概率论与数理统计

第三十三讲
极大似然估计

Fisher的极大似然思想

随机试验有多个可能结果，但在一次试验中，有且只有一个结果会出现.

如果在某次试验中，结果 ω 出现了，则认为该结果(事件$\{\omega\}$) 发生的概率P$\{\omega\}$最大.

例如：字“毻”读什么音?

问题 如何将Fisher的极大似然思想应用于参数估计?

假设总体 X 是离散型随机变量，其分布律为：

$$P\{X=a_k\}=p_k(\theta)\ \ (k=1,2,\ldots)$$

其中 $\theta\ (\theta\in\Theta)$ 是未知参数.

第33讲　极大似然估计

$X_1, \quad X_2, \quad \ldots, \quad X_n$ 是来自总体 X 的样本.

$x_1, \quad x_2, \quad \ldots, \quad x_n$ 是样本观测值. ？

即事件$\{X_1 = x_1, \quad X_2 = x_2, \quad \ldots, \quad X_n = x_n\}$发生了.

由 Fisher 的极大似然思想可以得到：

概率 $P\{X_1 = x_1, \quad X_2 = x_2, \quad \ldots, \quad X_n = x_n\}$ 最大.

第33讲　极大似然估计

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

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

$$= P\{X = x_1\}P\{X = x_2\}\ldots P\{X = x_n\} = L(\theta)$$

$$\boxed{P\{X = a_k\} = p_k(\theta) \quad (k = 1, 2, \ldots)}$$

第33讲　极大似然估计

定义1 设 X_1，X_2，…，X_n 是来自总体 X 的样本，

x_1，x_2，…，x_n 是样本观测值.

1) 若 X 是离散型总体，其分布律为：

$$P\{X=a_k\}=p_k(\theta)\quad(k=1,2,\dots)$$

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

2) 若 X 是连续型总体，其密度为 $f(x;\theta)$.

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

称 $L(\theta)$ 为**似然函数**.

例 设 X_1，X_2，…，X_n 是来自总体 $X \sim B(1, p)$ 的样本，x_1，x_2，…，x_n 是样本观测值. 试写出似然函数.

例 设 X_1，X_2，…，X_n 是来自总体 $X \sim N(\mu, \sigma^2)$ 的样本，x_1，x_2，…，x_n 是样本观测值. 试写出似然函数.

定义2 设 $X_1, X_2, \ldots, X_n$ 是来自总体 X 的样本，

$x_1, x_2, \ldots, x_n$ 是样本观测值.

$L(\theta)$ $(\theta \in \Theta)$ 是似然函数. 若存在统计量

$$\hat{\theta} = \hat{\theta}(x_1, x_2, \cdots, x_n)$$

使得：

$$L(\hat{\theta}) = \sup_{\theta \in \Theta} L(\theta)$$

则称 $\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$ 为θ 的**极大似然估计量**，

简记为**MLE** (**M**aximum **L**ikelihood **E**stimate).

极大似然估计求解的一般过程

1) 根据总体分布的表达式，写出似然函数：

$$L(\theta_1,\theta_2,\cdots,\theta_m)\quad(\theta=(\theta_1,\theta_2,\cdots,\theta_m)\in\Theta)$$

2) 因为 $L(\theta_1, \theta_2, \ldots, \theta_m)$ 与 $\ln L(\theta_1, \theta_2, \ldots, \theta_m)$ 有相同的极值点，称 $\ln L(\theta_1, \theta_2, \ldots, \theta_m)$ 为**对数似然函数**.

记为 $l(\theta_1, \theta_2, \ldots, \theta_m)$.

写出对数似然函数 $l(\theta_1, \theta_2, \ldots, \theta_m)$.

3) 求出 $l(\theta_1, \theta_2, \ldots, \theta_m)$的极大值点： $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_m$

即为 $\theta_1, \theta_2, \ldots, \theta_m$ 的**MLE**.

- **关于3)的注释**

若 $l(\theta_1, \theta_2, \ldots, \theta_m)$ 关于 θ_i ($i = 1, 2, \ldots, m$) 可导，则称

$$\begin{cases} \dfrac{\partial l(\theta_1,\theta_2,\cdots,\theta_m)}{\partial \theta_1} = 0 \\ \dfrac{\partial l(\theta_1,\theta_2,\cdots,\theta_m)}{\partial \theta_2} = 0 \\ \quad\vdots \\ \dfrac{\partial l(\theta_1,\theta_2,\cdots,\theta_m)}{\partial \theta_m} = 0 \end{cases}$$

为**对数似然方程组**.

第33讲　极大似然估计

例　设 X_1，X_2，…，X_n 是来自总体 $X\sim B(1,p)$ 的样本，x_1，x_2，…，x_n 是样本观测值. 试求未知参数 p 的极大似然估计.

例　设 X_1，X_2，…，X_n 是来自总体 $X\sim N(\mu，\sigma^2)$ 的样本，x_1，x_2，…，x_n 是样本观测值. 试求未知参数 μ，σ^2 的极大似然估计.

例 设 x_1，x_2，…，x_n 是来自总体 $X \sim f(x;\theta,c)$ 的样本观测值，其中

$$f(x;\theta,c)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{1}{\theta}(x-c)}, & x \geq c,\\ 0, & x < c\end{cases}$$

试求未知参数 θ，c 的极大似然估计.

解 似然函数为

$$L(\theta,c)=\prod_{i=1}^{n} f(x_i;\theta,c)=\prod_{i=1}^{n}\frac{1}{\theta}e^{-\frac{1}{\theta}(x_i-c)}$$

$$=\theta^{-n}e^{-\frac{1}{\theta}n(\overline{x}-c)}, \quad c\le x_{(1)}.$$

对数似然函数为

$$l(\theta,c)=\ln L(\theta,c)=-n\ln\theta-\frac{n}{\theta}(\overline{x}-c), \quad c\le x_{(1)}.$$

对数似然方程组为

$$\begin{cases} \dfrac{\partial l(\theta,c)}{\partial \theta} = -\dfrac{n}{\theta} + \dfrac{n}{\theta^2}(\bar{x} - c) = 0 \\ \dfrac{\partial l(\theta,c)}{\partial c} = \dfrac{n}{\theta} = 0 \end{cases}$$

- **问题** 怎么解上对数似然方程组?

$$\frac{\partial l(\theta,c)}{\partial c} = \frac{n}{\theta} > 0 \Longrightarrow l(\theta,c)$$ 关于 c 严格单调增加

似然函数

$$L(\theta,c)=\theta^{-n}e^{-\frac{1}{\theta}n(\bar{x}-c)},c\le x_{(1)}$$

$\Rightarrow$ c 的极大似然估计为　$\hat{c}=x_{(1)}$

将 c 的极大似然估计代入对数似然方程

$$\frac{\partial l(\theta,c)}{\partial\theta}=-\frac{n}{\theta}+\frac{n}{\theta^2}(\bar{x}-c)=0$$

得到 θ 的极大似然估计为　$\hat{\theta}=\bar{x}-x_{(1)}$

问题 未知参数的极大似然估计唯一吗?

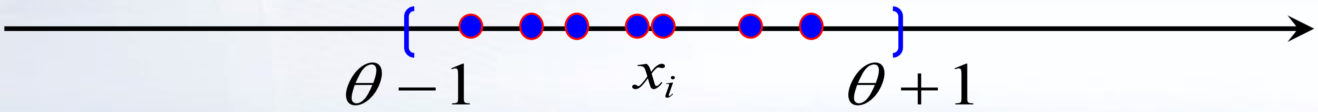
例 设 $X_1, X_2, \ldots, X_n$ 是来自总体 $X \sim U(\theta-1, \theta+1)$ 的样本，$x_1, x_2, \ldots, x_n$ 是样本观测值. 试求 θ 的极大似然估计.

解 因为总体的密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{2}, & \theta-1<x<\theta+1,\\ 0, & others\end{cases}$$

似然函数为

$$L(\theta)=\prod_{i=1}^{n} f(x_i;\theta) \;=2^{-n} \qquad \theta\in ?$$

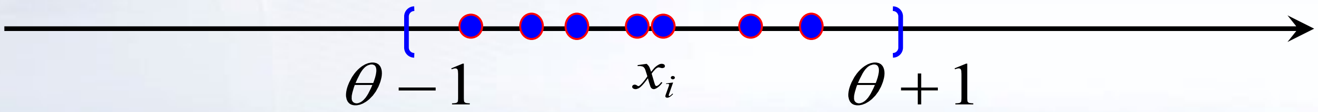


解　因为总体的密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{2}, \theta-1<x<\theta+1, \\ 0, \qquad\qquad\qquad others\end{cases}$$

似然函数为

$$L(\theta)=\prod_{i=1}^{n} f(x_i;\theta) = 2^{-n} \qquad x_{(n)}-1<\theta<x_{(1)}+1$$



即当 $x_{(n)}-1<\theta<x_{(1)}+1$ 时，似然函数 $L(\theta)$ 取得最大值 2^{-n} .

所以区间 $(x_{(n)}-1,\ x_{(1)}+1)$ 内任一点都是 θ 的极大似然估计.

极大似然估计的不变性

设 $\hat{\theta}$ 是 θ 的极大似然估计，$u=u(\theta)$ 是 θ 的函数，

且有单值反函数：

$$\theta=\theta(u)$$

则 $u(\hat{\theta})$ 是 $u(\theta)$ 的极大似然估计.

例　假设袋中有黑球和白球，其中白球所占比例为 $p(0 < p < 1)$未知. 每次有放回地从袋中随机摸取1个球出来观测其颜色后放回，共摸了m 个球，其中白球个数记为X. 共重复了n 次这样的试验，得到样本观察值为 x_1，x_2，…，x_n，试求

(1) p 的极大似然估计；

(2)袋中白球和黑球数之比 R 的极大似然估计.

解 (1) 先求 p 的极大似然估计

因为总体 $X \sim B(m,\ p)$

所以似然函数为

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

$$=p^{n\bar{x}}(1-p)^{n(m-\bar{x})}\prod_{i=1}^{n}\binom{m}{x_i}$$

对数似然函数为

$$l(p)=\ln L(p)=n\bar{x}\ln p+n(m-\bar{x})\ln(1-p)+\ln\prod_{i=1}^{n}\binom{m}{x_i}$$

对数似然函数方程

$$\frac{dl(p)}{dp}=\frac{n\bar{x}}{p}-\frac{n(m-\bar{x})}{1-p}=0$$

解得未知参数 p 的极大似然估计为 $\hat{p}=\dfrac{\bar{x}}{m}$

(2) 求白球和黑球数之比 R 的极大似然估计

因为白球和黑球数之比

$$R = \frac{p}{1-p}$$

所以由极大似然估计的不变性，有

$$\hat{R} = \frac{\hat{p}}{1-\hat{p}} = \frac{\overline{x}}{m-\overline{x}}$$

- **问题**　矩估计有不变性吗?