SCOREDUM - A stata command to test for fixed effects in Poisson and Logit models.

Paulo Guimaraes

CEFUP - Universidade do Porto

September 2012

Maximum Likelihood Estimation

- Suppose we have a probability model $f(Y; \mathbf{X}, \boldsymbol{\theta})$
- Y is the dependent variable, $\boldsymbol{\theta}$ is a vector of k unknown parameters and $\mathbf{X}$ is a set of covariates.
- Let $\mu_i = \mathbf{x}_i' \theta$
- The likelihood function is:

$$L(\boldsymbol{\theta}; \mathbf{Y}, \mathbf{X}) = \prod_{i=1}^{N} f(\boldsymbol{\theta}; y_i, \mathbf{x}_i)$$

- The m.l.e is the value $\widehat{\theta}$ that maximizes $L(\boldsymbol{\theta}; \mathbf{Y}, \mathbf{X})$

Testing Nested Hypothesis

- Suppose we want to test the null hypothesis $\mathbf{h}(\boldsymbol{\theta}) = 0$
- There are 3 likelihood based procedures for testing nested hypothesis:
  - Likelihood Ratio Tests
  - Wald Tests
  - Score Tests (or Lagrange Multiplier)
- These tests are all asymptotically chi-squared distributed with degrees of freedom equal to the number of restrictions imposed
- Let $\widehat{\theta}_r$ be the maximum likelihood estimator under the null hypothesis (the restricted MLE) and $\widehat{\theta}_u$ under the alternative (the unrestricted MLE).

Likelihood-Based Tests

- The Likelihood Ratio Test

$$LR = -2\left[lnL(\widehat{\theta}_r) - lnL(\widehat{\theta}_u)\right]$$

- The Wald Test

$$W = \mathbf{h}(\widehat{\boldsymbol{\theta}}_\mathbf{u})'\left[V\left[\mathbf{h}(\widehat{\theta}_u)\right]\right]^{-1}\mathbf{h}(\widehat{\boldsymbol{\theta}}_\mathbf{u})$$

- The Score Test

$$LM = \mathbf{s}(\widehat{\boldsymbol{\theta}}_\mathbf{r})'\left[\mathbf{I}(\widehat{\theta}_r)\right]^{-1}\mathbf{s}(\widehat{\boldsymbol{\theta}}_\mathbf{r})$$

Testing Group Comparisons

- Suppose the all N observations are classified into G mutually exclusive groups
- Groups can be panels or clusters
- Let $\boldsymbol{\theta}' = [\boldsymbol{\alpha}', \boldsymbol{\beta}']$
- $\boldsymbol{\alpha}$ is a subset of $\boldsymbol{\theta}$ with k_1 elements.
- $k = k_1 + k_2$ elements
- We want to test the following hypothesis

$$H_o : \boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 = \ldots = \boldsymbol{\alpha}_G \ .$$

The Score Test for Group Comparisons

- The score test is:

$$T = \mathbf{s}(\widehat{\boldsymbol{\vartheta}})' \left[-\mathbf{H}(\widehat{\boldsymbol{\vartheta}}) \right]^{-1} \mathbf{s}(\widehat{\boldsymbol{\vartheta}})$$

  the score equations and Hessian are calculated with respect to all the coefficients in $\boldsymbol{\vartheta}' = [\boldsymbol{\alpha}_1', \boldsymbol{\alpha}_2', ..., \boldsymbol{\alpha}_G'; \boldsymbol{\beta}']$, but evaluated at $\widehat{\boldsymbol{\vartheta}}' = [\widehat{\boldsymbol{\alpha}}', \widehat{\boldsymbol{\alpha}}', ..., \widehat{\boldsymbol{\alpha}}', \widehat{\beta}']$
- The $\widehat{\boldsymbol{\vartheta}}'$ are the m.l.e. solutions obtained under the null hypothesis
- The score test is asymptotically distributed as chi-square with $k_1(G-1)$ degrees of freedom

The Score Test for Group Comparisons

- Partitioning the score and Hessian with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ we get

$$T = -\begin{bmatrix} \mathbf{s}_\alpha \\ \mathbf{s}_\beta \end{bmatrix}' \begin{bmatrix} \mathbf{H}_{\alpha\alpha} & \mathbf{H}_{\alpha\beta} \\ \mathbf{H}_{\beta\alpha} & \mathbf{H}_{\beta\beta} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{s}_\alpha \\ \mathbf{s}_\beta \end{bmatrix}$$

- Since $\mathbf{s}_\beta(\widehat{\vartheta}) = \mathbf{0}$ the score test may be written as:

$$T = -\mathbf{s}'_\alpha \left[\mathbf{H}_{\alpha\alpha} + \mathbf{H}_{\alpha\beta}\left[-\mathbf{H}_{\beta\beta}\right]^{-1}\mathbf{H}_{\beta\alpha}\right]^{-1} \mathbf{s}_\alpha$$

$$T = -\mathbf{s}'_\alpha \left[\mathbf{H}^{-1}_{\alpha\alpha} + \mathbf{H}^{-1}_{\alpha\alpha}\mathbf{H}_{\alpha\beta}\left[\mathbf{H}_{\beta\beta} - \mathbf{H}_{\beta\alpha}\mathbf{H}^{-1}_{\alpha\alpha}\mathbf{H}_{\alpha\beta}\right]^{-1}\mathbf{H}_{\beta\alpha}\mathbf{H}^{-1}_{\alpha\alpha}\right] \mathbf{s}_\alpha$$

- $\mathbf{s}_\alpha$, $\mathbf{H}_{\alpha\alpha}$, $\mathbf{H}_{\alpha\beta}$ and $\mathbf{H}_{\beta\beta}$ can be easily calculated from the restricted model!

Simplifications

- If $k_2 = 0$ or $G = N$ then

$$T = -\mathbf{s}_{\alpha}' \mathbf{H}_{\alpha\alpha}^{-1} \mathbf{s}_{\alpha}$$

- Pearson tests are particular cases

$$T = -\sum_{g=1}^{G} \frac{s_{\alpha,\bullet g}^{2}}{h_{\alpha\alpha,\bullet g}}$$

- If Y is Poisson: $T = \sum_{g=1}^{G} \frac{(y_{\bullet g} - n_g \overline{y})^2}{n_g \overline{y}}$
- Y is Bernoulli: $T = \sum_{g=1}^{G} \frac{(y_{\bullet g} - n_g \overline{p})^2}{n_g \overline{p}(1-\overline{p})}$
- If Y is normal: $T = \sum_{g=1}^{G} \frac{n_g (\overline{y}_g - \overline{y})^2}{\sigma^2}$

SCOREDUM

Syntax

`scoredum [indepvars], group(varname) [options]`

Examples:

- Test for fixed effects in Poisson Regression

  ```
  poisson y x1 x2 x3
  scoredum, group(grvar)
  ```

- Test for "slippery slope" in `x1`

  ```
  scoredum x1, group(grvar)
  ```

- Test for differences in several coefficients across groups

  ```
  scoredum x1 x2 x3, group(grvar) cons
  ```

- Pearson chi-square test for count data

  ```
  poisson y
  scoredum, group(grvar)
  ```