Maximum Likelihood Method

Definitions

Let $X_1, \dots, X_n$ r.s with PDF $f(x\,;\theta)$. Then the function
$L(\theta\,; x) = L(\theta) = \prod_{i=1}^{n} f(x_i\,;\theta) = f(\underline{x}\,;\theta) = f(x_1,\dots,x_n\,;\theta)$
is considered a function of the parameter θ and is called
Likelihood function.

Let $L(\underline{\theta}) = L(\underline{\theta}, \underline{x})$ the likelihood function of a r.s. $X_1, \dots, X_n$. The estimator $\hat{\underline{\theta}}$ of $\underline{\theta} = (\theta_1, \dots, \theta_n)$ is called Maximum Likelihood Estimator (MLE) of $\underline{\theta}$ if:
$L(\hat{\underline{\theta}}) = L(\hat{\underline{\theta}}, \underline{x}) = \max\limits_{\underline{\theta}\in\Theta} L(\underline{\theta}\,;\underline{x})$ or $\hat{\underline{\theta}} = \arg\max\limits_{\theta\in\Theta} L(\underline{\theta})$.

Notes

- For convenience, instead of maximazing $L(\underline{\theta})$ we maximize $\ell(\theta) = \log(L(\theta))$, which is called Log Likelihood.

- When we maximize $L(\theta)$ we might have 1, more than 1 or no maximum points

Method of finding MLE (for a single parameter)

→ For a r.s. $X_1, \dots, X_n$ with PDF $f(x;\theta)$ we calculate
$L(\theta) = \prod_{i=1}^{n} f(x_i\,;\theta)$

→ We calculate $\ell(\theta) = \log L(\theta)$

→ We maximaze $\ell(\theta)$: ⤷ $\dfrac{\partial \ell(\theta)}{\partial\theta} = 0 \Rightarrow$ extremum

⤷ $\left.\dfrac{\partial^2 \ell(\theta)}{\partial\theta^2}\right|_{\theta=\hat{\theta}} < 0 \Rightarrow$ maximum

Example

Let $X_1, \dots, X_v$ r.s. of Poisson(λ). Find $\hat{\lambda}$ MLE.

Solution

$$L(\lambda) = \prod_{i=1}^{v} f(x_i\,;\,\lambda) = \prod_{i=1}^{v} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-v\lambda} \cdot \frac{\lambda^{\sum x_i}}{\prod_{i=1}^{v} x_i!} \Rightarrow$$

$$\Rightarrow \ell(\lambda) = \log L(\lambda) = \log\left(e^{-v\lambda} \frac{\lambda^{\sum x_i}}{\prod x_i!}\right) = -v\lambda + \log \lambda^{\sum x_i} - \log \prod_{i=1}^{v} x_i! \Rightarrow$$

$$\overset{\partial/\partial\lambda}{\Rightarrow} \frac{\partial \ell(\lambda)}{\partial \lambda} = -v + \frac{\sum_{i=1}^{v} x_i}{\lambda}$$

and $\frac{\partial \ell(\lambda)}{\partial \lambda} = 0 \Rightarrow -v + \frac{\sum x_i}{\lambda} = 0 \Rightarrow \hat{\lambda} = \frac{\sum x_i}{v} = \bar{X}$

Also $\left.\frac{\partial^2 \ell(\lambda)}{\partial \lambda^2}\right|_{\lambda=\hat{\lambda}} = -\left.\frac{\sum x_i}{\lambda^2}\right|_{\lambda=\hat{\lambda}} < 0$ so $\hat{\lambda} = \bar{X}$ is MLE.

Example

Let $X_1, \dots, X_v$ r.s. of Geom(p) with PDF $f(x;p) = p(1-p)^x$, $x = 0, 1, \dots$, $p \in (0,1)$. Find MLE.

Solution

$$L(p) = \prod_{i=1}^{v} f(x_i\,;\,p) = \prod_{i=1}^{v} p(1-p)^{x_i} = p^v (1-p)^{\sum x_i} \Rightarrow$$

$$\Rightarrow \ell(p) = \log L(p) = \log\left(p^v (1-p)^{\sum x_i}\right) = v \log p + \sum_{i=1}^{v} x_i \cdot \log(1-p) \Rightarrow$$

$$\Rightarrow \frac{\partial \ell(p)}{\partial p} = \frac{v}{p} - \frac{\sum x_i}{1-p} \text{ and } \frac{\partial \ell(p)}{\partial p} = 0 \Rightarrow -p \sum x_i + (1-p)v = 0 \Rightarrow$$

$$\Rightarrow \hat{p} = \frac{v}{\sum x_i + v} \left(= \frac{1}{\bar{x} + 1}\right)$$

Also $\left.\frac{\partial^2 \ell(p)}{\partial p^2}\right|_{p=\hat{p}} = -\frac{v}{p^2} - \left.\frac{\sum x_i}{(1-p)^2}\right|_{p=\hat{p}} < 0$ so $\hat{p}$ is MLE.

Note

If $L(\underline{\theta})$ has s unknown parameters, $\underline{\theta} = (\theta_1, \dots, \theta_s)$, then the MLE is the vector $(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_s)$ that maximizes $L(\underline{\theta}) = L(\theta_1, \dots, \theta_s)$. The method then goes as follows:

- → Calculate $L(\underline{\theta})$
- → Calculate $\ell(\underline{\theta}) = \log L(\underline{\theta})$

→ We set $\frac{\partial \ell(\underline{\theta})}{\partial \theta_i}\Big|_{\hat{\underline{\theta}}}=0 \quad \forall i=1,\dots,s$ and find $\hat{\underline{\theta}}$, then check if $\frac{\partial^2 \ell(\underline{\theta})}{\partial \theta_i^{2}}\Big|_{\hat{\underline{\theta}}}=0 \quad \forall i=1,\dots,s$ and $\frac{\partial \ell^2(\underline{\theta})}{\partial\theta_i\partial\theta_j}\Big|_{\underline{\theta}=\hat{\underline{\theta}}}=0$

Example

Let $X_1,\dots,X_v$ r.s. of $N(\theta_1,\theta_2)$. Find MLE of (θ_1,θ_2)

$(\Theta=\{(\theta_1,\theta_2) \mid -\infty<\theta_1<+\infty,\ 0<\theta_2<+\infty\}$

Solution

$$L(\theta_1,\theta_2)=\prod_{i=1}^{v} f(x_i;\theta_1,\theta_2)=\prod_{i=1}^{v}\frac{1}{\sqrt{2\pi\theta_2}}\cdot\exp\left\{\frac{1}{2\theta_2}\cdot(x_i-\theta_1)^2\right\}=$$
$$=(2\pi\theta_2)^{-v/2}\cdot\exp\left\{\frac{1}{2\theta_2}\cdot\sum_{i=1}^{v}(x_i-\theta_1)^2\right\}$$
$$\ell(\theta_1,\theta_2)=\log(L(\theta_1,\theta_2))=-\frac{v}{2}\log(2\pi\theta_2)-\frac{1}{2\theta_2}\cdot\sum_{i=1}^{v}(x_i-\theta_1)^2$$

$$\frac{\partial\ell(\theta_1,\theta_2)}{\partial\theta_1}=0\Rightarrow\frac{1}{2\theta_2}\sum_{i=1}^{v}2(x_i-\theta_1)=0\Rightarrow\frac{1}{\theta_2}\cdot\sum_{i=1}^{v}(x_i-\theta_1)=0$$
$$\Rightarrow\sum_{i=1}^{v}(X_i-\theta_1)=0\Rightarrow -v\theta_1+\sum_{i=1}^{v}X_i=0\Rightarrow\hat{\theta}_1=\frac{\sum X_i}{v}=\bar{X}$$

$$\frac{\partial\ell(\theta_1,\theta_2)}{\partial\ell_2}=0\Rightarrow-\frac{v}{2\theta_2}+\frac{1}{2\theta_2^2}\cdot\log\sum_{i=1}^{v}(x_i-\theta_1)^2=0\Rightarrow$$
$$\Rightarrow -v\theta_2+\sum_{i=1}^{v}(X_i-\bar{X})^2=0\Rightarrow\hat{\theta}_2=\frac{\sum(X_i-\bar{X})^2}{v}.$$
$$\left(E(\hat{\theta}_2)=\frac{v-1}{v}\theta_2\xrightarrow{v\to\infty}\theta_2\right)$$

$$\frac{\partial^2\ell(\theta_1,\theta_2)}{\partial\theta_1^2}\Big|_{\underline{\theta}=\hat{\underline{\theta}}}=\frac{v}{\theta_2}\Big|_{\underline{\theta}=\hat{\underline{\theta}}}<0$$

$$\frac{\partial^2\ell(\theta_1,\theta_2)}{\partial\theta_1^2}\Big|_{\underline{\theta}=\hat{\underline{\theta}}}=\dots=\frac{-\sum_{i=1}^{v}(X_i-\bar{X})^2}{2\theta_2^2}<0$$

$$\frac{\partial^2\ell(\theta_1,\ell_2)}{\partial\theta_1\partial\theta_2}\Big|_{\underline{\theta}=\hat{\underline{\theta}}}=\dots=0.$$ Finally, $\hat{\underline{\theta}}$ is MLE.

Theorem

If $X_1,\dots,X_v$ r.s. with PDF $f(x;\theta)$ (single parameter). and $\hat{\theta}$ is a MLE of θ and $\delta=\delta(\underline{x})$ an efficient estimator of θ, then $\hat{\theta} = \delta$.

Proof

Since δ is efficient we have: $\sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log f(x_i ; \theta) = K(\theta)[\delta(\underline{x}) - \theta]$

$L(\theta) = \prod_{i=1}^{n} f(x_i ; \theta)$ and $\ell(\theta) = \log L(\theta) = \sum_{i=1}^{n} \log(f(x_i ; \theta))$

so $\frac{\partial \ell(\theta)}{\partial\theta} = \frac{\partial}{\partial\theta}\left(\sum_{i=1}^{n} \log f(x_i ; \theta)\right) = \sum_{i=1}^{n} \frac{\partial}{\partial\theta} \log f(x_i ; \theta) = K(\theta) \cdot [\delta(x) - \theta]$

and so $\frac{\partial \ell(\theta)}{\partial\theta} = 0 \Rightarrow K(\theta)[\delta(x) - \theta] = 0 \Rightarrow \hat{\theta} = \delta(\underline{X})$

Also $\frac{\partial^2 \ell(\theta)}{\partial\theta^2} = -K(\theta) < 0$ (needs explanation, next lecture)

$$y_?(t) = c_1 y_1(t) + c_2 y_2(t)$$

$$\underset{c_1}{\underline{2}}\, e^t + \underset{c_2}{\underline{5}} \log\cos t$$

$$y(t) = c_1(t)\, e^t + c_2(t) \log\cos(t)$$

$$y(t) = \underbrace{\left(-\int_{t_0}^{t} (\cdots) + c_1\right)}_{c_1(t)} +$$

$$c_1'(t) = e^t \Longrightarrow c_1(t) = e^t + c_1$$

$$y_0(t) = c_1 e^t + c_2 e^{-2t}$$

$$y(t) = c_1(t) e^t + c_2(t) \cdot e^{-2t}$$

$$c_1(t) = \frac{1}{3}\left[\ln(e^{-t}+1) - e^{-t}\right] + C_1$$
$$c_2(t) = \frac{1}{3}\left[\ln(e^{t}+1) - e^{t}\right] + C_2$$

$$y(t) = \left\{\frac{1}{3}\left[\ln(e^{-t}+1) - e^{-t}\right] + C_1\right\} \cdot e^t + \left\{\frac{1}{3}\left[\ln(e^{t}+1) - e^{t}\right] + C_2\right\} \cdot e^{-2t} =$$

$$= C_1 e^t + C_2 \cdot e^{-2t} + \frac{1}{3}\left[\ln(e^{-t}+1) - e^{-t}\right] \cdot e^t + \frac{1}{3}\left[\ln(e^{t}+1) - e^{t}\right] \cdot e^{-2t}$$