STAT6038 week 3 lecture 8

Rui Qiu

2017-03-09

ANOVA (Analysis of Variance Table In general,

Source	df	SS	MS	F	p-value
Regression (Model)	$k = p - 1$	$SS_{Regression}$	$MS_{Reg} = \frac{SS_{Reg}}{k}$	$\frac{MS_{Reg}}{MS_{Errors}}$	
Residuals (Errors)	$n - 1$	SS_{Errors}	$MS_{Errors} = \frac{SS_{Errors}}{n-p}$		
Total	$n - 1$	SS_{Total}			

Table 1: ANOVA table in details

$p =$ number of parameters in the model $\{\beta_0, \beta_1, \ldots, \beta_k\}$

$k =$ number of variables or number of slope coefficients (excluding β_0) $\{\beta_1, \beta_2, \ldots, \beta_k\}$.

For simple linear regression (SLR), ($p = 2, \beta_0, \beta_1$ and $k = 1, \beta_1$).

F is Fisher test statistics.

For SLR,

Source	df	SS	MS	F	p-value
Regression (Model)	1	$\sum(\hat{Y}_i - \bar{Y})^2$	$\sum(\hat{Y}_i - \bar{Y})^2$	MSR/MSE	
Residuals (Errors)	$n - 2$	$\sum(Y_i - \hat{Y}_i)^2$	$\frac{\sum(Y_i - \hat{Y}_i)^2}{n-2}$		
Total	$n - 1$	$\sum(Y_i - \bar{Y})^2$			

Table 2: ANOVA table for SLR

Note that we use s_y^2 to estimate SS_{Total}

$$s_y^2 = \frac{1}{n-1}\sum(Y_i - \bar{Y})^2.$$