Confidence Limits (Interval Estimation)

Estimating the interval or what is called confidence limits is the process of finding the true value of the parameter between an upper and lower confidence limit such that we are confident with an amount of $1-\propto$ that the true value of the parameter lies between these limits, i.e.:

$$Pr.(L \leq Parameter \leq U = 1-\propto \cdots 1$$

Where:-

L:- represents the lower limit

U:- represents the upper limit

The confidence limits for the parameter are placed within the following limits:-

$$\hat{\sigma} \pm t\,(\frac{\propto}{2}, n-2)\, S(\hat{\sigma})$$

Where:-

$\hat{\sigma}$:- The estimated parameter.

$S(\hat{\sigma})$:- The standard deviation of the estimated parameter $t\,(\frac{\propto}{2}, n-2)$ = the tabular t value.

It is worth noting that the confidence limits are the limits in which the null hypothesis is accepted, and the larger the sample size (n), the smaller the difference between the lower and upper limits (the confidence interval becomes smaller), i.e. it becomes narrow.

1. **Confidence Limits for β_1 (C.L. for β_1)**

   We have $Cal.t = \frac{\hat{\beta}_1-\beta_1}{S(\hat{\beta}_1)} \sim t\,(\frac{\propto}{2}, n-2)$

   The confidence limits β_1 are obtained from the t-statistic as follows:

$$P\left(-t_{\frac{\propto}{2}} \leq Cal.t \leq t_{\frac{\propto}{2}}\right) = 1-\propto$$

$$\Rightarrow P\left(-t_{\frac{\propto}{2}} \leq \frac{\hat{\beta}_1-\beta_1}{S(\hat{\beta}_1)} \leq t_{\frac{\propto}{2}}\right) = 1-\propto$$

$$\Rightarrow P\left(-t_{\frac{\propto}{2}} S(\hat{\beta}_1) \leq \hat{\beta}_1-\beta_1 \leq t_{\frac{\propto}{2}} S(\hat{\beta}_1)\right) = 1-\propto$$

$$\Rightarrow P\left(-\hat{\beta}_1 - t_{\frac{\propto}{2}} S(\hat{\beta}_1) \leq -\beta_1 \leq -\hat{\beta}_1 + t_{\frac{\propto}{2}} S(\hat{\beta}_1)\right) = 1-\propto$$

$$\Rightarrow P\left(\hat{\beta}_1 + t_{\frac{\propto}{2}} S(\hat{\beta}_1) \geq \beta_1 \geq \hat{\beta}_1 - t_{\frac{\propto}{2}} S(\hat{\beta}_1)\right) = 1-\propto$$

$$\Rightarrow P\left(\hat{\beta}_1 + t_{\frac{\alpha}{2}} S(\hat{\beta}_1) \le \beta_1 \le \hat{\beta}_1 - t_{\frac{\alpha}{2}} S(\hat{\beta}_1)\right) = 1-\alpha$$

وبهذا نستنتج ان حدود الثقة الى β_1 هي :-

$$L = \hat{\beta}_1 - t_{(\frac{\alpha}{2}, n-2)} S(\hat{\beta}_1)$$

$$U = \hat{\beta}_1 + t_{(\frac{\alpha}{2}, n-2)} S(\hat{\beta}_1)$$

That is, the confidence limits for the slope of the population regression line (coefficient) are shown as follows:

$$\hat{\beta}_1 \pm t_{(\frac{\alpha}{2})} S(\hat{\beta}_1)$$

2. **Confidence Interval for β_0 (C.I. for β_0)**

$H_0 : \beta_0 = \beta_{00}$
$H_1 : \beta_0 \ne \beta_{00}$

$$\frac{\hat{\beta}_0 - \beta_0}{S(\hat{\beta}_0)} \sim t\ (D.f., n-2)$$

In the same way in $\hat{\beta}_1$, the confidence limits to β_0 will be:

$$L = \hat{\beta}_0 - t_{\frac{\alpha}{2}} S(\hat{\beta}_0)$$

$$U = \hat{\beta}_0 + t_{\frac{\alpha}{2}} S(\hat{\beta}_0)$$

Thus, the confidence limits for the constant term (Intercept) in the simple linear regression model are written in another formula:

$$\hat{\beta}_0 \pm t_{\frac{\alpha}{2}} S(\hat{\beta}_0)$$

3. **C.L. for true mean value of Y at $x = x_0$**

On the same previous basis, we have the confidence limits for the average response at $x = x_0$ as follows:

$$\text{C.L. for } E\,Y_0 = \hat{Y}_0 \pm t_{\frac{\alpha}{2}} S(\hat{\beta}_0)$$

4. Confidence limits for population variance σ^2:-

$$\therefore \frac{(n-2)\text{mse}}{\sigma^2} \sim X^2_{n-2}$$

$H_0 : \sigma^2 = \sigma_0^2$
$H_1 : \sigma^2 \ne \sigma_0^2$

It is calculated as follows:

$$\therefore X^2_{1-\frac{\propto}{2},n-2} \leq \frac{(n-2)mse}{\sigma^2} \leq X^2_{\frac{\propto}{2},n-2}$$

The confidence limits for σ^2 will be:-

$$\frac{(n-2)mse}{X^2_{\frac{\propto}{2},n-2}} \leq \sigma^2 \leq \frac{(n-2)mse}{X^2_{1-\frac{\propto}{2},n-2}}$$

مثال :- اذا كانت لديك المعلومات التالية

$$\sum X_i = 491 , \sum X_i^2 = 26157 , \sum Y_i = 1410 , \sum Y_i^2 = 202094 , n = 10 , \quad \overline{Y} = 141 , \overline{X} = 49 - i$$

Required: 1. Test the hypothesis that$\beta_0 = 0$ at a significance level of 0.05 2. Test the hypothesis Ho :$\beta_1 = 0$ against H1 :$\beta_1 \neq 0$ and also β_0 =100 against $\beta_0 = 0 \neq 100$ at a probability level of 0.05 3. Find the confidence limits β_0, $\beta_1$4. Confidence limits for the mean response $\hat{Y}_0$ at point $X_0 = 44$ 5. Find the confidence limits for σ^2

Solution :-

$H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

$$Cal.\, t = \frac{\hat{\beta}_0 - \beta_{00}}{S(\hat{\beta}_0)}$$

$$\hat{Y}_i = \hat{\beta}_0 + X_i$$

$$S_{xy} = \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}$$

$$= 71566 - \frac{(491)(1410)}{10} = 2335$$

$$S_{xx} = \sum (X_i)^2 - \frac{(\sum X_i)^2}{n}$$

$$= 26157 - \frac{(491)^2}{10} = 2048.9$$

$$S_{yy} = \sum (Y_i)^2 - \frac{(\sum Y_i)^2}{n}$$

$$= 202094 - \frac{(1410)^2}{10} = 3284$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{2335}{2048.9} = 1.1369$$

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{X}$$

$$= 141 - (1.1396)(49.1) = 85.0439$$

$$\hat{Y}_i = 85.0439 + 1.1396(X_i)$$

$$S^2(\hat{\beta}_1) = \frac{\delta^2}{S_{xx}}$$

$$\delta^2 = \frac{S_{yy} - \hat{\beta}_1^2\, S_{xx}}{n-2} = Mse$$

$$= \frac{3.284 - (1.1396)^2(2048.9)}{10-2} = 77.9$$

$$S^2(\hat{\beta}_1) = \frac{77.9}{2048.9} = 0.038$$

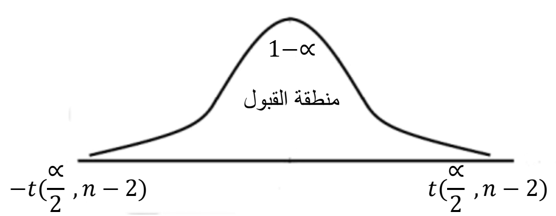
$$\Longrightarrow S(\hat{\beta}_1) = 0.1949$$

$$S^2(\hat{\beta}_0) = \hat{\delta}^2\left[\frac{1}{n} + \frac{\overline{X}^2}{S_{xx}}\right]$$

$$= 77.9\left[\frac{1}{10} + \frac{(49.1)^2}{2048.9}\right] = 99.45$$

$$\Longrightarrow S(\hat{\beta}_0) = 9.97$$

$$\text{Cal. t} = \frac{85.0439-0}{9.972} = 8.528$$



The calculated t value, which is equal to (8.528), is greater than the positive tabular t value, which is equal to (2.306). Therefore, the null hypothesis is rejected and the alternative hypothesis is accepted, i.e. $\beta_0 \neq 0$ at a probability level of 0.05. This means that there is importance to the intercept in calculating the expected value of Y and that the regression line does not pass through the origin.

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

$$\text{Cal. t} = \frac{\hat{\beta}_1 - \beta_{10}}{S(\hat{\beta}_1)}$$

$$= \frac{1.1396 - 0}{0.19494} = 5.58$$

The calculated t value, which is equal to (5.85), is greater than the positive tabular t value, which is equal to (2.306). Therefore, the null hypothesis is rejected and the alternative hypothesis is accepted, i.e. $\beta_0 \neq 0$. This means that X is important in explaining the changes that occur in Y. This means that the relationship between X and Y is represented by a straight line.

$$H_0 : \beta_1 = 100$$

$$H_1 : \beta_1 \neq 100$$

$$\text{Cal. t} = \frac{\hat{\beta}_1 - \beta_{10}}{S(\hat{\beta}_1)}$$

$$= \frac{85.0439 - 100}{9.972} = -1.50$$

The absolute value of t, which is equal to (1.306), is less than the table value of t, which is equal to (2.306), so the null hypothesis is accepted, i.e. the value of β_0 = 100.

$$\hat{\beta}_1 - t_{\frac{\alpha}{2},n-2}\, S(\hat{\beta}_1) \leq \beta_1 \leq \hat{\beta}_1 + t_{\frac{\alpha}{2},n-2}\, S(\hat{\beta}_1)$$

$$1.396 - (2.306)(0.194) \leq \beta_1 \leq 1.396 + (2.306)(0.194)$$

$$0.6902 \leq \beta_1 \leq 1.5660$$

$$\hat{\beta}_0 - t_{\frac{\alpha}{2},n-2}\, S(\hat{\beta}_0) \leq \beta_0 \leq \hat{\beta}_0 + t_{\frac{\alpha}{2},n-2}\, S(\hat{\beta}_0)$$

$$85.0456 - (2.306)(9.97) \leq \beta_0 \leq 85.0456 + (2.306)(9.97)$$

$$62.05 \leq \beta_0 \leq 108.03$$