

Deviation of \hat{Y}_i from the true value Y_i

That is the mean-square error value of prediction (or estimation) $\tilde{Y} = Y - \hat{Y}$.

We want to evaluate $E((Y - \hat{Y})^2) = E(\tilde{Y}^2)$

$$E((Y - \hat{Y})^2) = E(Y - \bar{Y} + \bar{Y} - \hat{Y})^2$$
$$= E(Y - \bar{Y})^2 + E(\hat{Y} - \bar{Y})^2$$

$$- 2 E(Y - \bar{Y})(\hat{Y} - \bar{Y})$$

(equal to zero b/c the two factors are independent of each other.)

$$= E(Y - \bar{Y})^2 + E(\hat{Y} - \bar{Y})^2$$

$$= \text{var}(Y) + \text{var}(\hat{Y})$$

$$= \sigma^2 + \text{var}(\hat{\beta}_0 + \hat{\beta}_1 X)$$

$$= \sigma^2 + \text{var}(\hat{\beta}_0) + \text{var}(\hat{\beta}_1 X)$$

$$+ 2 \text{Cov}(\hat{\beta}_0, X \hat{\beta}_1)$$

$$= \sigma^2 + \frac{\sigma^2}{n} + \frac{\sigma^2 \bar{X}^2}{S_X^2} + \frac{X^2 \sigma^2}{S_X^2} + 2X \left(-\frac{\sigma^2 \bar{X}}{S_X^2} \right)$$

$$\sigma_{\tilde{Y}}^2 = \sigma^2 + \frac{\sigma^2}{n} + \frac{\sigma^2}{S_X^2} (X - \bar{X})^2$$

$$MSE = \sigma^2 \left(1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{S_x^2} \right) = \sigma_{\hat{Y}}^2$$

for the given value of $x = x_0$.

$$\sigma_{\hat{Y}} = \sigma \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{S_x^2}}$$

Under the assumption of normality of error term

$$Z = \frac{Y(x=x_0) - \hat{Y}(x=x_0)}{\sigma \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_x^2}}} \sim N(0, 1)$$

If σ^2 is known (which is not possible in real life scenarios)

then C.I. is obtained by Z-statistic.

If σ^2 is to be estimated, then we use the following unbiased formula:

$$S^2 = \frac{1}{n-2} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Then $(1-\alpha)100\%$ of C.I. on $E(Y - \hat{Y})^2$ is obtained as (for the given value of $x_i = x_0$)

$$\hat{Y}(x_0) \pm t_{\frac{\alpha}{2}, n-2} \cdot S \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_x^2}}$$

where $S_x^2 = \sum_{i=1}^n (x_i - \bar{x})^2$ and $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$.