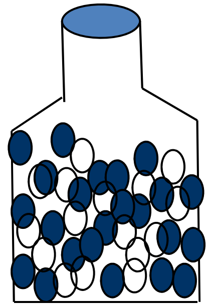
浙江大学
ZHEJIANG UNIVERSITY

第45讲　　极大似然估计

极（最）大似然估计的原理介绍



考察以下例子：

假设在一个罐中放着许多白球和黑球，并假定已经知道两种球的数目之比是1:3，但不知道哪种颜色的球多。如果用放回抽样方法从罐中取5个球，观察结果为：黑、白、黑、黑、黑，估计取到黑球的概率p.

$$\text{解：设}X=\begin{cases}1, & \text{取到黑球},\\ 0, & \text{取到白球}.\end{cases}\text{则}X\sim B(1,p).$$

p为取到黑球的概率,未知, $p=\dfrac{1}{4}$或$\dfrac{3}{4}$.

抽取容量为5的样本$X_1,\ldots,X_5$,观测值为1,0,1,1,1.

当$p=\dfrac{1}{4}$时, 出现本次观察结果的概率为$\left(\dfrac{1}{4}\right)^4\left(\dfrac{3}{4}\right)=\dfrac{3}{1024}$.

当$p=\dfrac{3}{4}$时, 出现本次观察结果的概率为$\left(\dfrac{3}{4}\right)^4\left(\dfrac{1}{4}\right)=\dfrac{81}{1024}$.

由于 $\dfrac{3}{1024}<\dfrac{81}{1024}$, 因此认为$p=\dfrac{3}{4}$比$p=\dfrac{1}{4}$更有可能,

于是 $\hat{p}$ 取为$\dfrac{3}{4}$更合理.

设离散型总体$X \sim p(x;\theta), \theta \in \Theta, \theta$未知. $X_1,\ldots,X_n$为样本,其观察值为$x_1,\ldots,x_n$, 则事件$\{X_1 = x_1,\ldots,X_n = x_n\}$发生的概率为

似然函数: $$L(\theta) = \prod_{i=1}^{n} p(x_i;\theta)$$

该参数对应分布最有可能产生已观测样本值。

极大似然原理: $$L(\hat{\theta}(x_1,\ldots,x_n)) = \max_{\theta \in \Theta} L(\theta).$$

$\hat{\theta}(x_1,\ldots,x_n)$称为$\theta$的**极大似然估计值**, 相应统计量

$\hat{\theta}(X_1,\ldots,X_n)$称为$\theta$的**极大似然估计量(MLE).**

连续型总体X概率密度为$f(x;\theta),\theta\in\Theta,\theta$未知. $X_1,\ldots,X_n$为样本, 则样本在观察值$\left(x_1,x_2,\cdots,x_n\right)$邻域发生的概率

$$\prod_{i=1}^{n}P(x_i<X_i<x_i+\Delta x_i)\approx\prod_{i=1}^{n}f\left(x_i;\theta\right)\Delta x_i,\ \Delta x_i\text{与参数}\theta\text{无关}.$$

因此, 似然函数取为$L\left(\theta\right)=\prod_{i=1}^{n}f\left(x_i;\theta\right)$

极大似然原理: $L(\hat{\theta}(x_1,\ldots,x_n))=\max_{\theta\in\Theta}L(\theta).$

说明: 1.未知参数可能不是一个, 设为$\theta=(\theta_1,\theta_2,\cdots,\theta_k)$;

2.求$L(\theta)$的最大值时, 可转换为求$\ln L(\theta)$的最大值,

$\ln L(\theta)$称为对数似然函数.

利用$\frac{\partial \ln L(\theta)}{\partial \theta_i}\Big|_{\theta=\hat{\theta}}=0, i=1,2,\ldots,k$,解得$\hat{\theta}_i, i=1,2,\ldots,k$.

3. 若$L(\theta)$关于某个θ_i 是单调增(减)函数, 则θ_i的极大似然估计为θ_i的最大(小)值(与样本有关);

4. 若$\hat{\theta}$是θ的极大似然估计, 则$g(\theta)$的极大似然估计为$g(\hat{\theta})$.

例1：设X的概率密度为$f(x;\theta)=\begin{cases}\sqrt{\theta}x^{\sqrt{\theta}-1}, & 0\le x\le 1,\\ 0, & 其他.\end{cases}$

$X_1,\ldots,X_n$是样本, 求θ的极大似然估计量.

若已获得$n=10$的样本值如下,

0.43	0.01	0.30	0.04	0.54
0.14	0.99	0.18	0.98	0.02

求θ的极大似然估计值.

$$\text{解}: L(\theta)=\prod_{i=1}^{n} f(x_i;\theta)=\prod_{i=1}^{n}\sqrt{\theta}x_i^{\sqrt{\theta}-1}=\theta^{\frac{n}{2}}\left(\prod_{i=1}^{n}x_i\right)^{\sqrt{\theta}-1}$$

$$\ln L(\theta)=\frac{n}{2}\ln\theta+\left(\sqrt{\theta}-1\right)\sum_{i=1}^{n}\ln x_i$$

$$\frac{d\ln L(\theta)}{d\theta}=\frac{n}{2}\cdot\frac{1}{\theta}+\frac{1}{2\sqrt{\theta}}\sum_{i=1}^{n}\ln x_i=0$$

$$\Rightarrow \frac{n}{\sqrt{\theta}}=-\sum_{i=1}^{n}\ln x_i \Rightarrow \sqrt{\theta}=-n\Big/\sum_{i=1}^{n}\ln x_i$$

- 参数θ的极大似然估计量为：$\hat{\theta}=\dfrac{n^2}{\left(\sum_{i=1}^{n}\ln X_i\right)^2}.$

- 将上面的样本值代入估计量，得θ的极大似然估计值为：$\hat{\theta}=0.305.$

- 比较第44讲例3得到矩估计量：$\hat{\theta}=\left(\dfrac{\bar{X}}{1-\bar{X}}\right)^2$

例2: 设总体$X \sim N(\mu,\sigma^2), X_1,\ldots,X_n$是样本，$\mu,\sigma^2$均未知. 求$\mu,\sigma^2$的极大似然估计.

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

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

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

$$\frac{\partial}{\partial \sigma^2} \ln L(\mu, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \mu)^2 = 0$$

$$\Rightarrow \begin{cases} \hat{\mu} = \bar{X} \\ \hat{\sigma}^2 = B_2 \end{cases}$$

例3: 设X服从均匀分布$U(a,b)$, a和b未知, 样本$X_1,\cdots,X_n$.

(1) 求a和b的极大似然估计.

(2) 求$E(X)$的极大似然估计.

(3) 若已获得$n=5$的样本值如下,

0.34　0.59　0.16　0.96　0.84

求$a,b,E(X)$的极大似然估计值.

解：(1) 似然函数

$$L(a,b)=\begin{cases}\dfrac{1}{(b-a)^n}, & a\le x_i\le b, i=1,\dots,n.\\ 0, & \text{其他}.\end{cases}$$

关于a单调增,关于b单调减.

并且，在得到样本值$x_1,\dots,x_n$后,只有当

a的取值$\le \min\{x_1,\dots,x_n\}$,

b的取值$\ge \max\{x_1,\dots,x_n\}$,

时才能使似然函数$L(a,b)$不为零.

因此，a达到最大值$\min\{x_1,\ldots,x_n\}$，b达到最小值$\max\{x_1,\ldots,x_n\}$，就能使$L(a,b)$达到最大. 所以

$$\hat{a}=\min\{X_1,\ldots,X_n\}=X_{(1)},$$

$$\hat{b}=\max\{X_1,\ldots,X_n\}=X_{(n)}.$$

比较矩估计量：

$$\hat{a}=\bar{X}-\sqrt{3B_2},$$

$$\hat{b}=\bar{X}+\sqrt{3B_2}.$$

(2)$E(X)=\dfrac{a+b}{2}$是参数a,b的函数，因此$E(X)$极大似然

估计量为 $\widehat{E(X)}=\dfrac{\hat{a}+\hat{b}}{2}=\dfrac{X_{(1)}+X_{(n)}}{2}.$

(3)将样本值分别代入$a,b,E(X)$极大似然估计量，分别得：

$$\hat{a}=0.16,\ \hat{b}=0.96,\ E(\hat{X})=0.56.$$