Hypothesis testing in finite and large samples

Econometrics for minor Finance, Lecture 5

Tunga Kantarcı, Fall 2025

Hypothesis testing in finite samples: Multiple restrictions

The F statistic for testing q linear restrictions is

$$F = \frac{\dfrac{SSR_r - SSR_{ur}}{q}}{\dfrac{SSR_{ur}}{n-k-1}}$$

Hypothesis testing in finite samples: Multiple restrictions

SSR_r : Residual sum of squares from the restricted model, with q restrictions imposed

SSR_{ur} : Residual sum of squares from the unrestricted model. Full regression

q : Number of linear restrictions being tested

n : Sample size

k : Number of slope parameters in the unrestricted model

1 : For the constant term

Hypothesis testing in finite samples: Multiple restrictions

Restricted model with q restrictions imposed on β:

$$y = X_r \beta_r + u$$

The SSR in this model:

$$SSR_r = \sum_{i=1}^{n} (y_i - \hat{y}_{r,i})^2$$

Unrestricted model with all k regressors:

$$y = X\beta + u$$

The SSR in this model:

$$SSR_{ur} = \sum_{i=1}^{n} (y_i - \hat{y}_{ur,i})^2$$

Hypothesis testing in finite samples: Multiple restrictions

$$F \sim F\left(q,\, n-k-1\right)$$

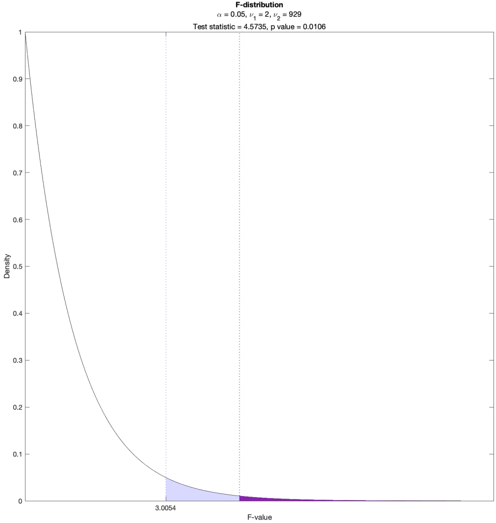
If the errors u are i.i.d. normal with constant variance, then in finite samples the statistic has an exact $F(q,\, n-k-1)$ distribution. Hence, in finite samples with normal errors we can compare the F statistic with critical values from the F distribution.

Hypothesis testing in large samples: Multiple restrictions

$$F \overset{a}{\sim} F\,(q,\; n-k-1)$$

If the errors u are not assumed to be normal but if they are i.i.d. with constant variance, and if the sample size is large, the statistic is asymptotically distributed as $F(q,\, n-k-1)$. Hence, in large samples we can compare the F statistic with critical values from the F distribution.

Hypothesis testing in large samples: Multiple restrictions



Hypothesis testing in finite samples: Multiple restrictions: Alternative form

$$F = \frac{\frac{SSR_r - SSR_{ur}}{q}}{\frac{SSR_{ur}}{n-k-1}}$$

If the restricted model is intercept-only, then

$$SSR_r = SST$$

and since

$$SSE_{ur} = SST - SSR_{ur}$$

the numerator becomes

$$\frac{SSR_r - SSR_{ur}}{q} = \frac{SST - SSR_{ur}}{q} = \frac{SSE_{ur}}{q}$$

Hypothesis testing in finite samples: Multiple restrictions: Alternative form

So the alternative form of the test is

$$F = \frac{\frac{SSE_{ur}}{q}}{\frac{SSR_{ur}}{n-k-1}}$$

Hypothesis testing in finite samples: Multiple restrictions: Alternative form

$$SST = SSE + SSR$$

SST, total sum of squares, measures total variation in y:

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

SSE, explained sum of squares, is the part explained by the model:

$$SSE = \sum_{i=1}^{n} (\hat{y}_i - \bar{\hat{y}})^2$$

SSR, residual sum of squares, is the leftover unexplained variation:

$$SSR = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$