(*e*) **Non Stochastic :** The explanatory variable(s) is non stochastic variable and is measured without error. *i.e.*

U_i is independent of the explanatory variable(s).

$$E(X_i U_j) = X_i E(U_j) = 0 \; i, j = 1, 2, \ldots n$$

2.6. LEAST SQUARE ESTIMATION

The first method of estimating the relationship is termed as least suare estimation method. Here the line of best fit is said to be that which maximizes the sum of the squared residuals between the points of the graph and the points of the straight line.

Let n observations on Y and X are denoted as –

$$Y_1, Y_2, \ldots\ldots Y_n$$
$$X_1, X_2, \ldots\ldots X_n$$

And their mean value is

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \text{ and } \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

The estimated line is given as

$$\hat{Y} = \hat{\alpha} + \hat{\beta}X$$

Where $\hat{\alpha}$ and $\hat{\beta}$ are parameters estimated and $\hat{Y}$ is the estimated value of Y.

The scatter diagram from the given observation will be

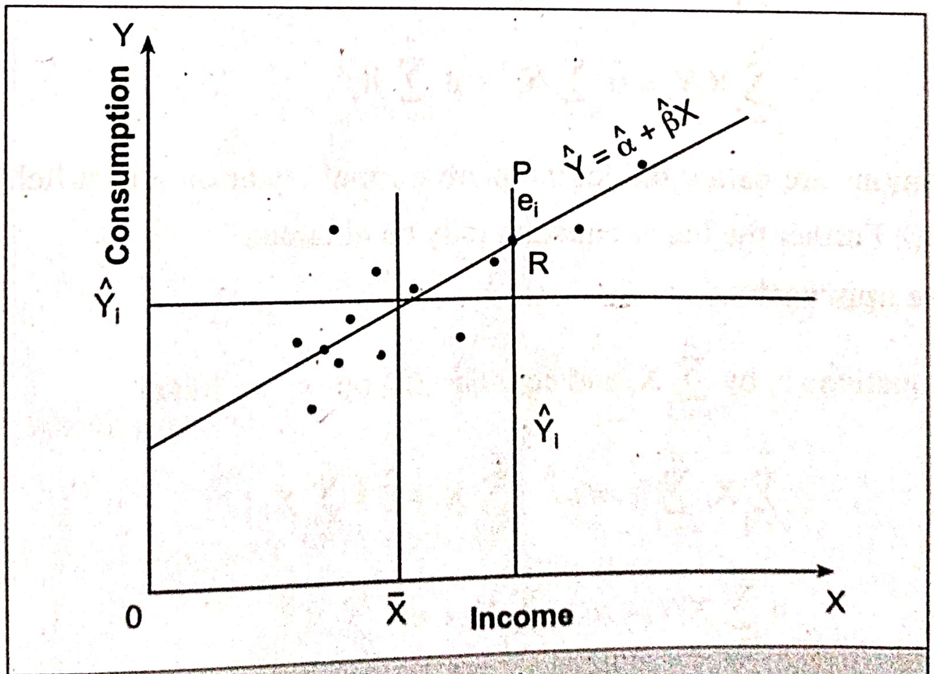

Fig. 2.1

Consider the point P on the scatter diagram, here the difference between the actual and the estimated value of Y is the error. Therefore,

$$e_i = Y_i - \hat{Y}_i = PR$$

The deviation of actual values from the estimated line will be positive or negative depend upon the position of actual points of lying above or below the estimated line. Square of this residuals will be positive

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_1 - \bar{Y}_i)^2$$

The principle of least square is to chose the values of estimators that will minimize the sum of squared deviations and the necessary condition is to take the partial deviation of $\sum_{i=1}^{n} e_i^2$ with respect to $\hat{\alpha}$ and $\hat{\beta}$

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{\alpha} - \hat{\beta}X_i)^2$$

Differentiating w.r.t $\hat{\alpha}$ and $\hat{\beta}$ respectively

$$\frac{\partial}{\partial\hat{\alpha}}\left[\sum_{i=1}^{n} e_i^2\right] = -2\sum_{i=1}^{n} (Y_i - \hat{\alpha} - \hat{\beta}X_i) = 0$$

And $$\frac{\partial}{\partial\hat{\beta}}\left[\sum_{i=1}^{n} e_i^2\right] = -2\sum_{i=1}^{n} X_i(Y_i - \hat{\alpha} - \hat{\beta}X_i) = 0.$$

Rearranging we get two normal equation

$$\sum_{i=1}^{n} Y_i = n\hat{\alpha} + \hat{\beta}\sum_{i=1}^{n} X_i \qquad \text{...(1)}$$

$$\sum_{i=1}^{n} X_iY_i = \hat{\alpha}\sum_{i=1}^{n} X_i^2 + \hat{\beta}\sum_{i=1}^{n} X_i^2 \qquad \text{...(2)}$$

These equations are called the least square normal equation and it helps to find out the value of $\hat{\alpha}$ and $\hat{\beta}$. Further the linear equation may be obtained.

Alternative approach :

Multiply equation (1) by $\sum_{i=1}^{n} X_i$ and equation (2) by 'n' we have

$$\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i = n\hat{\alpha}\sum_{i=1}^{n} X_i + \hat{\beta}\left(\sum_{i=1}^{n} X_i\right)^2 \qquad \text{...(3)}$$

$$n\sum_{i=1}^{n} X_iY_i = n\hat{\alpha}\sum_{i=1}^{n} X_i + n\hat{\beta}\sum_{i=1}^{n} X_i^2 \qquad \text{...(4)}$$

Subtracting equation (3) from equation (4)

$$n\sum_{i=1}^{n} X_iY_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i = n\hat{\beta}\sum_{i=1}^{n} X_i\,\hat{\beta}\left(\sum_{i=1}^{n} X_i\right)^2$$

or
$$\sum_{i=1}^{n} X_iY_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n} = \hat{\beta}\left\{\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}\right\}$$

$$\hat{\beta}_{yx} = \frac{\sum_{i=1}^{n} X_iY_i - \left(\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i\right)/n}{\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2/n}$$

Similarly

$$\hat{\beta}_{xy} = \frac{\sum_{i=1}^{n} X_iY_i - \left(\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i\right)/n}{\sum_{i=1}^{n} Y_i^2 - \left(\sum_{i=1}^{n} Y_i\right)^2/n}$$

If the deviations are taken from actual mean, then the sum of the deviations is always zero. Therefore

$$\hat{\beta}_{yx} = \frac{\sum_{i=1}^{n} x_iy_i}{\sum_{i=1}^{n} x_i^2} \quad \text{and} \quad \hat{\beta}_{xy} = \frac{\sum_{i=1}^{n} x_iy_i}{\sum_{i=1}^{n} y_i^2}$$

Where, $x_i = X_i - \overline{X}$ and $y_i = Y_1 - \overline{Y}$

Or,
$$\hat{\beta}_{yx} = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_1 - \overline{Y})}{\sum_{i=1}^{n} (X_i - \overline{X})^2}$$

and
$$\hat{\beta}_{xy} = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}$$

Now, estimated regression equation is

$$Yi = \hat{\alpha} + \hat{\beta}X_i$$

Taking summation and dividing by 'n' we have

$$\frac{\sum_{i=1}^{n} Y_i}{n} = \hat{\alpha} + \hat{\beta}\frac{\sum_{i=1}^{n} X_i}{n}$$

or
$$\overline{Y} = \hat{\alpha} + \hat{\beta}\overline{X}$$

or
$$\overline{Y} = \hat{\beta}\overline{X} = \hat{\alpha}$$

Now computing the values of $\hat{\alpha}$ and $\hat{\beta}$ we can estimate the regression line Y on X and X on Y.

2.6.1. Properties of Least Square Estimators

Property I : The least square estimators are unbiased estimators.

We have $$\beta_{xy} = \frac{\sum x_i y_i}{\sum x_i^2}$$

We wish to relate $\hat{\beta}$ to actual β. We know

$$Y_i = \alpha + \beta X_i + U_i$$

Taking deviation from actual mean –

$$y_i = \beta x_i + (u_i - \bar{u})$$

Where $$y_i = Y_i - \bar{Y} \text{ and } x_i = X_i - \bar{X}$$

Putting the values in the $\hat{\beta}$ formula

$$\beta_{yx} = \frac{\sum_{i=1}^{n} x_i (\beta x_i + u_i - \bar{u})}{\sum_{i=1}^{n} x_i^2}$$

or, $$\hat{\beta} = \beta \frac{\sum x_i^2}{\sum x_i^2} + \frac{\sum x_i u_i}{\sum x_i} - \frac{\bar{u} \sum x_i}{\sum x_i^2}$$

or, $$\hat{\beta} = \beta + \frac{\sum x_i u_i}{\sum x_i}$$ ($\because$ Error term has zero mean)

Taking expectation of both side

$$E(\hat{\beta}) = \beta + E\left(\frac{\sum x_i u_i}{\sum x_i}\right)$$

or $$E(\hat{\beta}) = \beta + 0$$

Thus, $$E(\hat{\beta}) = \beta$$ $\{\because E(u_i) = 0\}$

It follows that $\hat{\beta}$ is an unbiased estimator of β.

Property II : The OLS estimators are linear :

We know $E(\hat{\beta}) = \beta$

It indicates that $\hat{\beta}$ is a linear function of β and we know

$$\hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2}$$

Or $$\hat{\beta} = \frac{\sum x_i(Y - \bar{Y})}{\sum x_i^2} \qquad (\because y_1 - Y_i - \bar{Y})$$

$$= \frac{\sum x_i Y_i}{\sum x_i^2} - \frac{\bar{Y} \sum x_i}{\sum x_i^2}$$

Or $$\hat{\beta} = \frac{\sum x_i Y_i}{\sum x_i^2} \qquad (\because \sum x_i = 0)$$

$$= \frac{\sum x_i}{\sum x_i^2} Y_i$$

Let $= \frac{x_i}{\sum x_i^2} = k$ a constant.

Then, $$\hat{\beta} = \sum k Y_i$$

Thus, $\hat{\beta}$ is a linear constant.

Property III : The OLS estimators are best estimators :

Here we will have to prove that variance of $\hat{\beta}$ is the best and is possible if we prove that has the smallest variance. We know

$$\hat{\beta} - \beta = \frac{\sum x_i u_i}{\sum x_i^2}$$

By definition

$$\text{Variance of } \hat{\beta} = E(\hat{\beta} - \beta)^2$$

Or $$\text{Variance } (\hat{\beta}) = E\left(\frac{\sum x_i u_i}{\sum x_i^2}\right)^2$$

$$= \left[\frac{1}{(\sum x_i^2)^2}(x_1^2 u_1^2 + x_2^2 u_2^2 + \dots x_n^2 u_n^2 + \dots 2x_1 x_2 u_1 u_2 + \dots 2x_{n-1} x_n u_{n-1} u_n\right]$$

$$= \frac{1}{(\sum x_i^2)^2}[x_1^2 E(u_i^2) + x_2^2 E(u_2^2) + \dots x_n^2 E(u_n^2) + 2x_1 x_2 E(u_1 u_2) + \dots 2x_{n-1} x_n E(u_{n-1} u_n)]$$

Or $$\text{variance } (\hat{\beta}) = \frac{1}{(\sum x_i^2)^2}(x_1^2 \sigma_u^2 + x_2^2 \sigma_u^2 + \dots + x_n^2 \sigma_u^2)$$

Since, $E(u_1^2) = \sigma_u^2$ and $E(u_i u_j) = 0$

Or $$\text{variance } (\hat{\beta}) = \frac{\sigma_u^2 \sum x_i^2}{(\sum x_i^2)^2} = \frac{\sigma_u^2}{\sum x_i^2}$$

Or $$\text{variance } (\hat{\beta}) \leq \text{variance } (u)$$

Hence, $\hat{\beta}$ is the best estimator.

Now, we need to prove that the variance is least.

Let us take another arbitrary estimator of β

$$\hat{\beta}^* = \sum_{i=1}^{n} C_i Y_i$$

Where $C_i = \frac{x_i}{\sum x_i^2} + d_i$ (d_i is any constant)

Or,

$$\hat{\beta}^* = \sum C_i (\alpha + \beta X_i + u_i)$$
$$= \alpha \sum C_i + \beta \sum C_i X_i + \sum C_i u_i$$

Taking expectations

$$E(\hat{\beta})^* = \alpha \sum C_i + \beta \sum C_i X_i + E(u_i) \sum C_i$$
$$= \alpha \sum C_i + \beta \sum C_i X_i \qquad (\because \ E(u_i) = 0)$$

If $\sum C_i = 0$ and $\sum C_i X_i = 1$, then and only then

$$E(\hat{\beta})^* = \beta$$

This condition can be fulfilled if

$$\sum d_i = 0 \text{ and } \sum d_i X_i = 1 = \sum d_i x_i$$

Therefore, $\text{Var}(\hat{\beta})^* = E(\hat{\beta}^* - \beta)^2 = E(\sum C_i u_i)^2]$

Solving we get,

$$\text{Var}(\hat{\beta})^* = \sigma_u^2 \sum C_i^2 \qquad [\because \ E(u_i^2 = \sigma_u^2]$$

Where

$$C_i = \frac{x_i}{\sum x_i^2} + d_i$$

$$\therefore \qquad C_i^2 = \frac{x_i^2}{(\sum x_i^2)^2} + d_i^2 + 2\frac{x_i}{\sum x_i^2} d_i$$

$$\text{Var}(\hat{\beta}^*) = \sigma_u^2 \left[\frac{\sum x_i^2}{(\sum x_i^2)^2} + \sum d_i^2 + 2d_i \frac{\sum x_i}{\sum x_i^2}\right]$$
$$= \sigma_u^2 \left[\frac{1}{\sum x_i^2} + \sum d_i^2\right] \qquad \left(\because \ \frac{\sum x_i}{\sum x_i^2} = 0\right)$$
$$= \frac{\sigma_u^2}{\sum x_i^2} + \sigma_u^2 \sum d_i^2$$
$$= \text{Var}(\hat{\beta}) + \sigma_u^2 \sum d_i^2$$

If $d_i = 0$. Then only $\text{Var}(\hat{\beta})$ will be equal to $\text{Var}(\hat{\beta}^*)$ otherwise it would be greater $\text{Var}(\hat{\beta})$.

$$\therefore \qquad \text{Var}(\hat{\beta}^*) \geq \text{Var}(\hat{\beta})$$