

Economics 696F, Homework 3

Due Friday, February 17, 2006

For this homework, you may use any standard statistical package, such as Matlab or Stata. Make sure your code is clearly commented and explained in your own words.

1. Consider a logit model: y is binary, equal to 0 or 1, with

$$Pr(y = 1|x) = \frac{\exp(x'\beta)}{1 + \exp(x'\beta)}.$$

Here, x is a $k \times 1$ vector which may include a constant.

- (a) Write a routine to calculate the maximum likelihood estimator of β given a random sample of size n . The routine should take as inputs a vector y and a matrix x , and produce the value of β which maximizes the likelihood function. The maximization should be done using the Newton-Raphson routine.

The routine should start at $\beta_0 = \underline{0}$ and iteratively update the value of β according to

$$\beta_{k+1} = \beta_k - [\nabla_{\beta\beta}L(\beta_k)]^{-1} \cdot \nabla_{\beta}L(\beta_k),$$

where $L(\beta)$ is the log likelihood function. (Note: the gradient and Hessian of the logit log-likelihood have a closed form expression which is relatively easy to calculate, so it should not be necessary to use “numerical” derivatives.) Stop when $\|\beta_{k+1} - \beta_k\|$ is less than some user-specified constant.

For this part and the next, generate some fake data from the logit model and use it to show that the routine behaves as expected.

- (b) Suppose we assume that the sample was a choice-based sample (with strata corresponding to the two possible values of y) with known population shares. Modify the routine to calculate the weighted ML, using the weighting scheme discussed in LN5. The routine should now accept as an input the population shares Q_s , in addition to y and x .