二．极大似然法（最大似然法）

1.离散型随机变量极大似然估计量的求法

2.连续型随机变量极大似然估计量的求法

例1. 总体 $X \sim B(1,p)$ 分布，抽样得（1,0,1,1,0）。

既然（1,0,1,1,0）在一次抽样下发生，则有理由认为它发生的概率应该最大。

该样本点发生的概率为：

$$P(X_1=1, X_2=0, X_3=1, X_4=1, X_5=0) = p^3(1-p)^2$$

设 $L(p) = p^3(1-p)^2 = p^3 - 2p^4 + p^5$

令 $\dfrac{dL(p)}{dp} = 0$，求 $L(p)$ 的极大值，

得 $\hat{p} = \dfrac{3}{5}$，称其为 p 的极大似然估计。

总体分布列$\boxed{P(X=x_i\ ,\ \theta)=p_i}$, $i=1,2,\ldots,$ $(X_1, X_2\cdots X_n)$是容量为n的样本,$(x_1, x_2\cdots x_n)$为样本观察值,求参数θ的极大似然估计

步骤:(1)似然函数: $$L(\theta)=\prod_{i=1}^{n}P(X=x_i)$$

(先整理称号再取对数)

(2)对数似然函数: $$\ln L(\theta)=\ln\prod_{i=1}^{n}P(X=x_i)$$

$$=\sum_{i=1}^{n}\ln P(X=x_i)$$

(3)求导数求极值: $$\frac{d\ln(L(\theta))}{d\theta}=0$$ 得到θ的极大似然估计。

例1.总体$X \sim B(1, p)$, $P(X = k) = p^k (1-p)^{1-k}, k = 0,1,$

$(X_1, X_2 \cdots X_n)$是样本，$(x_1, x_2 \cdots x_n)$是样本观察值，

求参数p的极大似然估计。

解：似然函数：$L(p) = \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i} = p^{\sum_{i=1}^{n} x_i} (1-p)^{n-\sum_{i=1}^{n} x_i}$

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

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

解得$\hat{p} = \overline{X} = n_A / n$

例2. $X \sim P(\lambda)$，求$\hat{\lambda}_{矩}$，$\hat{\lambda}_{极大}$。

解：$EX=\lambda$，则$\hat{\lambda}_{矩}=\overline{X}$

$$似然函数L(\lambda)=\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}=\lambda^{\sum_{i-1}^{n}x_i}\left(\prod_{i=1}^{n}\frac{1}{x_i!}\right)e^{-n\lambda}$$

$$\ln L(\lambda)=\ln(\lambda^{\sum_{i-1}^{n}x_i}e^{-n\lambda}(x_i!)^{-n})=\sum_{i-1}^{n}x_i\ \ln\lambda-n\lambda+\ln\left(\prod_{i=1}^{n}\frac{1}{x_i!}\right)$$

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

得$\hat{\lambda}_{极大}=\overline{X}$。

例3. $X \sim B(n,p)$ 抽样$(X_1, X_2 \cdots X_m)$, 求 p 的极大似然估计。

解: $$L(p)=\prod_{i=1}^{m} C_n^{x_i} p^{x_i}(1-p)^{n-x_i}=\left(\prod_{i=1}^{m} C_n^{x_i}\right) p^{\sum_{i-1}^{m} x_i}(1-p)^{nm-\sum_{i=1}^{m} x_i}$$

$$\ln L(p)=\left(\sum_{i=1}^{m} x_i\right)\ln p+\left(nm-\sum_{i=1}^{m} x_i\right)\ln(1-p)+\ln\left(\prod_{i=1}^{m} C_n^{x_i}\right)$$

$$\frac{d\ln(L(p))}{dp}=\left(\sum_{i=1}^{m} x_i\right)\frac{1}{\hat{p}}-\left(nm-\sum_{i=1}^{m} x_i\right)\frac{1}{1-\hat{p}}=0$$

$$\hat{p}_{极}=\frac{1}{nm}\sum_{i=1}^{m} x_i=\frac{1}{n}\overline{X}_m \qquad \hat{p}_{矩}=\frac{\overline{X}_m-B_2}{\overline{X}_m}=\frac{1}{n}\overline{X}_m$$

例4. $X \sim G(p)$ 抽样$(X_1, X_2 \cdots X_n)$ 求 p 的极大似然估计。

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

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

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

$$\hat{p}_{极}=\frac{n}{\sum_{i=1}^{n} x_i}=\frac{1}{\overline{X}_n} \qquad \hat{p}_{矩}=\frac{1}{\overline{X}_n}$$

例5.$X \sim \begin{pmatrix} 1 & 2 & 3 \\ \theta^2 & 2\theta(1-\theta) & (1-\theta)^2 \end{pmatrix}$ 抽样(1,2,1)，求 θ 的极大似然估计。

解：$L(\theta)=P(X_1=1，X_2=2，X_3=1)$

$$=P(X_1=1)P(X_2=2)P(X_3=1)$$

$$=\theta^2 \qquad 2\theta(1-\theta) \qquad \theta^2$$

$$=2\theta^5-2\theta^6$$

令 $\dfrac{d\ln(L(\theta))}{d\theta}=10\hat{\theta}^4-12\hat{\theta}^5=0,$ 得 $\hat{\theta}=\dfrac{5}{6}$。

2. 连续型： 总体 X 的分布密度 $\boxed{f(x,\theta)}$ (θ为未知参数), $(X_1, X_2 \cdots X_n)$ 为总体的样本, $(x_1, x_2 \cdots x_n)$ 为样本观察值。**既然样本点 $(x_1, x_2 \cdots x_n)$ 在一次抽样下发生，则有理由认为它发生的可能性（密度）应该最大。**

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

(先整理称号再取对数)

(2) 对数似然函数: $\ln L(\theta) = \ln \prod_{i=1}^{n} f(x_i, \theta)$

(3) 求极值： 令 $\dfrac{d \ln(L(\theta))}{d\theta} = 0$ 得到θ的极大似然估计。

(多个参数求偏导)

例1. $X \sim e(\lambda)$，求λ的矩估计和极大似然估计。

解:(1)矩估计：$EX = 1/\lambda$，令$\overline{X}_n = EX$，得$\hat{\lambda}_{矩} = 1/\overline{X}_n$

(2)极大似然估计：$L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$

$\ln L(\lambda) = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i$；令 $\dfrac{d \ln L(\lambda)}{d\lambda} = \dfrac{n}{\hat{\lambda}} - \sum_{i=1}^{n} x_i = 0$

$$\hat{\lambda}_{极大} = 1/\overline{X}_n$$

若$\hat{\theta}$是θ的极大似然估计量则$g(\hat{\theta})$是$g(\theta)$的估计量。

例2. $X \sim N\left(\mu, \sigma^2\right)$ 求参数μ，σ^2的极大似然估计。

$\left(\hat{\mu}_{矩} = \overline{X}_n,\ \ \sigma^2_{矩} = B_2\right)$

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

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

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

$$\begin{cases} \dfrac{\partial\left(\ln L\left(\mu, \sigma^2\right)\right)}{\partial\mu} = \dfrac{2}{2\sigma^2}\sum_{i=1}^{n}\left(x_i - \hat{\mu}\right) \Rightarrow \sum_{i=1}^{n} x_i - n\hat{\mu} = 0 \\ \dfrac{\partial\left(\ln L\left(\mu, \sigma^2\right)\right)}{\partial\sigma^2} = -\dfrac{n}{2\hat{\sigma}^2} + \dfrac{1}{2\hat{\sigma}^4}\sum_{i=1}^{n}\left(x_i-\mu\right)^2 = 0 \end{cases}$$

得$\begin{cases} \hat{\mu}_{极大} = \overline{X} \\ \hat{\sigma}^2_{极大} = B_2 \end{cases}$

$$\hat{\sigma}_{极大} = \sqrt{\hat{\sigma}^2} = \sqrt{B_2}$$

例3. $f(x,\theta)=\begin{cases}(\theta+1)x^{\theta}, & 0<x<1\\ 0, & \text{其他}\end{cases}$，求$\theta$的极大似然估计。

解：$L(\theta)=\prod_{i=1}^{n}(\theta+1)x_i^{\theta}=(\theta+1)^n\left(\prod_{i=1}^{n}x_i\right)^{\theta}$

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

$$\frac{d\ln L(\theta)}{d\theta}=\frac{n}{\hat{\theta}+1}+\sum_{i=1}^{n}\ln x_i=0 \quad \text{得}\ \hat{\theta}_{\text{极大}}=-1-\frac{n}{\sum_{i=1}^{n}\ln x_i}$$

例5. $X \sim U(a,b)$ $(X_1, X_2 \cdots X_n)$是样本，$(x_1, x_2 \cdots x_n)$是观察值，

求参数a,b的极大似然估计。

解：$L(a,b)=\prod_{i=1}^{n}\frac{1}{b-a}=(b-a)^{-n}$，$f(x,\theta)=\begin{cases}\frac{1}{b-a} & \boxed{a<x<b} \\ 0 & 其他\end{cases}$

令$\frac{\partial L(a,b)}{\partial a}=0$ $\boxed{\text{无解或参数消失}}$

我们的目的不是取对数也不是求导，目的是样本点发生概率最大。

即似然函数$L(a,b)=(b-a)^{-n}$取最大值。

$b-a$越小 $\Rightarrow$ $L(a,b)$越大，但$a<(X_1, X_2 \cdots X_n)<b$，

$a<X_{(1)}\le X_{(2)}\le\cdots X_{(n)}<b \Rightarrow \hat{a}_{极大}=X_{(1)}$，$\hat{b}_{极大}=X_{(n)}$

例6.某种电子元件的使用寿命 X 的密度函数为:

$$f(x,\theta)=\begin{cases}2e^{-2(x-\theta)}, & \boxed{x>\theta} \\ 0, & x\le\theta\end{cases}$$

求 θ 的极大似然估计。

解:

$$L(\theta)=\prod_{i=1}^{n}2e^{-2(x_i-\theta)}=2^n e^{-2\sum_{i=1}^{n}(x_i-\theta)}$$

$$\ln L(\theta)=\ln 2^n-2\sum_{i=1}^{n}(x_i-\theta)\text{；}\quad \frac{d\ln L(\theta)}{d\theta}$$

参数 θ 消失

θ 越大 $L(\theta)$ 越大，但 $x>\theta$

$$\hat{\theta}_{极大}=X_{(1)}$$

矩法定义： 构造前t个 A_k　求出前t个 EX^k　令 $A_k = EX^k$ 。

极大似然法：（1）似然函数 $L(\theta)=\prod_{i=1}^{n} P(X = x_i)$ 或 $f(x_i,\theta)$

（2）对数似然函数　$\ln L(\theta)$

（3）求极大值　令 $\frac{d\ln(L(\theta))}{d\theta}=0$

若连续型试验中(3)式无解，则直接求$L(\theta)$的极大值.