- Global F-test (ANOVA)
- Generalized/partial F-test
- t-test for indiv. slopes

Global F-test:

$$\boxed{H_0: \beta_1 = \beta_2 = \beta_3 = \cdots = \beta_{p-1} = 0}$$

H_1: at least one $\beta_j \neq 0$

$\Updownarrow$

$$H_0: Y_i = \beta_0 + \underbrace{\beta_1 X_{1i} + \cdots + \beta_{p-1} X_{p-1,i}}_{\beta_1 = 0,\ \beta_{p-1} = 0} + \varepsilon_i$$

$$\boxed{Y_i = \beta_0 + \varepsilon_i}$$ "Reduced model" (intercept only)

$$H_1: \boxed{Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_{p-1} X_{p-1,i} + \varepsilon_i}$$

"Full Model"

So we can construct an F-stat as:

$$F = \frac{(SSE_R - SSE_F) / (df_{SSE_R} - df_{SSE_F})}{SSE_F / df_{SSE_F}} \sim F_{df_1, df_2}$$

$df_1 = df_{SSE_R} - df_{SSE_F}$

$df_2 = df_{SSE_F}$

Global

$(= p-1)$

$= (n-p)$

Full Model:

source	SS	df
Regression	*	$p-1$
Error	*	$n-p$

$p = \#$ of parameters in full model

Decision:

If $F > F^{*}_{df_1, df_2, 1-\alpha}$ then

reject H_0 & conclude there exists

at least one significant predictor

among $X_1, \dots, X_{p-1}$ for predicting Y.

Partial F-test:

Idea: test only a subset of predictors,

say $X_r, \dots, X_{p-1}$

H_0: $\beta_{\underline{r}} = \beta_{r+1} = \cdots = \beta_{p-1} = 0$

H_1: at least one of $\{\beta_j\}_{j=r}^{p-1}$ is not zero.

$r = \#$ of parameters in reduced model

$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_{r-1} X_{r-1\,i} + \varepsilon_i$

$\Updownarrow$

Reduced H_0: $Y_i = \beta_0 + \sum_{j=1}^{r-1} \beta_j x_{ji} + \varepsilon_i \Leftrightarrow Y = X\beta^* + \varepsilon$

where $\beta^* = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{r-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

Full H_1: $Y_i = \beta_0 + \sum_{j=1}^{p-1} \beta_j x_{ji} + \varepsilon_i$

$\Leftrightarrow Y = X\beta + \varepsilon$

$$F = \frac{(SSE_R - SSE_F) / (df_{SSE_R} - df_{SSE_F})}{SSE_F / df_{SSE_F}} \sim F_{df_1, df_2}$$

$$df_1 = df_{SSE_R} - df_{SSE_F}$$

$$(n-r) - (n-p) = p-r$$

$$df_2 = n-p$$

Decision:

if $F > F^*_{df_1, df_2, 1-\alpha}$ then at least one significant predictor exists among $\{X_r, \dots, X_{p-1}\}$.

Ex:

$H_0: Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$ $\Rightarrow r=3$

$H_1: Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_5 X_{5i} + \varepsilon_i$ $\Rightarrow p=6$

num df $= p-r = 6-3 = 3$

Specific Predictor F-test

$H_0: \beta_{p-1} = 0$

vs

$H_1: \beta_{p-1} \neq 0$

same exact story, just let $r = p-2$ //

How can we make a t-test to check if a specific predictor is useful or not?

idea: $\dfrac{\hat{\beta}_1 - 0}{\hat{SE}(\hat{\beta}_1)} \xrightarrow{\text{in MLR}} \dfrac{\hat{\beta}_j - 0}{\hat{SE}(\hat{\beta}_j)}$ $\quad j = 1, \dots, p$

$j+1$ st diag. entry

$$SE(\hat{\beta}_j) = \sqrt{Var(\hat{\beta}_j)} = \sqrt{\sigma^2 (X^T X)^{-1}_{(j+1, j+1)}}$$

Recall

$$\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1}) \rightarrow \begin{bmatrix} Var(\hat{\beta}_0) & & & \\ & Var(\hat{\beta}_1) & & \\ & & \ddots & \\ & & & Var(\hat{\beta}_j) \\ & & & & \ddots \end{bmatrix}$$

So we can construct a t-stat:

$$t = \frac{\hat{\beta}_j - 0}{\hat{SE}(\hat{\beta}_j)} = \frac{\hat{\beta}_j}{\hat{\sigma}\sqrt{(X^TX)^{-1}_{j+1,j+1}}}$$

If $|t| > t^*_{\underline{n-p}, \boxed{\underline{1-\alpha/2}}} \Rightarrow$ reject H_0

Interpretation:

we have evidence that suggests X_j is a significant predictor of Y <u>given</u> the other predictors in the model.

This t-test is totally equivalent to the specific predictor's partial F-test!!

$(t^2 = F.)$

Adjusting R^2

Recall R^2 comes from the ANOVA breakdown:

$$R^2 = 1 - \frac{SSE}{SST}$$

Fact: When we compare models of different sizes, R^2 will always be biggest for the largest model.

It doesn't take into account the "cost" of losing degrees of freedom.

To handle this, we can use

Adjusted R^2:

$$R_a^2 = 1 - \frac{SSE/n-p}{SST/n-1} = 1 - \frac{MSE}{MST}$$

Advantage: R_a^2 doesn't always increase as more predictors get added.

$\Rightarrow$ R_a^2 is useful as a model selection criterion