The F Test

Caleb Leedy

19 January 2024

I. Computation

To compute an F test we can use Equation 1.

$$F = \frac{(SSE_{Reduced} - SSE_{Full})/(DFE_{Reduced} - DFE_{Full})}{SSE_{Full}/DFE_{Full}}. \quad (1)$$

Let's assess what this would mean to compare models `Parametric 1` and `Outcome Robust` from Table 1.

Table 1: *This table identifies the different constraints for each model type.*

Type	Constraints
Parametric 1	$\sum_{k,t} c_{kt} = 1$
Parametric 2	$\sum_{k,t:(k,t)\neq(4,4)} c_{kt} = 0, c_{44} = 1$
Outcome Robust	$c_{11} + c_{21} + c_{31} + c_{41} = 0, c_{22} + c_{42} = 0, c_{33} + c_{43} = 0$, and $c_{44} = 1$.
Response Robust	$c_{11} = \pi_{00}, c_{21} + c_{22} = \pi_{10}, c_{31} + c_{33} = \pi_{01}$, and $c_{41} + c_{42} + c_{43} + c_{44} = \pi_{11}$
Double Robust	$c_{11} + c_{21} + c_{31} + c_{41} = 0, c_{22} + c_{42} = 0, c_{33} + c_{43} = 0, c_{44} = 1, c_{11} = \pi_{00}, c_{21} + c_{22} = \pi_{10}, c_{31} + c_{33} = \pi_{01}$, and $c_{41} + c_{42} + c_{43} + c_{44} = \pi_{11}$

Consider the case where we run estimate each model on a data set with $n = 1000$ observations, and define the following notation: let $\hat{\theta}^{(P)}$ and $\hat{\theta}^{(OR)}$ be the estimated values of $\theta = E[Y_2]$ for the `Parametric 1` model and the `Outcome Robust` model respectively. Define the estimated coefficients to be $\hat{c}_j^{(P)}$ and $\hat{c}_j^{(OR)}$ where $j = 1, \ldots 9$ for the `Parametric 1` and `Outcome Robust` models respectively. The one can compute the SSE with the following:

$$SSE = n^{-1} \sum_{i=1}^{n} (\hat{y}_{2i} - \theta)^2 \text{ where } \hat{y}_{2i} = \sum_{j=1}^{9} \hat{c}_j \hat{\gamma}_j$$

and $\hat{\gamma}_0 := \hat{\gamma}_{00} = \frac{\delta_{00i}}{\pi_{00}} E[Y_2 \mid x_i]$, $\hat{\gamma}_1 := \hat{\gamma}_{11} = \frac{\delta_{10i}}{\pi_{10}} E[Y_2 \mid x_i, y_{1i}]$, etc. Note that this is different from the previous estimate of the variance which used the Monte Carlo variance (standard deviation) defined by

$$\frac{1}{n-1} \sum_{b=1}^{B} (\hat{\theta}_b - \bar{\theta}_B)^2$$

where B is the number of Monte Carlo estimates and $\bar{\theta}_B = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_b$.

Likewise, one can compute the degrees of freedom by noticing that each model has a degrees of freedom equal to nine minus the number of constraints. This means that we can compute the model degrees of freedom and error degrees of freedom for each model type, which we do in Table 2

Table 2: *This table displays the degrees of freedom for each model*

Model	Model Degrees of Freedom	Error Degrees of Freedom
Parametric 1	$9-1=8$	$n-1-8=n-9$
Parametric 2	$9-1=8$	$n-1-8=n-9$
Outcome Robust	$9-4=5$	$n-1-5=n-6$
Response Robust	$9-4=5$	$n-1-5=n-6$
Double Robust	$9-8=1$	$n-1-1=n-2$

So continuing our first example, if $SSE^{(P)} = 993$ and $SSE^{(OR)} = 3428$ with $n = 1000$ then the F statistic is

$$F = \frac{(3428-993)/(3)}{993/(1000-9)} = 810.$$

The critical value we want to compare this with is the 0.95 quantile of $F_{3,993}$ which is 2.61. So there *is* a significant difference between the fit of these two model. (The p-value is basically zero.)