Testing Hypothesis after Probit Estimation

Quantitative Microeconomics

R. Mora

Department of Economics
Universidad Carlos III de Madrid

Outline

1. Introduction
2. Exclusion Restrictions
3. Linear Hypothesis

The Probit Model and ML Estimation

The Probit Model

- $U_m = \beta_m x_m + \varepsilon_m$
- $U_h = \beta_h x_h + \varepsilon_h$
- $\varepsilon_h, \varepsilon_m \sim N(0, \Sigma)$ such that $\varepsilon \sim N(0,1)$
- $Pr(work = 1) = \Phi(\beta x)$ where Φ is the cdf of the standard normal

$$\hat{\beta}^{ML} = \arg\max \sum_i \{work_i \log(\Phi(\beta x_i)) + (1 - work_i)\log(1 - \Phi(\beta x_i))\}$$

- in `gretl`, a quasi-Newton algorithm is used (the BFGS algorithm)

Asymptotic Properties and Testing

under general conditions, the MLE is consistent, asymptotically normal, and asymptotically efficient

- we can construct (asymptotic) t tests and confidence intervals (just as with OLS, 2SLS, and IV)
- exclusion restrictions
  - the Lagrange multiplier requires estimating model under the null
  - the Wald test requires estimation of only the unrestricted model
  - the likelihood ratio (LR) test requires estimation of both models

The Likelihood Ratio Test

The LR test

- it is based on the difference in loglikelihood functions
- as with the F tests in linear regression, restricting models leads to no-larger loglikelihoods

$$LR = 2\left(l_{ur} - l_r\right) \xrightarrow{a} \chi_q$$

where q is the number of restrictions

Basic Commands in gretl for Probit Estimation

- probit: computes Maximum Likelihood probit estimation
- omit/add: LR or Wald tests for the joint significance
- $yhat: estimates probabilities
- $lnl: returns the log-likelihood for the last estimated model
- logit: computes Maximum Likelihood logit estimation

- in this Session, we are going to learn how to use omit, add, and $lnl

Example: Simulated Data

The Probit Model

- $U_m = 0.3 + 0.05 * educ + 0.5 * kids + \varepsilon_m$
- $U_h = 0.8 - 0.02 * educ + 2 * kids + \varepsilon_h$
- $\varepsilon_h, \varepsilon_m \sim N(0, \Sigma)$ such that $\varepsilon \sim N(0,1)$

- education brings utility if you work, dissutility if you don't
- having a kid brings more utility if you don't work
- $\beta x = -0.5 + 0.07 * educ - 1.5 * kids$

probit Output

probit *work const educ kids*

```
Convergence achieved after 6 iterations

Model 1: Probit, using observations 1-5000
Dependent variable: work

             coefficient   std. error   t-ratio     slope
  -----------------------------------------------------------
  const      -0.434462     0.0812490     -5.347
  educ        0.0659247    0.00576068    11.44      0.0240325
  kids       -1.47598      0.0407604    -36.21     -0.521270

Mean dependent var   0.366800   S.D. dependent var   0.364545
McFadden R-squared   0.233290   Adjusted R-squared   0.232378
Log-likelihood      -2519.525   Akaike criterion     5045.049
Schwarz criterion    5064.601   Hannan-Quinn         5051.902

Number of cases 'correctly predicted' = 3859 (77.2%)
f(beta'x) at mean of independent vars = 0.365
Likelihood ratio test: Chi-square(2) = 1533.26 [0.0000]

           Predicted
              0      1
  Actual 0  2495    671
         1   470   1364
```

`omit` *varlist* `--wald --quiet`

- *varlist* is a subset of controls in the last model estimated
- it gives the likelihood-ratio test for the joint significance of the variables in *varlist*
- if the `--wald` option is given, the statistic is an asymptotic Wald chi-square value based on the covariance matrix of the original model
- using the `--quiet` option:
  - only the result of the test is printed
  - the restricted model does not become the last estimated model in `gretl`'s memory (for access to `$coeff`, `$yhat`, `$uhat`, and `$lnl`)

Example: the LR test

omit *educ kids* --quiet

```
  Null hypothesis: the regression parameters are zero for the variables
     educ, kids

   Likelihood ratio test:
     Chi-square(2) = 1533.26, with p-value = 0
```

Example: the Wald test

omit *educ kids* ——wald

```
Null hypothesis: the regression parameters are zero for the variables
educ, kids
Asymptotic test statistic:
Wald chi-square(2) = 1362.14,      p-value = 1.636 8e-2
F-form: F(2, 4   ) ∓ 681.072,      p-value = 2.71422e-262
```

add *varlist* --quiet

```
? probit work const
Convergence achieved after 4 iterations

Model 2: Probit, using observations 1-5000
Dependent variable: work

             coefficient   std. error     z      slope
  -----------------------------------------------------
  const       -0.340341    0.0181027    -18.80

Mean dependent var   0.366800   S.D. dependent var   0.376494
McFadden R-squared   0.000000   Adjusted R-squared         NA
Log-likelihood      -3286.153   Akaike criterion     6574.306
Schwarz criterion    6580.823   Hannan-Quinn         6576.590

Number of cases 'correctly predicted' = 3166 (63.3%)
f(beta'x) at mean of independent vars = 0.376

           Predicted
               0      1
  Actual 0  3166      0
         1  1834      0

? add educ kids --quiet

  Null hypothesis: the regression parameters are zero for the
variables
    educ, kids

  Asymptotic test statistic:
    Wald chi-square(2) = 1362.14, with p-value = 1.63698e-296

    F-form: F(2, 4997) = 681.072, with p-value = 2.71422e-262
```

Linear Hypothesis Testing Using the LR

- since we can recover the log-likelihood, it is possible to compute tailor-made likelihood ratio tests

$\beta x = -0.5 + 0.07 * educ - 1.5 * kids$

- $H_0 : 2 * \beta_{educ} = -\beta_{kids}$

- estimate the unrestricted model and store the log-likelihood, l_{ur}
- estimate the restricted model and store the log-likelihood, l_r
- compute the likelihood ratio, $LR = 2 * (l_{ur} - l_r)$
- compute its asymptotic p-value under the null: $\Pr(\chi^2_1 > LR)$

Testing Linear Hypothesis in `gretl`

- `$lnl`: returns the log-likelihood for the last estimated model
- `pvalue(`*c* `[,` *argument* `,...],`*value*`)`: Returns $\Pr(X > x)$, where
  - the distribution X is determined by the character *c*
  - required parameter(s) for X are set with *argument* (,...)
  - x is determined by *value*

Examples

- p1 = pvalue(z, 2.2) # z: standard normal
- p2 = pvalue(X, 3, 5.67) # X: chi-square
- p2 = pvalue(F, 3, 30, 5.67) # F: Snedeckor's F

Example: $H_0 : 2 * \beta_{educ} = -\beta_{kids}$

Unrestricted model: $\beta x = \beta_0 + \beta_{educ} * educ + \beta_{kids} * kids$

Restricted model: $\beta x = \beta_0 + \beta_{educ} * (educ - 2 * kids)$

```
 outfile --write null
 # estimating unrestricted model and storing loglikelihood
 probit work const educ kids --quiet
 scalar lur= $lnl

 # estimating restricted model and storing loglikelihood
 genr x=educ-2*kids
 probit work const x --quiet
 scalar lr= $lnl

 # computing the LR statistic and p-value
 scalar LR=2*(lur-lr)
 scalar pval = pvalue(X, 1, LR)

 # printout
 outfile --close
 printf "\nLikelihood Ratio test\nH0: 2*beta_educ+beta_kids=0\nLR
%.8g   p-value %.8g,\n", LR, pval
```

Example's Output

```
Likelihood Ratio test
H0: 2*beta_educ+beta_kids=0
LR 1139.6918   p-value 7.8025612e-250
```

Summary

- `gretl` allows for testing exclusion restrictions after probit estimation
- the likelihood ratio and the wald tests are available
- it is not difficult to test homogeneous linear hypothesis with a little bit of programming