MAXIMUM LIKELIHOOD ESTIMATION

Maximum Likelihood Estimation is a systematic technique for estimating parameters in a probability model from a data sample. Suppose a sample $x_1, ..., x_n$ has been obtained from a probability model specified by mass or density function $f_X(x;\theta)$ depending on parameter(s) θ lying in parameter space Θ. The **maximum likelihood estimate** or **m.l.e.** is produced as follows;

STEP 1 Write down the **likelihood function**, $L(\theta)$, where

$$L(\theta) = \prod_{i=1}^{n} f_X(x_i;\theta)$$

that is, the product of the n mass/density function terms (where the ith term is the mass/density function evaluated at x_i) viewed as a function of θ.

STEP 2 Take the natural log of the likelihood, collect terms involving θ.

STEP 3 Find the value of $\theta \in \Theta$, $\hat{\theta}$, for which $\log L(\theta)$ is maximized, for example by differentiation. If θ is a single parameter, find $\hat{\theta}$ by solving

$$\frac{d}{d\theta}\left\{\log L(\theta)\right\} = 0$$

in the parameter space Θ. If θ is vector-valued, say $\theta = (\theta_1, ..., \theta_k)$, then find $\hat{\theta} = (\hat{\theta}_1, ..., \hat{\theta}_k)$ by simultaneously solving the k equations given by

$$\frac{\partial}{\partial\theta_j}\left\{\log L(\theta)\right\} = 0 \quad j = 1, ..., k$$

in parameter space Θ. Note that, if parameter space Θ is a bounded interval, then the maximum likelihood estimate may lie on the boundary of Θ.

STEP 4 Check that the estimate $\hat{\theta}$ obtained in STEP 3 truly corresponds to a maximum in the (log) likelihood function by inspecting the second derivative of $\log L(\theta)$ with respect to θ. In the single parameter case, if the second derivative of the log-likelihood is negative at $\theta = \hat{\theta}$, then $\hat{\theta}$ is confirmed as the m.l.e. of θ (other techniques may be used to verify that the likelihood is maximized at $\hat{\theta}$).

EXAMPLE Suppose a sample $x_1, ..., x_n$ is modelled by a Poisson distribution with parameter denoted λ, so that

$$f_X(x;\theta) \equiv f_X(x;\lambda) = \frac{\lambda^x}{x!}e^{-\lambda} \quad x = 0, 1, 2, ...$$

for some $\lambda > 0$. To estimate λ by maximum likelihood, proceed as follows.

STEP 1 Calculate the likelihood function $L(\lambda)$.

$$L(\lambda) = \prod_{i=1}^{n} f_X(x_i;\lambda) = \prod_{i=1}^{n}\left\{\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right\} = \frac{\lambda^{x_1+...+x_n}}{x_1!....x_n!}e^{-n\lambda}$$

for $\lambda \in \Theta = R^+$.

STEP 2 Calculate the log-likelihood $\log L(\lambda)$.

$$\log L(\lambda) = \sum_{i=1}^{n} x_i \log\lambda - n\lambda - \sum_{i=1}^{n}\log(x_i!)$$

STEP 3 Differentiate $\log L(\lambda)$ with respect to λ, and equate the derivative to zero to find the m.l.e..

$$\frac{d}{d\lambda}\left\{\log L(\lambda)\right\} = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n = 0 \Rightarrow \hat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

Thus the maximum likelihood estimate of λ is $\hat{\lambda} = \bar{x}$

STEP 4 Check that the second derivative of $\log L(\lambda)$ with respect to λ is negative at $\lambda = \hat{\lambda}$.

$$\frac{d^2}{d\lambda^2}\left\{\log L(\lambda)\right\} = -\frac{\sum_{i=1}^{n} x_i}{\lambda^2} < 0 \quad \text{at } \lambda = \hat{\lambda}$$

For the Accident Statistics Data on a previous handout we had $n = 647$ data as follows

Number of accidents	Frequency
0	447
1	132
2	42
3	21
4	3
5	2

so that the estimate of λ if a Poisson model is assumed is

$$\hat{\lambda}_{ML} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{(447\times 0) + (132\times 1) + (42\times 2) + (21\times 3) + (3\times 4) + (2\times 5)}{647} = 0.465$$