MLR.1 (Linear in parameters): $y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + u$.

MLR.2 (Random sampling)

MLR.3 (No perfect collinearity)

MLR.4: (Zero conditional mean) $E(u | x_1, \ldots, x_k) = 0$

MLR.5 (Homoskedasticity) $Var(u | x_1, \ldots, x_k) = \sigma^2$

Theorem 3.2. $Var(\hat{\beta}_j) = \dfrac{\sigma^2}{SST_j(1-R_j^2)}$ R-squared from a regression of x_j on all other independent variables

SST_j: total sample variation in x_j

example: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$.

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{SST_j(1-R_1^2)}.$$

R_1^2 is the R-squared from the simple regression of x_1 on x_2.

$$x_1 = \delta x_2 + e \qquad R_1^2 = \frac{\sum(\hat{\delta} x_{2,i} - \bar{x}_1)^2}{\sum(x_{1,i} - \bar{x}_1)^2}$$

$R_1^2 \uparrow \Longrightarrow Var(\hat{\beta}_1) \uparrow$

High degree of linear relationship between x_1 and x_2 can lead to large variances for OLS slope estimators.

$Var(\hat{\beta}_j) \to \infty$ as $R_j^2 \to 1$, High relation between two

$R_j^2 = 1$ is ruled out by MLR.3.

or more independent variables is called Multicollinearity.

variance inflation factor (VIF)

$$VIF_j = \frac{1}{1-R_j^2} \in [1, +\infty) , \quad Var(\beta_j) = \frac{\sigma^2}{SST_j} \cdot VIF_j$$

$$Var(\tilde{\beta}_1) = Var\left(\frac{\sum (x_{1,i} - \bar{x}_1)(y_{1,i} - \bar{y})}{\sum (x_{1,i} - \bar{x})^2} \right)$$

$$= Var\left(\frac{\sum (x_{1,i} - \bar{x}_1)(\beta_1(x_{1,i} - \bar{x}_1) + \beta_2(x_{2,i} - \bar{x}_2) + u_i)}{\sum (x_{1,i} - \bar{x})^2} \right.$$

$$= Var\left(\beta_1 + \frac{\beta_2 \sum (x_{1,i} - \bar{x}_1)(x_{2,i} - \bar{x}_2)}{\sum (x_{1,i} - \bar{x})^2} + u_i \frac{\sum (x_{1,i} - \bar{x})}{\sum (x_{1,i} - \bar{x})^2} \right.$$

$$= Var\left(u_i \frac{\sum (x_{1,i} - \bar{x})}{\sum (x_{1,i} - \bar{x})^2} \right)$$

$$= \frac{\sigma^2}{SST_1}$$

$$Var(\hat{\beta}_j) \geq Var(\tilde{\beta}_j)$$

$\xi = 0.$

x_j is uncorrelated with other independent variables.

$R_j^2 = 0$

$$\Longrightarrow Var(\hat{\beta}_j) = Var(\tilde{\beta}_j), \quad \hat{\beta}_j = \tilde{\beta}_j$$

$$SSR = \sum_{i=1}^{n} \hat{u}_i^2$$

$$\hat{\sigma}^2 = \frac{SSR}{\boxed{n-k-1}}$$ degrees of freedom (df)

k: # of regressors.

$= n - (k+1)$

number of observations; number of estimated parameters.

under MLR.1 to MLR.5, $E(\hat{\sigma}^2) = \sigma^2$

$$\widehat{Var}(\hat{\beta}_j) = \frac{\hat{\sigma}^2}{SST_j(1-R_j^2)}$$

$\hat{\sigma}^2$ is an unbiased estimator of σ^2 but $\hat{\sigma}$ isn't.

$$sd(\hat{\beta}_j) = \sqrt{Var(\hat{\beta}_j)} = \sqrt{\frac{\sigma^2}{SST_j(1-R_j^2)}}$$

$$se(\hat{\beta}_j) = \sqrt{\widehat{Var}(\hat{\beta}_j)} = \sqrt{\frac{\hat{\sigma}^2}{SST_j(1-R_j^2)}}$$

Gauss-Markov Theorem:

under MLR.1 through MLR.5, $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ are the best linear unbiased estimators (BLUE) of $\beta_1, \beta_2, \dots, \beta_k$ respectively.

and asymptotically efficient

best: having smallest variance.

linear: if and only if $\hat{\beta}_j = \sum_{i=1}^{n} w_{ij} y_i$

i.e.

if $\tilde{\beta}_j = \sum_{i=1}^{n} w_{ij} y_i$ and $E(\tilde{\beta}_j) = \beta_j$, $Var(\hat{\beta}_j) \le Var(\tilde{\beta}_j)$