MSc Development Economics: Quantitative Methods
Maximum Likelihood Estimation in Stata

Selma Telalagić

University of Oxford

1st November 2013

Recap of Maximum Likelihood Estimation

- Our running example:
- What is the relationship between wages and schooling?
- A statistical model describing this is

$$W_i = \alpha + \beta S_i + u_i$$

- This is the true relationship; how do we estimate α and β?

What is Maximum Likelihood Estimation?

- A way of estimating the parameters of a statistical model i.e. α and β
- Based on a likelihood approach
- What is the most likely value of a parameter that is consistent with the observed data?
- Maximisation takes into account changing errors and parameters jointly to yield the parameter that gives the maximum likelihood

The Log-Likelihood Function

- Assuming errors are normally distributed
-

$$\ell_{W_1,\ldots,W_n|S_1,\ldots,S_n}(\alpha,\beta,\sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(W_i - \alpha - \beta S_i)^2$$

Yesterday's questions about MLE

- Why are the Wald test, LR test and Score test asymptotically equivalent? Proved by Engle:
  - Engle, Robert F. (1983). "Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics". In Intriligator, M. D.; and Griliches, Z. Handbook of Econometrics II. Elsevier. pp. 796–801.

MLE in Stata

- A good reference is Steenbergen (2003), see http://www.unc.edu/~monogan/computing/r/MLE_in_Stata.pdf
- We will be using wage and education data
- `clear`
- `use sa_wage_curve_1`
- `sum wphy logwphy edyrs`
- `graph twoway scatter wphy edyrs`
- `graph twoway scatter logwphy edyrs`

How do we know to use logs?

- `reg wphy edyrs`
- `rvpplot edyrs`
- `reg logwphy edyrs`
- `rvpplot edyrs`

Running these regressions with maximum likelihood

- First, we need to define a program that specifies the parameters and the log-likelihood function
- `version 9`
- `program MLE`
- `args lfn beta sigma`
- `quietly replace ‘lfn’= ln(normalden($ML_y1, ‘beta’, ‘sigma’))`
- `end`

Specify the model

- `ml model lf MLE (logwphy = edyrs) /sigma,tech(bfgs 5 dfp 5 nr 5 bhhh 5)`

ML Check

- Check the program for errors
- `ml check`

ML Search

- Look for starting values
- `ml search`

ML Max

- Find the parameter values that maximise the log-likelihood function
- `ml max`
- Compare to OLS:
- `reg logwphy edyrs`
- Why the difference?

ML Graph

- Graphs the value of the log-likelihood function against the iteration number
- Useful to see if convergence is smooth
- Stata may not have found the global maximum
- Beware of sudden jumps in convergence; the log-likelihood should be concave when approaching convergence, not flat or convex
- Beware if convergence takes a long time (up to 15 is fine; 100 is not!)
- Beware if the final log-likelihood is reported as 'not concave'; this implies the log-likelihood is flat at that point and Stata doesn't know which direction to go in
  - Add option `difficult` to `ml maximise`
- Re-run the estimates; beware if it stops at a different value of the log-likelihood or provides different estimates of the parameters
- Beware if Stata is unable to calculate a standard error
- `ml max`
- `ml graph`

Inference: Wald test & Score test

- `test [varname]`
- The Score test is not a standard command in Stata and there are currently no user-written commands that calculate it
- You would have to program the test manually and I advise against it because it is not a widely-used test

Likelihood Ratio Test

- Null hypothesis: Education has no effect on wages
- Define a constraint
- `constraint 1 [eq1]edyrs = 0`
- Run unrestricted model first
- `ml model lf MLE (logwphy = edyrs) /sigma,tech(bfgs 5 dfp 5 nr 5 bhhh 5)`
- `ml max`

Likelihood Ratio Test cont.

- Save the results and inform Stata you are intending to run the LR test
- `lrtest, saving(0)`
- Run the restricted model
- `ml model lf MLE (logwphy = edyrs) /sigma,tech(bfgs 5 dfp 5 nr 5 bhhh 5) constraint(1)`
- `ml max`
- Perform the LR test
- `lrtest`

Summary

- Define the program with parameters and distribution assumption
- Specify the model
- Check your program and search for an initial value
- Maximise the log-likelihood
- Graph to check convergence
- Run Wald and LR tests on hypotheses