Exercise - Week 8

Part 1: Poisson Regression Model

- Load the dataset, CRIME1.DTA. This contains data on arrests during the year 1986 and other information on 2,725 men born in either 1960 or 1961 in California. Each man in the sample was arrested at least once prior to 1986. The variable *narr86* is the number of times the man was arrested during 1986. The variable *pcnv* is the proportion of arrests prior to 1986 that led to conviction, *avgsen* is average sentence length served for prior convictions, *tottime* is months spent in prison since age 18 prior to 1986, and *qemp86* is the number of quarters (0 to 4) that the man was legally employed in 1986.

- We are interested in understanding the determinants of the number of arrests for this group of young men.

1 Tabulate the variable narr86 and see how that variable is distributed. Also scrutinize simple statistics of other variables.

2 First, run the following OLS regression:

$$narr86_i = \beta_0 + \beta_1 pcnv_i + \beta_2 avgsen_i + \beta_3 tottime_i + \beta_4 qemp86_i + \beta_5 black_i + u_i$$

What is the effect of pcnv on the number of arrested numbers?

3 Estimate the same regression model using the Poisson model by typing **poisson narr86 pcnv avgsen tottime qemp86 black**. How does the interpretation of β_1 differ from that in the OLS regression? What is the effect of pcnv?

4 Compute the mean of predicted number of arrests (**predict pnarr86,n** and **su pnarr86**). How close is it with the sample mean of the number of arrested times?

5 Compute the marginal effects of each independent variable on the predicted number of arrested times (**mfx compute, predict(n)**). Interpret them.

Part 2: Tobit Model

- We want to estmate the effect of income on the willingness to pay (WTP), controlling for age, sex, and smell, while taking into account the fact that WTP cannot be negative (the variable is censored). This will be done by using a Tobit model.

1 First, tabulate the variable, WTP, to see what is the proportion of individuals who report a zero WTP. Run the OLS regression of WTP on log income, age, sex, and smell:

$$WTP_i = \beta_0 + \beta_1 \ln y_i + \beta_2 age_i + \beta_3 sex_i + \beta_4 smell_i + u_i.$$

2 Consider the following Tobit model about the latent variable, WTP^*, and observed variable, WTP:

$$\begin{aligned} WTP_i^* &= \beta_0 + \beta_1 \ln y_i + \beta_2 age_i + \beta_3 sex_i + \beta_4 smell_i + u_i \\ WTP_i &= \max\left(WTP_i^*, 0\right). \end{aligned}$$

Estimate the Tobit model by typing **tobit wtp lny age sex smell, ll(0)**. Note that ll(0) represents the lower limit for left censoring. Compare with the result obtained from an OLS regression. How are the coefficients on age and log income changing?

3 (Marginal Effects of Tobit model) We are interested in predicting how the WTP would change when household income increases.

- First, what is the marginal effect of log income on the latent WTP?
- Considering the censoring, what is the marginal effect of log income on the observed WTP? ($\partial E\left(WTP_i|X\right)/\partial \log income_i = \beta_1 \Phi\left(\beta' X/\sigma\right) = \beta_1 \Pr\left(WTP_i^* > 0|X\right)$). Note that the value of marginal effect does depend on the values of independent variables and thus we will compute the marginal effect given the mean value of each independent variable. To calculate this value, we need to compute $\Pr\left(WTP_i^* > 0|\overline{X}\right)$, where $\overline{X}$ means the mean values of independent variables. Type **mfx compute, predict(pr(0,.))**, where the value of $\Pr\left(WTP_i^* > 0|\overline{X}\right)$ will be reported.
- Compare the above two marginal effects.