Testing – 6.5 Non nested hypotheses

Michel Bierlaire

Adjusted likelihood ratio index

The likelihood ratio index is defined as

$$\rho^2 = 1 - \frac{\mathcal{L}(\widehat{\beta})}{\mathcal{L}(0)}, \tag{1}$$

and called "rho-squared" by analogy to the R^2 in regression analysis. Note that it is not the square of anything. And even if it generally lies between zero (when $\mathcal{L}(\widehat{\beta}) = \mathcal{L}(0)$) and one (when the model perfectly fits the data and $\mathcal{L}(\widehat{\beta}) = 0$), there is no absolute interpretation of its value.

For the same estimation data set, the ρ^2 of a model always increases or at least stays the same whenever new variables are added to the model, even if they are actually irrelevant. This is the motivation to use the adjusted likelihood ratio index (rho bar squared):

$$\bar{\rho}^2 = 1 - \frac{\mathcal{L}(\widehat{\beta}) - K}{\mathcal{L}(0)}. \tag{2}$$

where K is the number of unknown parameters in the model.

The adjusted likelihood ratio index $\bar{\rho}^2$ can be used for testing non nested hypotheses of discrete choice models. Under the null hypothesis that model 1 is the true specification, compared to model 2, the following holds asymptotically:

$$\Pr(\bar{\rho}_2^2 \geq z + \bar{\rho}_1^2) \leq \Phi\{-\sqrt{-2z\mathcal{L}(0) + (K_1 - K_2)}\}, \qquad z > 0, \tag{3}$$

where

- $\bar{\rho}_\ell^2$ is the the adjusted likelihood ratio index for model ℓ,

- K_ℓ is the number of parameters in model ℓ,
- $\Phi(\cdot)$ is the standard normal cumulative distribution function.

This can be used to answer the following question: if we select the model with the largest $\bar{\rho}^2$, what is the probability that we make a mistake? In other words, if the true model is model 1, what is the probability that $\bar{\rho}_2^2$ is larger than $z + \bar{\rho}_1^2$, for some threshold z?

When all N observations in the sample have all J alternatives, we have

$$\mathcal{L}(0) = -N \log J \tag{4}$$

the bound becomes

$$\Pr(\bar{\rho}_2^2 \geq z + \bar{\rho}_1^2) \leq \Phi\{-\sqrt{2Nz \log J + (K_1 - K_2)}\}, \qquad z > 0. \tag{5}$$

Consider an example with $J = 2$ alternatives, $N = 300$ observations and two models with the same number of parameters ($K_1 = K_2$). Then we have

$$\Pr(\bar{\rho}_2^2 \geq 0.0001 + \bar{\rho}_1^2) \leq 41.92\% \tag{6}$$
$$\Pr(\bar{\rho}_2^2 \geq 0.001 + \bar{\rho}_1^2) \leq 25.95\% \tag{7}$$
$$\Pr(\bar{\rho}_2^2 \geq 0.01 + \bar{\rho}_1^2) \leq 2.07\% \tag{8}$$

For $z = 0.1$ and above, the probability is practically zero.