PRATIMA BHAKTA

4th Semester

Econometrics (ML Method)

Maximum Likelihood Methods

The maximum likelihood method is another method for obtaining estimates of the parameters of a population from a random sample.

Assume that the variable $X \sim N(\mu, \sigma_x^2)$. A (continuous) random variable X is said to be normally distributed if its Probability distribution function has the following form:

$$f(x_i) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \cdot \exp\left\{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma_x}\right)^2\right\}$$

We want to obtain maximum likelihood estimates of the parameters μ and σ_x^2 from a sample of n independent observations of X. We form the likelihood function $L(X_1, X_2 -, X_n ; \mu, \sigma_x^2)$, starting from the individual probabilities of the sample values

$$f(x_1) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \cdot \exp\left\{-\frac{1}{2}\left(\frac{x_1-\mu}{\sigma_x}\right)^2\right\}$$

$$f(x_2) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \cdot \exp\left\{-\frac{1}{2}\left(\frac{x_2-\mu}{\sigma_x}\right)^2\right\}$$

- - - - - - - - - - - - - - - - -

$$f(x_n) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \cdot \exp\left\{-\frac{1}{2}\left(\frac{x_n-\mu}{\sigma_x}\right)^2\right\}$$

The total probability of obtaining all the values in the sample is the product of the individual probabilities given that each observation is independent of the others. The joint probability of the n sample values is given by the likelihood function

$$L = f(x_1, x_2, \ldots, x_n; \mu, \sigma_x^2)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_x^2}} \cdot \exp\left\{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma_x}\right)^2\right\}$$

where $\prod_{i=1}^{n}$ denotes the product of n terms.

The maximum likelihood principle as applied to a simple linear regression model:-

our model is

$$Y = b_0 + b_1 X + u.$$

where, $u \sim N(0, \sigma_u^2)$ and $E(u_i u_j) = 0$

We take a random sample of X and Y values and we want to obtain maximum likelihood estimates of b_0 and b_1.

Given $u \sim N(0, \sigma_u^2)$

it can be shown that,

$$Y_i = N\{(b_0 + b_1 X_i), \sigma_u^2\}$$

that is Y_i is normally distributed with mean

$$E(Y_i) = E(b_0 + b_1 X_i) + E(u) = b_0 + b_1 X_i$$

and variance equal to the variance of the component σ_u^2.

For any particular sample value Y_i the individual probability is

$$f(Y_i) = \frac{1}{\sqrt{2\pi\sigma_u^2}} \cdot \exp\left[-\frac{1}{2\sigma_u^2}\{Y_i - (b_0 + b_1 X_i)\}^2\right]$$

The joint probability (likelihood function) of all n sample values is the product of the individual probabilities, since, by assumption, the Y values are independent of one another by assumption, the Y values are independent of one another

$$L = f(Y_1, Y_2 --- Y_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_u^2}} \cdot \exp\left[-\frac{1}{2\sigma_u^2}\{Y_i - (b_0 + b_1X_i)\}^2\right]$$

or, applying the rule of exponentials,

$$L = \left\{\frac{1}{\sqrt{2\pi\sigma_u^2}}\right\}^n \cdot \exp\left[\sum_{i=1}^{n}\left(-\frac{1}{2\sigma_u^2}\right)\{Y_i - (b_0 + b_1X_i)\}^2\right]$$

Given that σ_u^2 is constant we may take the relevant term out of the summation

$$L = \left\{\frac{1}{\sqrt{2\pi\sigma_u^2}}\right\}^n \cdot \exp\left\{-\frac{1}{2\sigma_u^2}\sum_{i=1}^{n}(Y_i - \tilde{b}_0 - b_1X_1)^2\right\}$$

To minimise L we take the partial derivatives with respect to $\tilde{b}_0$ and $\tilde{b}_1$ and equal to zero

$$\frac{\partial L}{\partial \tilde{b}_0} = 0 \text{ and } \frac{\partial L}{\partial \tilde{b}} = 0$$

our problem of maximisation of the likelihood function reduces to the minimisation of the exponent, ignoring the term outside the summation which is constant and hence does not affect the derivatives; thus we minimise

$$\Sigma \left(Y_i - \tilde{b}_0 - \tilde{b}_1 x_1 \right)^2$$

The necessary condition for a minimum is that the partial derivatives with respect to $\tilde{b}_0$ and $\tilde{b}_1$ be zero

$$\frac{\partial \, \Sigma \left(Y_i - \tilde{b}_0 - \tilde{b}_1 x_i \right)^2}{\partial \tilde{b}_0} = 0$$

$$\frac{\partial \, \Sigma \left(Y_i - \tilde{b}_0 - \tilde{b}_1 x_i \right)^2}{\partial \tilde{b}_1} = 0$$

working out the partial derivatives we obtain the equations

$$\Sigma Y_i = n\tilde{b}_0 + \tilde{b}_1 \Sigma X_i$$

$$\Sigma Y_i X_i = \tilde{b}_0 \Sigma X_i + \tilde{b}_1 \Sigma X_i^2$$

These equations are identical with the normal equations of OLS. Thus the maximum likelihood estimates ($\tilde{b}_0$ and $\tilde{b}_1$) are identical to the OLS estimates.

Scanned with CamScanner

: Maximum Likelihood Estimation of two-variable Regression Model :

Assume that in the two variable model:

$Y_i = b_0 + b_1 X_i + u_i$ the Y_i are normally and independently distributed with mean $= b_0 + b_1 X_i$ and variance σ_u^2. As a result, the joint probability (likelihood function) function of $Y_1, Y_2, \text{---} Y_n$, given the preceding mean and variance, can be written as

$$L = f(Y_i) = \frac{1}{\sigma_u \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{(Y_i - b_0 - b_1 X_i)^2}{\sigma_u^2}\right\}$$

If $Y_1, Y_2 \text{---} Y_n$ are known or given but b_0, b_1 and σ_u^2 are not known, the above function is called a likelihood function, denoted by $L(b_0, b_1, \sigma_u^2)$ and written as

$$L(b_0, b_1, \sigma_u^2) = \frac{1}{\sigma_u^n (\sqrt{2\pi})^n} \exp\left\{-\frac{1}{2} \sum \frac{(Y_i - b_0 - b_1 X_i)^2}{\sigma_u^2}\right\}$$

The method of maximum likelihood, consists in estimating the unknown parameters in such a manner that the probability of observing the given y's is as high (or maximum) as possible. Therefore, we have to find

(or maximum) as possible. [illegible]
the maximum of the above function. This is a straight-forward exercise in differential calculus. For differentiation it is easier to express the above function in the log term as follows:

$$\log L = -N \log \sigma - \frac{N}{2}\log(2\pi) - \frac{1}{2}\sum \frac{(Y_i - b_0 - b_1 X_i)^2}{\sigma_u^2}$$

$$= -\frac{N}{2}\log \sigma_u^2 - \frac{N}{2}\log(2\pi) - \frac{1}{2}\sum \frac{(Y_i - b_0 - b_1 X_i)^2}{\sigma_u^2}$$

Differentiating partially with respect to b_0, b_1 and σ_u^2, we obtain,

$$\frac{\partial \log L}{\partial b_0} = -\frac{1}{\sigma_u^2}\sum (Y_i - b_0 - b_1 X_i)(-1)$$

$$\frac{\partial \log L}{\partial b_1} = -\frac{1}{\sigma_u^2}\sum (Y_i - b_0 - b_1 X_i)(-X_i)$$

$$\frac{\partial \log L}{\partial \sigma_u^2} = -\frac{N}{2\sigma_u^2} + \frac{1}{2\sigma_u^4}\sum (Y_i - b_0 - b_1 X_i)^2$$

Setting these equations equal to zero (the first-order condition for optimization) and letting $\tilde{b}_0$, $\tilde{b}_1$ and $\tilde{\sigma}_u^2$ denote the ML estimators, we obtain,

$$\frac{1}{\tilde{\sigma}_u^2} \Sigma (Y_i - \tilde{b}_0 - \tilde{b}_1 X_i) = 0 \quad \text{------} \quad (1)$$

$$\frac{1}{\tilde{\sigma}_u^2} \Sigma (Y_i - \tilde{b}_0 - \tilde{b}_1 X_i) X_i = 0 \quad \text{------} \quad (2)$$

$$-\frac{N}{2\tilde{\sigma}_u^2} + \frac{1}{2\tilde{\sigma}_u^4} \Sigma (Y_i - \tilde{b}_0 - \tilde{b}_1 X_i)^2 = 0 \quad \text{---} \quad (3)$$

After simplifying, equations (1) and (2) yield,

$$\Sigma Y_i = N\tilde{b}_0 + \tilde{b}_1 \Sigma X_i$$

$$\Sigma Y_i X_i = \tilde{b}_0 \Sigma X_i + \tilde{b}_1 \Sigma X_i^2$$

Therefore, the ML estimators, the $\tilde{b}$'s, are the same as the OLS estimators.

These above equations are identical with the normal equations of OLS. Thus the maximum likelihood estimates ($\tilde{b}_0$ and $\tilde{b}_1$) are identical to the OLS estimates.