Econ 342 Feb. 4, 2008
Shakeeb Khan Due: Feb. 15
Spring 2008

Problem Set 3

1. Let $y_i = \max(x_i'\beta_0 + \epsilon_i, 0)$. Formally derive the expression for $E[y_i|x_i, y_i > 0]$ and $E[y_i|x_i]$ assuming ϵ_i is normally distributed with variance σ_0^2.

2. Now suppose we alter the basic censored regression model in the following way:

$$y_i^* = x_i\beta_0 + \epsilon_i \quad \epsilon_i \sim N(0, \sigma_0^2) \tag{1}$$
$$y_i = cI[y_i^* \leq 0] + y_i^* I[y_i^* > 0] \tag{2}$$

   where c is some constant.

   (a) Find $E[y_i|x_i, c]$

   (b) How does the MLE change with c?

3. Here we will use Matlab or Gauss to use the NLS Probit estimator using the Gauss Newton Method. (See the answer to question 2 on page 500 of Hayashi to see how.)

   (a) First, a simulation study. Simulate data for the following model:

$$y_i = I[\alpha_0 + x_i\beta_0 + \epsilon_i > 0]$$

   where $\alpha_0 = \beta_0 = 1$, $x_i \sim N(0,1)$, $\epsilon_i \sim N(0,1)$, and ϵ_i, x_i are independent. Generate 100 i.i.d observations and estimate α_0, β_0. Repeat this 101 times so you can report the following statistics: mean bias and RMSE. Do this again for 400 observations at what rate should RMSE decline? Does it?

   (b) Now we'll try NLS Probit on a real data set. Download the Excel file hitstat.xls from the class web page: the columns correspond to at bats, home runs, rbis, stolen bases and price of the player. First construct a dummy left hand side variable which is 1 if the price exceeds 15 and 0 otherwise. Do NLS Probit of this dummy on an intercept, homeruns, rbis and stolen bases. Construct a 95% confidence interval for each of the coefficients.