F - Test

(test <u>multiple</u> hypotheses about $\beta_1, \beta_2, \cdots \beta_k$)

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + u.$$

unrestricted model.

$H_0: \beta_{k-q+1} = 0 \cdots\cdots \beta_k = 0$ (three exclusion restrictions).

H_1: H_0 is not true.

$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_{k-q} x_{k-q} + u$ (restricted model for testing)

F statistic $\quad F \equiv \dfrac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)} \sim F_{q,\, n-k-1}.$

q : numerator degrees of freedom $= df_r - df_{ur}$

df: # observations − # estimated parameters. $= [n-(k-q+1)] - (n-(k+1))$
$= q$

SSR_r is the sum of squared residuals from the restricted model.

SSR_{ur} is from the unrestricted model.

$SSR_r \geq SSR_{ur} \Rightarrow F \geq 0.$

Choose significance level 5%

Let c be the 95^{th} percentile in the $F_{q,n-k-1}$
critical value depends on q, $n-k-1$.

$F > c \rightarrow$ reject H_0

$x_{k-q+1}, \cdots\cdots, x_k$ are jointly statistically significant.
(can't say which has partial effect on y)

if H_0 is not rejected, the variables are jointly insignificant.

take the null to be $H_0: \beta_k = 0$ and $q = 1$

t^2_{n-k-1} has an $F_{1,n-k-1}$ distribution, the two approaches lead to exactly the same outcome.

t statistic is more flexible for testing a single hypothesis, and more easy to obtain.

R-squared Form of the F-statistic.

$SSR_r = SST(1 - R_r^2)$, $SSR_{ur} = SST(1 - R_{ur}^2)$.

$\sum (y_i - \bar{y})^2$

$$\Rightarrow F = \frac{(R_{ur}^2 - R_r^2)/q}{(1 - R_{ur}^2)/(n-k-1)}$$

p-value for F Tests

$p\text{-value} = P(\mathcal{F} > F)$.

F Statistic for overall significance of a Regression.

model $y = \beta_0 + \beta_1 x_1 + \cdots \beta_k x_k + u.$

$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$

$\Rightarrow$ restricted model: $y = \beta_0 + u.$

$$F = \frac{R^2 / k}{(1-R^2)/(n-k-1)}$$