6.3 ESTIMATION OF REGRESSION PARAMETERS

Estimation of α and β by least squares method (OLS) or classical least squares (CLS) involves finding values for the estimates $\hat{\alpha}$ and $\hat{\beta}$ which will minimise the sum of the squared residuals: Σe_i^2.

From fitted regression line:

$Y_i = \hat{\alpha} + \hat{\beta}X + e_i$; we obtain:

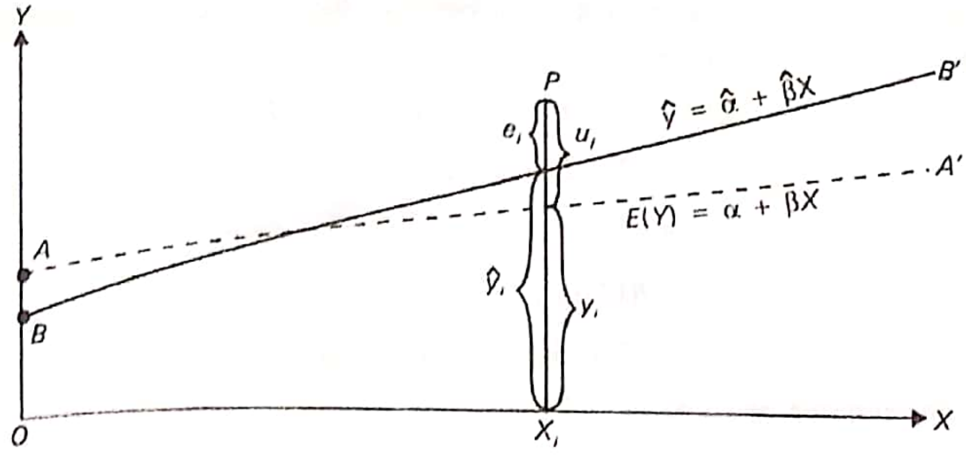

Fig. 6.4

$$e_i = Y_i - (\hat{\alpha} + \hat{\beta}_i)$$

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

To find the values of α and β that minimise this sum, we have to differentiate with respect to $\hat{\alpha}$ and $\hat{\beta}$ and set the partial derivatives equal to zero.

$$\partial/\partial\hat{\alpha}\left[\Sigma e_i^2\right] = -2\Sigma (Y_i - \hat{\alpha} - \hat{\beta}X_i) = 0$$

$$\partial/\partial\hat{\beta}\left[\Sigma e_i^2\right] = -2\Sigma X_i (Y_i - \hat{\alpha} - \hat{\beta}X_i) = 0$$

or, equivalently, $\Sigma Y_i = n\hat{\alpha} + \hat{\beta}\Sigma X_i$ ...(6.1)

$$\Sigma X_i Y_i = \hat{\alpha}\Sigma X_i + \hat{\beta}\Sigma X_i^2 \qquad ...(6.2)$$

From (6.1) we have,

$$n\hat{\alpha} = \Sigma Y_i - \hat{\beta}\Sigma X_i$$

$$\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X} \qquad ...(6.3)$$

Substituting value of $\hat{\alpha}$ in (6.2) we get

$$\Sigma X_i Y_i = (\overline{Y} - \hat{\beta}\overline{X})\, \Sigma X_i + \hat{\beta}\Sigma X_i^2$$

$$\Sigma X_i Y_i = \overline{Y}\,\Sigma X_i - \hat{\beta}\overline{X}\,\Sigma X_i + \hat{\beta}\Sigma X_i^2$$

$$\Sigma X_i Y_i - \overline{Y}\,\Sigma X_i = \hat{\beta}(\Sigma X_i^2 - \overline{X}\,\Sigma X_i)$$

$$\therefore \quad \hat{\beta} = \frac{\Sigma X_i\, Y_i - \overline{Y}\,\Sigma X_i}{\Sigma X_i^2 - \overline{X}\Sigma X_i}$$

$$= \frac{n\Sigma X_i\, Y_i - \Sigma Y_i \Sigma X_i}{n\Sigma X_i^2 - (\Sigma X_i)^2} \qquad ...(6.4)$$

(6.4) can also be written in a somewhat different way.

Numerator of (6.4) is:

$$
\begin{aligned}
n\Sigma X_i Y_i - \Sigma X_i \Sigma Y_i &= n\Sigma X_i Y_i - \Sigma Y_i \Sigma X_i + (\Sigma X_i \Sigma Y_i - \Sigma X_i \Sigma Y_i) \\
&= n\Sigma X_i Y_i - \Sigma Y_i \Sigma X_i - \Sigma X_i \Sigma Y_i + \Sigma X_i \Sigma Y_i \\
&= n\Sigma X_i Y_i - n\overline{X}\,\Sigma Y_i - n\overline{Y}\Sigma X_i + n^2\,\overline{X}\,\overline{Y} \\
&= n(\Sigma X_i Y_i - \overline{X}\,\Sigma Y_i - \overline{Y}\Sigma X_i + n\,\overline{X}\,\overline{Y}) \\
&= n\{\Sigma\,(X_i - \overline{X})\,(Y_i - \overline{Y})\}
\end{aligned}
$$

Denominator of (6.4) is:

$$
\begin{aligned}
n\Sigma X_i^2 - (\Sigma X_i)^2 &= n\Sigma X_i^2 - 2(\Sigma X_i)^2 + (\Sigma X_i)^2 \\
&= n\Sigma X_i^2 - 2\Sigma X_i \Sigma X_i + (\Sigma X_i)^2 \\
&= n\Sigma X_i^2 - 2n\overline{X}\,\Sigma X_i + n^2\,\overline{X}^2 \\
&= n(\Sigma X_i^2 - 2\,\overline{X}\Sigma X_i + n\overline{X}^2) \\
&= n\Sigma\,(X_i - \overline{X})^2
\end{aligned}
$$

$$\therefore \qquad \hat{\beta} = \frac{n\Sigma\,(X_i - \overline{X})\,(Y_i - \overline{Y})}{n\Sigma\,(X_i - \overline{X})^2}$$

Now denoting $(X_i - \overline{X})$ as x_i and $(Y_i - \overline{Y})$ as y_i we get;

$$\hat{\beta} = \frac{\Sigma x_i y_i}{\Sigma x_i^2} \qquad \text{...(6.5)}$$

6.4 STATISTICAL PROPERTIES OF LEAST SQUARES ESTIMATOR

(i) *Linearity*

$$\hat{\beta} = \frac{\Sigma(X_i - \overline{X})(Y_i - \overline{Y})}{\Sigma(X_i - \overline{X})^2} \qquad \text{(from 6.5)}$$

$$\hat{\beta} = \frac{\Sigma Y_i(X_i - \overline{X}) - \overline{Y}\,\Sigma(X_i - \overline{X})}{\Sigma(X_i - \overline{X})^2}$$

$$\hat{\beta} = \frac{\Sigma Y_i(X_i - \overline{X})}{\Sigma(X_i - \overline{X})^2} \qquad [\because \overline{Y}\,\Sigma(X_i - \overline{X}) = 0]$$

$$\hat{\beta} = \frac{\Sigma Y_i x_i}{\Sigma x_i^2}$$

Let us suppose that,

$$\frac{x_i}{\Sigma x_i^2} = k_i \ (i = 1, \ldots, n)$$

$$\therefore \qquad \hat{\beta} = \sum_{i=1}^{n} k_i Y_i \qquad \ldots(6.6)$$

Similarly (6.3) gives $\hat{a} = \overline{Y} - \hat{\beta}\overline{X} = \frac{1}{n}\Sigma Y_i - \overline{X}\,\Sigma k_i Y_i$

$$\therefore \qquad \hat{\alpha} = \Sigma\left[\frac{1}{n} - \overline{X} k_i\right] Y_i \qquad \ldots(6.7)$$

Thus both $\hat{\alpha}$ and $\hat{\beta}$ are expressed as linear functions of the Y's.

(ii) *Unbiasedness:*

$$\begin{aligned} \hat{\beta} &= \Sigma k_i Y_i && \text{(from 6.6)} \\ &= \Sigma k_i(\alpha + \beta X_i + U_i) \\ &= \alpha \Sigma k_i + \beta \Sigma k_i X_i + \Sigma k_i U_i && \ldots(6.8) \end{aligned}$$

$\because \quad k_i = \dfrac{x_i}{\Sigma x_i^2} \qquad \therefore \ \Sigma k_i = \dfrac{\Sigma x_i}{\Sigma x_i^2} = 0;$

and, $\quad \Sigma k_i X_i = \Sigma k_i \ (x_i + \overline{X}) = \dfrac{\Sigma x_i^2}{\Sigma x_i^2} = 1 \qquad \text{...(6.9)}$

Substituting these values in (6.8) we obtain,

$$\hat{\beta} = \beta + \Sigma k_i U_i \qquad \text{...(6.10)}$$

$$E(\hat{\beta}) = E(\beta) + \Sigma k_i E(U_i) = \beta$$

Equation (6.7) gives,

$$\hat{\alpha} = \Sigma\left(\frac{1}{n} - \overline{X} k_i\right) Y_i$$

$$= \Sigma\left(\frac{1}{n} - \overline{X} k_i\right) (\alpha + \beta X_i + U_i)$$

$$= \alpha + \beta \frac{1}{n} \Sigma X_i + \frac{1}{n} \Sigma U_i - \alpha \overline{X} \Sigma k_i - \beta \overline{X} \Sigma k_i X_i - \overline{X} \Sigma k_i U_i$$

$$= \alpha + \beta \overline{X} + \frac{1}{n} \Sigma U_i - \beta \overline{X} - \overline{X} \Sigma k_i U_i$$

$$= \alpha + \frac{1}{n} \Sigma U_i - \overline{X} \Sigma k_i U_i \qquad \text{...(6.11)}$$

$$\therefore \ E(\hat{\alpha}) = \alpha + \frac{1}{n} \Sigma E(U_i) - \overline{X} \Sigma k_i E(U_i)$$

$$E(\hat{\alpha}) = \alpha$$

Thus, we prove that $\hat{\alpha}$ and $\hat{\beta}$ are unbiased estimators of α and β.

(iii) *Minimum variance of* $\hat{\alpha}$ *and* $\hat{\beta}$: Now we have to establish that out of the class of linear unbiased estimators of α and β; $\hat{\alpha}$ and $\hat{\beta}$ possess the smallest sampling variances. For this we shall first obtain the variance of $\hat{\beta}$ and then establish that it is the minimum variance.

$$\text{Var}(\hat{\beta}) = E[(\hat{\beta} - \beta)^2]$$

$$= E[\Sigma k_i U_i)^2] \qquad \text{...from equation (6.10)}$$

$$= E\left[k_1^2 U_1^2 + k_2^2 U_2^2 ... + k_n^2 U_n^2 + 2k_1 k_2 U_1 U_2 + ... + 2k_{n-1} k_n U_{n-1} U_n\right]$$

$$= E\left[k_1^2 U_1^2 + k_2^2 U_2^2 ... + k_n^2 U_n^2\right] + E\left[2k_1 k_2 U_1 U_2 + ... + 2k_{n-1} k_n U_{n-1} U_n\right]$$

$$= E\left[\Sigma\left(k_i^2 U_i^2\right)\right] + 2E\left[\sum_{i \neq j} k_i k_j U_i U_j\right]$$

$$= \Sigma k_i^2\, E(U_i^2) + 2\Sigma k_i k_j E(U_i\, U_j) \;=\; \sigma^2 \Sigma k_i^2 \quad [\because E(U_i U_j) = 0]$$

$$\Sigma k_i = \frac{\Sigma x_i}{\Sigma x_i^2}$$

$$\therefore \qquad \Sigma k_i^2 = \frac{\Sigma x_i^2}{(\Sigma x_i^2)^2} = \frac{1}{\Sigma x_i^2}$$

$$\therefore \;\; \text{Var}\,(\hat{\beta}) = \sigma^2 \Sigma k_i^2 = \frac{\sigma^2}{\Sigma x_i^2} \qquad \text{...(6.12)}$$

$$\text{Var}\,(\hat{\alpha}) = E\left[(\hat{\alpha} - \alpha)^2\right]$$

$$= E\left[\Sigma\left(\frac{1}{n} - \overline{X} k_i\right)^2 U_i^2\right] \qquad \text{(from equation 6.11)}$$

$$= \sigma^2 \Sigma\left(\frac{1}{n} - \overline{X} k_i\right)^2$$

$$= \sigma^2 \Sigma\left(\frac{1}{n^2} - \frac{2}{n}\,\overline{X} k_i + \overline{X}^2 k_i^2\right)$$

$$= \sigma^2\left(\frac{1}{n} - \frac{2\overline{X}}{n}\Sigma k_i + \overline{X}^2\, \Sigma k_i^2\right)$$

$$= \sigma^2\left(\frac{1}{n} + \frac{\overline{X}^2}{\Sigma x_i^2}\right) \qquad \left(\because \Sigma k_i = 0 \text{ and } \Sigma k_i^2 = \frac{1}{\Sigma x_i^2}\right)$$

Again; $$\frac{1}{n} + \frac{\overline{X}^2}{\Sigma x_i^2} = \frac{\Sigma x_i^2 + n\overline{X}^2}{n\Sigma x_i^2} = \frac{\Sigma X_i^2}{n\Sigma x_i^2}$$

$$\therefore \qquad \text{Var}\,(\hat{\alpha}) = \sigma^2\left(\frac{1}{n} + \frac{\overline{X}^2}{\Sigma x_i^2}\right) = \sigma^2\left(\frac{\Sigma X_i^2}{n\Sigma x_i^2}\right) \qquad \text{...(6.13)}$$

$\hat{\alpha}$ and $\hat{\beta}$ are also 'Best' estimators: In order to establish that $\hat{\beta}$ possesses the minimum variance property (Best), we compare its variance with that of some alternative, unbiased estimator of β, say β^*.

Suppose $\beta^* = \Sigma w_i Y_i$ where constant $w_i \neq k_i$, but $w_i = k_i + c_i$

$$\therefore \qquad \beta^* = \Sigma w_i (\alpha + \beta X_i + U_i)$$

$$= \alpha \Sigma w_i + \beta \Sigma w_i X_i + \Sigma w_i U_i \qquad \text{...(6.14)}$$

$$\therefore \quad E(\beta^*) = \alpha\Sigma w_i + \beta\Sigma w_i X_i \qquad [\because E(U_i) = 0].$$

Since β^* is assumed to be an unbiased estimator which implies that $\Sigma w_i = 0$ and $\Sigma w_i X_i = 1$ in the above equation.

But, $\quad \Sigma w_i = \Sigma(k_i + c_i) = \Sigma k_i + \Sigma c_i$

$\therefore \quad \Sigma c_i = 0 \quad \because \Sigma k_i = \Sigma w_i = 0$

Again, $\Sigma w_i X_i = \Sigma (k_i + c_i) X_i = \Sigma k_i X_i + \Sigma c_i X_i$

$\therefore \quad \Sigma c_i X_i = 0 \quad \because \Sigma w_i X_i = 1$ and $\Sigma k_i X_i = \Sigma k_i x_i = 1$.

Also $\therefore$ $\Sigma c_i x_i = \Sigma c_i X_i + \bar{X}\Sigma c_i = 0$

Thus we have shown that if β^* is to be unbiased estimator then following results must hold true.

$$\Sigma w_i = 0, \ \Sigma w_i X_i = 1, \ \Sigma c_i = 0, \ \Sigma c_i X_i = \Sigma c_i x_i = 0 \qquad \text{...(6.15)}$$

The variance of this assumed estimator β^* is then

$$\text{Var}(\beta^*) = E(\beta^* - \beta)^2]$$

$$= E\left[(\Sigma w_i U_i)^2\right] \qquad \text{(from 6.14)}$$

$$= \sigma^2 \Sigma w_i^2$$

[By following exactly the same arguments that we used in obtaining Var ($\hat{\beta}$).]

$$\therefore \quad \text{Var}(\beta^*) = \sigma^2 \Sigma w_i^2$$

But $\quad \Sigma w_i^2 = \Sigma(k_i + c_i)^2 = \Sigma k_i^2 + 2\Sigma k_i c_i + \Sigma c_i^2$

$$\Sigma k_i c_i = \Sigma c_i \cdot \frac{x_i}{\Sigma x_i^2} = \frac{\Sigma c_i x_i}{\Sigma x_i^2} = 0 \qquad \text{(By 6.15)}$$

$\therefore \quad \Sigma w_i^2 = \Sigma k_i^2 + \Sigma c_i^2$; so that

$$\text{Var}(\beta^*) = \sigma^2\left(\Sigma k_i^2 + \Sigma c_i^2\right) = \sigma^2 \Sigma k_i^2 + \sigma^2 \Sigma c_i^2$$

$$\text{Var}(\beta^*) = \text{Var}(\hat{\beta}) + \sigma^2 \Sigma c_i^2$$

Σc_i^2 must be positive; so that Var $(\beta^*) >$ Var $(\hat{\beta})$

In case $\Sigma c_i^2 = 0$, then Var $(\beta^*) =$ Var $(\hat{\beta})$.

This proves that $\hat{\beta}$ possesses minimum variance property.

In the similar way we can prove that the least squares constant intercept $\hat{\alpha}$ possesses minimum variance, in other words it is also a 'Best' estimator.

We take a new estimator α^*, which we assume to be a linear function of the Y_i's, with weights $w_i = k_i + c_i$, as earlier.

Illustration, 9—A sample of 20 observations on X and Y gave the following data :—

$\Sigma Y = 21 \cdot 9$ $\quad \Sigma (Y - \bar{Y})^2 = 86 \cdot 9$

$\Sigma X = 186 \cdot 2$ $\quad \Sigma (X - \bar{X})^2 = 215 \cdot 4,$ $\quad \Sigma (X - \bar{X})(Y - \bar{Y}) = 106 \cdot 4$

Answer the followings :—

(a) Estimate the regression of Y on X.

(b) Estimate the regression of X on Y.

(c) Compute the mean value of Y corresponding to $X = 10$.

(d) Compute the mean value of X corresponding to $Y=1.5$

(M.A., Meerut, 1975)

Solution : (a) Let regression line of Y on X be

$$Y=\alpha+\beta X$$

we know that

$$\hat{\beta}_{YX}=\frac{\Sigma (X-\bar{X})(Y-\bar{Y})}{\Sigma (X-\bar{X})^2}=\frac{106.4}{215.4}=0.49$$

$$\bar{X}=\frac{\Sigma X}{n}=\frac{186.2}{20}=9.31$$

$$\bar{Y}=\frac{\Sigma Y}{n}=\frac{21.9}{20}=1.09$$

$$\hat{\alpha}=\bar{Y}-\hat{\beta}\bar{X}=1.09-(0.49\times 9.31)=-3.47$$

Thus, estimated regression line of Y on X is

$$\mathbf{Y=-3.47+0.49\ X}$$

(b) Now, let the regreression line of X on Y be

$$X=\gamma+\delta Y$$

$$\hat{\delta}_{XY}=\frac{\Sigma (X-\bar{X})(Y-\bar{Y})}{\Sigma (Y-\bar{Y})^2}=\frac{106.4}{86.9}=1.22$$

$$\hat{\gamma}=\bar{X}-\hat{\delta}\bar{Y}=9.31-(1.22\times 1.09)=7.98$$

Thus estimated regression line of X on Y is

$$\mathbf{X=7.98+1.22\ Y}$$

(c) when $X=10$

then $Y=-3.47+(0.49\times 10)=\mathbf{1.43}$

(d) when $Y=1.5$

then $X=7.98+(1.22\times 1.5)=\mathbf{9.81}$

Illustration 10—The following data were obtained in a sample study :—

$\Sigma X=56,\quad \Sigma Y=40,\quad \Sigma X^2=524,\quad \Sigma Y^2=256,$

$\Sigma XY=364,\quad N=20$

Answer all of the following :—

(a) Estimate the regression line $Y=\alpha+\beta X$.

(b) Estimate the regression line $X=\gamma+\delta Y$

(c) Compute the value of Y corresponding to a value 7 for X.

(d) Compute the value of X corresponding to a value 3 for Y.

(M.A., Meerut, 1975)

Solution : (a) Estimated regression line is

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

where, $$\hat{\beta}=\frac{\Sigma XY-\frac{\Sigma X.\Sigma Y}{n}}{\Sigma X^2-\frac{(\Sigma X)^2}{n}}$$

$$=\frac{364-\frac{56\times 40}{20}}{524-\frac{56\times 56}{20}}=\frac{252}{367\cdot 2}=0\cdot 686$$

$$\bar{X}=\frac{\Sigma X}{n}=\frac{56}{20}=2\cdot 8,\quad \bar{Y}=\frac{\Sigma Y}{n}=\frac{40}{20}=2$$

$$\hat{\alpha}=\bar{Y}-\hat{\beta}\ \bar{X}=2-0\cdot 686\times 2\cdot 8=2-1\cdot 921=0\cdot 079$$

Thus estimated regression line becomes

$$\mathbf{Y=0\cdot 079+\cdot 686X}$$

(b) Estimated regression line is

$$X=\gamma+\delta Y$$

where, $$\hat{\delta}=\frac{\Sigma XY-\frac{\Sigma X.\Sigma Y}{n}}{\Sigma Y^2-\frac{(\Sigma Y)^2}{n}}$$

$$=\frac{364-\frac{56\times 40}{20}}{256-\frac{40\times 40}{20}}=\frac{252}{256-80}$$

$$=\frac{252}{176}=1\cdot 43$$

$$\hat{\gamma}=\bar{X}-\hat{\delta}\ \bar{Y}=2\cdot 8-1\cdot 43\times 2=2\cdot 8-2\cdot 86=-0\cdot 06$$

Now the estimated regression line becomes

$$\mathbf{X=-0\cdot 06+1\cdot 43\ Y}$$

(c) when $X=7$

then $Y=0\cdot 079+\cdot 686\times 7=\mathbf{4\cdot 881}$

(d) when $Y=3$

then $X=-0\cdot 06+1\cdot 43\times 3=4\cdot 23$