

ANOVA (Analysis of Variance) table

In general:

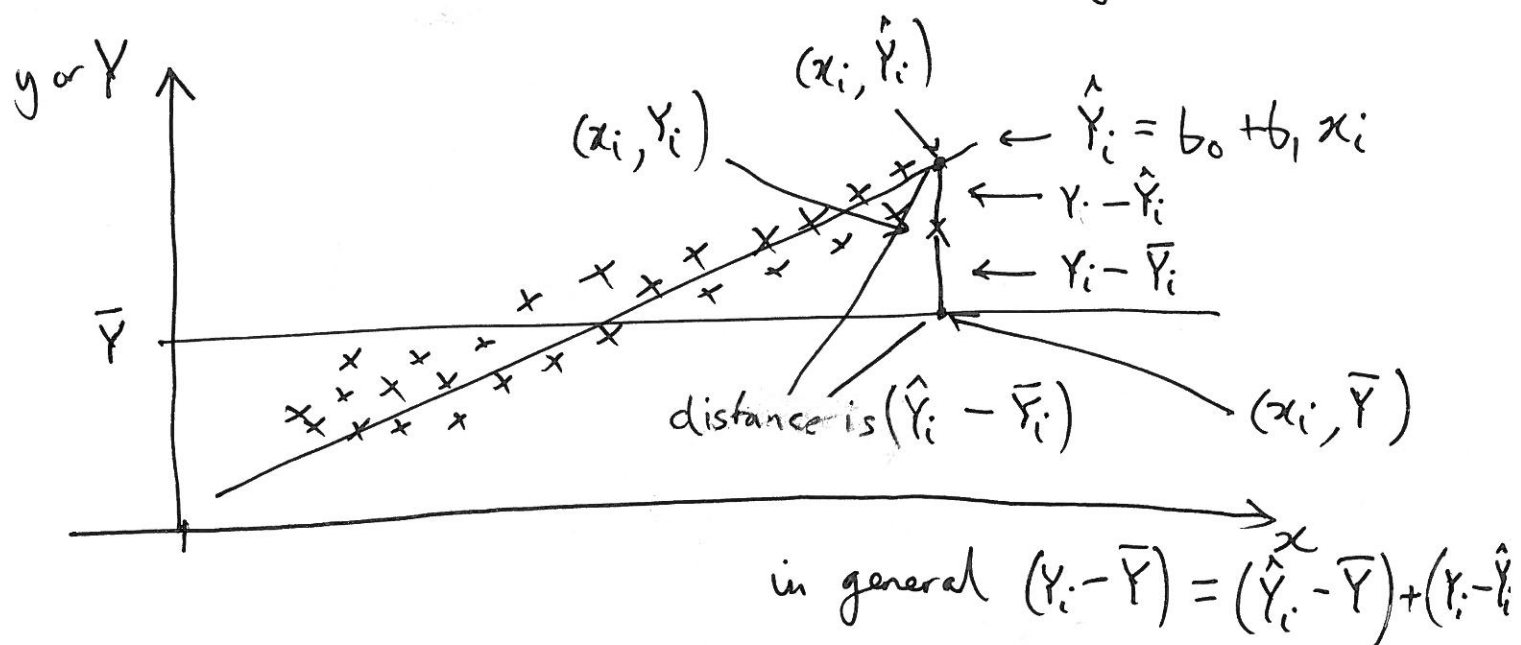
Source	df ← degrees of freedom	SS ← sum of squares	MS ← mean squares	F ← Fisher test statistic	p-value
Regression (Model)	$k = p - 1$	$SS_{\text{Regression}}$	$MS_{\text{reg}} = \frac{SS_{\text{reg}}}{k}$	$\frac{MS_{\text{reg}}}{MS_{\text{error}}}$
Residuals (Errors)	$n - p$	SS_{Errors}	$MS_{\text{error}} = \frac{SS_{\text{errors}}}{n - p}$		
Total	$n - 1$	SS_{Total}			

$p = \# \text{ parameters in the model } \{ \beta_0, \beta_1, \beta_2, \dots, \beta_k \}$

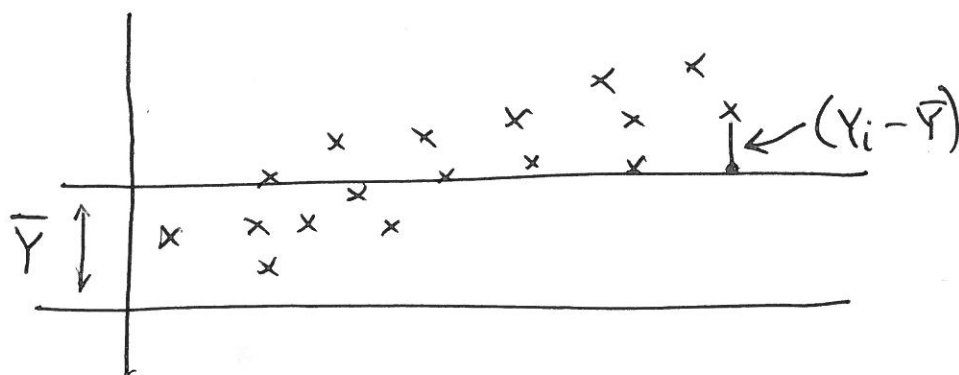
$k = \# \text{ variables or } \# \text{ slope coefficients (excl. } \beta_0) \{ \beta_1, \dots, \beta_k \}$

For simple linear regression ($p = 2$; β_0, β_1 & $k = 1$; β_1)

Source	df	SS	MS	F
Regression	1	$\sum (\hat{Y}_i - \bar{Y})^2$	$\frac{\sum (\hat{Y}_i - \bar{Y})^2}{1}$ ← $\div \text{ by } 1$	
Residuals	$n - 2$	$\sum (Y_i - \hat{Y}_i)^2$	$\frac{\sum (Y_i - \hat{Y}_i)^2}{n - 2}$	
Total	$n - 1$	$\sum (Y_i - \bar{Y})^2$ ← $\Delta_y^2 = \frac{1}{n-1} \sum (Y_i - \bar{Y})^2$		

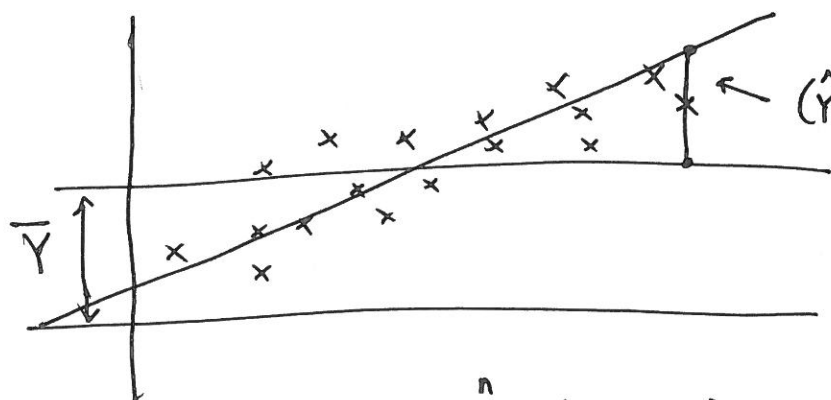


$$\text{Total SS} = \sum_{i=1}^n (Y_i - \bar{Y})^2 \quad \text{total squared distance between data \& a mean (or null model } Y_i = \beta_0 + \epsilon_i \text{)}$$



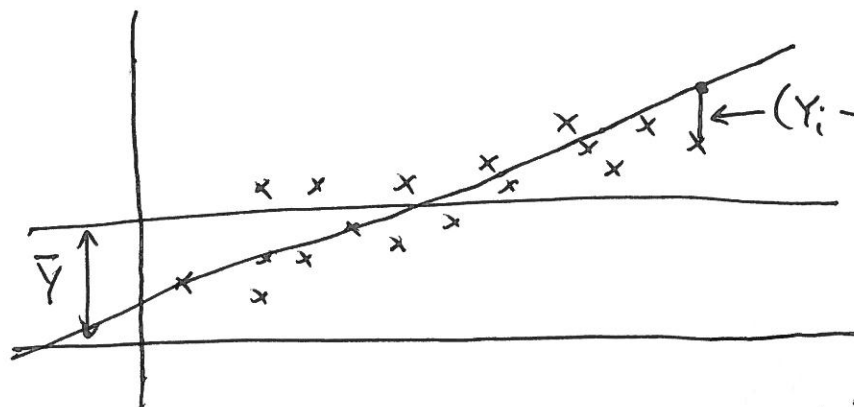
is a measure of the total variation in the observed Y data

$$\text{Regression SS} = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 \quad \text{total squared distance between fitted values from model } Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \text{ \& mean (null model)}$$



is a measure of the variation "explained" by the model involving x

$$\text{Error SS} = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad \text{total squared distance between data \& the fitted values}$$



is a measure of the "unexplained" variation

$$\sum (Y_i - \bar{Y})^2 = \sum (\hat{Y}_i - \bar{Y}_i)^2 + \sum (Y_i - \hat{Y}_i)^2 + \left(\begin{array}{l} \text{a cross} \\ \text{product} \\ \text{term} = 0 \end{array} \right)$$

(see Q 4 (a) of Tutorial 1) →

$$SS_{\text{Total}} = SS_{\text{Regression}} + SS_{\text{Errors}}$$