Partialing Out and F-tests

Partialling-out in OLS

OLS estimation of this equation gives estimates $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)$.

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \epsilon_t$$

Interpretations of $\hat{\beta}_1$

- $\frac{\partial y_t}{\partial x_{1t}} = \hat{\beta}_1$.
- The change in y predicted from a one-unit change in x_1, *holding all other included variables constant.*
- The change in y predicted from variation in x_1 that is *independent of all other included variables.*
- Controlling for x_2 and x_3, $\hat{\beta}_1$ measures effects of changes in x_1 that are independent of x_2 and x_3.

Partialing Out

If we run this regression:

$$x_{1t} = \hat{\delta}_0 + \hat{\delta}_2 x_{2t} + \hat{\delta}_3 x_{3t} + \hat{r}_{1t}$$

Then $\hat{r}_{1t}$ is the remaining variation in x_1 that is independent of x_2 and x_3. Suppose we run a simple regression of y on $\hat{r}_{1t}$:

$$y_t = \hat{\alpha} + \hat{\beta}_1 \hat{r}_{1t} + e_t$$

This $\hat{\beta}_1$ is *numerically identical* to the one from the full regression.

An important consideration

- What if x_1, x_2, x_3 are all highly correlated?
- Then $\hat{r}_{1t}$ won't have a lot of variation left.
- Most of the variation in x_1 is captured in x_2 and x_3.
- Remember - more variation in the right hand side variables means
  - smaller standard errors
  - more precise estimates of the coefficients
  - bigger t-statistics
  - smaller confidence intervals.

Takeaway

- With little variation in $\hat{r}_{1t}$, $\hat{\beta}_1$ will have a large standard error (will not be precisely estimated).
- Even if x_1, x_2, x_3 all belong in the model, individual t-statistics may be small if they are highly correlated.
  - Multicollinearity, not perfect multicollinearity.
- This is one motivation for using an F-test - a joint test of several coefficients at once.

F-tests

Suppose we want to run the regression

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \epsilon_t$$

and test the joint hypothesis

$$\begin{aligned} H_0 &: \beta_2 = 4,\ \beta_3 = -2 \\ H_a &: \beta_2 \neq 4,\ \beta_3 \neq -2 \end{aligned}$$

F-tests

One idea:

- variance-weighted squared distance of $(\widehat{\beta}_2, \widehat{\beta}_3)$ from (4,-2)

Another idea:

- compare the variance of ϵ_t when $(\widehat{\beta}_2, \widehat{\beta}_3)$ are forced to be (4,-2) vs. when they are freely estimated.

These turn out to be the same thing: an F-test.

F-test

$$\begin{aligned} \text{full:} \quad & y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \epsilon_{ft} \\ \text{restricted:} \quad & y_t = \beta_0 + \beta_1 x_{1t} + 4x_{2t} - 2x_{3t} + \epsilon_{rt} \end{aligned}$$

$$F = \frac{\frac{1}{q}\left(\sum_{t=1}^{T} \epsilon_{rt}^2 - \sum_{t=1}^{T} \epsilon_{ft}^2\right)}{\frac{1}{T-k}\sum_{t=1}^{T} \epsilon_{ft}^2} = \frac{(SSR_r - SSR_f)/q}{SSR_f/(T-k)}$$

* q = 2 restrictions, k = 4 total coefficients

Distribution for F-test

As sums of squared Normals:

- the numerator and the denominator both have χ^2 distributions
- each divided by the degrees of freedom
- this has a special distribution called the "F distribution":

$$F \sim \frac{\chi^2/q}{\chi^2/(T-k)} \sim \mathcal{F}_{(q,T-k)}$$

Linear Algebra

The F-statistic is algebraically equivalent to variance-weighted squared distance of $(\widehat{\beta}_2, \widehat{\beta}_3)$ from the joint hypothesis (4,-2)

$$\text{hypothesis: } \underbrace{\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{R}} \underbrace{\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}}_{\mathbf{b}} = \underbrace{\begin{bmatrix} 4 \\ -2 \end{bmatrix}}_{\mathbf{r}}$$

$$H_0 : \underset{q \times K}{\mathbf{R}} \cdot \underset{K \times 1}{\mathbf{b}} = \underset{q \times 1}{\mathbf{r}}$$

$$H_a : \mathbf{R} \cdot \mathbf{b} \neq \mathbf{r}$$

Weighted distance

How to get a single number (test statistic) that

- captures distance between multiple coefficient estimates and their hypothesized values
- has a distribution against which we can compare that number/distance/statistic?

One idea: squared distance, weighted by variance.

Weighted distance

Squared distance:

$$\underbrace{(\mathbf{R}\cdot\mathbf{b}-\mathbf{r})^T}_{1\times q}\cdot\underbrace{(\mathbf{R}\cdot\mathbf{b}-\mathbf{r})}_{q\times 1}$$

Variance of the distance:

$$Var\,(\mathbf{R}\cdot\mathbf{b}-\mathbf{r}) = \underset{q\times k}{\mathbf{R}}\cdot\underset{k\times k}{Var(\mathbf{b})}\cdot\underset{k\times q}{\mathbf{R}^T}$$

Aside on variance

General rules for variance when X is a random variable but a and b are constants:

$$Var(aX + b) = a^2 Var(X) = a \cdot Var(X) \cdot a$$

$Var(\mathbf{R} \cdot \mathbf{b} - \mathbf{r}) = \mathbf{R} Var(\mathbf{b}) \mathbf{R}^T$ is the matrix version of this.

Variance-weighted squared distance

Variance-weighted, squared distance:

$$F = \frac{1}{q}(\mathbf{R} \cdot \underset{1\times q}{\mathbf{b}} - \mathbf{r})^T \cdot \left(\underset{q\times k}{\mathbf{R}} \cdot \underset{k\times k}{Var(\mathbf{b})} \cdot \underset{k\times q}{\mathbf{R}^T} \right)^{-1} \cdot (\mathbf{R} \cdot \underset{q\times 1}{\mathbf{b}} - \mathbf{r})$$

Variance-weighted squared distance

This is equivalent to

$$F = \frac{1}{q}\left(\begin{bmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} - \begin{bmatrix} 4 \\ -2 \end{bmatrix}\right)^T \left(\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \underset{4\times 4}{Var(\mathbf{b})} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\right)^{-1} \left(\begin{bmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}\right.$$

Variance weighted squared distance

Which further reduces to

$$F = \frac{1}{q}\begin{pmatrix}\hat{\beta}_2 - 4 & \hat{\beta}_3 + 2\end{pmatrix}\begin{bmatrix} Var(\hat{\beta}_2) & 0 \\ 0 & Var(\hat{\beta}_3)\end{bmatrix}^{-1}\begin{pmatrix}\hat{\beta}_2 - 4 \\ \hat{\beta}_3 + 2\end{pmatrix}$$

A simplest possible example

Can we use this logic to test just one restriction?

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \epsilon_t$$
$$H_0 : \beta_1 = 7$$
$$H_a : \beta_1 \neq 7$$

A simplest possible example

In this case, $q = 1$:

$$F = \frac{1}{1}\frac{(\hat{\beta}_1 - 7)^2}{Var(\hat{\beta}_1)} = t^2$$
$$t = \frac{\hat{\beta}_1 - 7}{se(\hat{\beta}_1)}$$

In the special case where we just test one restriction, the F-stat is the square of the t-stat for that restriction.