9/15/25

Agenda

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

Global F-test:

Test for checking if <u>any</u> predictor is significant.

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

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

$\Updownarrow$

$df_{SSE_R} = n-1$

$H_0: \quad Y_i = \beta_0 + \varepsilon_i$ (reduced model)

$H_1: \quad Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_{p-1} X_{(p-1)i} + \varepsilon_i$

(full model)

$df_{SSE_F} = n-p$

We can construct an F-stat based on the full/reduced models as:

"SSY" - "SSE" ← from SLR

$$F = \frac{(SSE_R - SSE_F) / (df_{SSE_R} - df_{SSE_F})}{SSE_F / (df_{SSE_F})} \overset{H_0}{\sim} F_{df_1, df_2}$$

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

$df_2 = df_{SSE_F}$

ANOVA Table (Global)

Source	SS	df
Regression	(num) *	$(n-1)-(n-p) = p-1$
Error	(den) *	$n-p$

p = # of parameters in the full model

Decision:

If $F > F^{*}_{df_1, df_2}(1-\alpha)$ then we reject H_0 & say we have evidence that at least one of $X_1, \dots, X_{p-1}$ is a significant pred. of Y.

Partial F-test

Idea: test only a subset of predictors,

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

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

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

$$\Updownarrow$$

$df_{SSE_R} = n-r$

$H_0: Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_{r-1} X_{(r-1)i} + \varepsilon_i$

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

$df_{SSE_F} = n-p$

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

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

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

$df_2 = df_{SSE_F} = n-p$

Decision: If $F > F^*_{df_1, df_2}(1-\alpha)$ we can say we have evidence that at least one of $\{X_r, \ldots, X_{p-1}\}$ is a sig. pred. of Y.

Specific Predictor F-test:

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

vs.

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

Same exact story except $r = p-2$

Source	SS	df
Res	$SSE_R - SSE_F$	1
Error	SSE_F	$n-p$

$$F = \frac{(SSE_R - SSE_F)/1}{SSE_F/(n-p)} \sim F_{1, n-p}$$

$\Updownarrow$ there is an associated t-stat.

Idea: $t = \dfrac{\hat{\beta}_{p-1} - 0}{\widehat{SE}(\hat{\beta}_{p-1})} =$

$\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1})$

$Var(\hat{\beta}_{p-1}) = \sigma^2 [\underline{(X^T X)^{-1}}]_{p,p}$ ← p^{th} diag element of $(X^T X)^{-1}$

$\widehat{SE}(\hat{\beta}_{p-1}) = \hat{\sigma} \sqrt{(X^T X)^{-1}_{p,p}}$

With this t stat, we have the usual $t^2 = F$ ← is for the specific pred. F-test

<u>Interpretation</u>:

If $F > F^*_{1,n-p}(1-\alpha) \iff |t| > t^*_{n-p}(1-\alpha/2)$ then we have evidence that suggests X_{p-1} is a sig. pred. of Y, conditional on $\underbrace{X_1 \cdots X_{p-2}}_{\text{all other predictors}}$ being in the model.

Recall R^2 comes from the ANOVA breakdown

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

(*) When we add more predictors into the model R^2 will always go up or stay the same...

So R^2 on its own is not useful for model selection — it always picks the full model.

It does not take into account the "cost" of estimating unnecessary predictors & the associated reduction in degrees of freedom.

To handle this, we can adjust R^2:

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

Advantage of R_a^2 is:

R_a^2 doesn't always pick the biggest model.

$\Rightarrow$ It can be used for model selection.

Why does adding an (perfectly) uncorrelated pred. leave SSE unchanged?

<u>EX</u>: X_1 is a sig. pred of Y
X_2 is totally uncorrelated w/ Y.

$df = n-2$ Red: $Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i$ $\leftarrow$

$df = n-3$ <u>Full</u>: $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$ $\leftarrow$

SSE is the same for each case