_{i=1}^{n}x_i-\mu^*$ **即** θ^*,μ^* **为** θ,μ **的MLE．**

谢 谢 大 家