Maximum Likelihood Estimation for Generalized Linear Models

2012 ICPSR Summer Program

Exercise Two

Due Tuesday, July 3, 2012

Purpose

The purpose of this exercise is to estimate ordered probit and ordered logit models.

Data

The data for this exercise (available as a Stata 11 file in `ANES2008subset.dta`) are from the 2008 American National Election Study

- `V08001` is the respondent's case id, a unique identifier for each respondent. Always keep a unique identifier for each observation in your data set.
- `V085044a` is the vote choice in the 2008 presidential election. Recode so that vote for Obama=1, Vote for McCain=0, vote for anyone else is missing data.
- `V083097` is party identification on a 3-point scale. Recode so that Democrat is the highest value, Independent the middle value, Republican the lowest value, and all other values coded as missing data.
- `V081101` is respondent's gender. Recode this as a dummy variable where female=1 and male=0.
- `V083248` is respondent's household income in 25 categories.
- `V083057` is the respondent's retrospective evaluation of the his/her personal finances. Recode this so that the order of the responses makes sense.
- `V083083` is the respondent's retrospective evaluation of the national economy. Recode this so that the order of the responses makes sense.

Exercise

The models you will estimate are:

$$\begin{aligned} \texttt{National Economy}_i &= f[\beta_0 + \beta_1 \texttt{Democrat}_i + \beta_2 \texttt{Republican}_i + \beta_3 \texttt{Income}_i + \\ &\quad \beta_4 \texttt{Gender}_i + u_i] \end{aligned}$$

$$\begin{aligned} \texttt{Personal Finances}_i &= f[\beta_0 + \beta_1 \texttt{Democrat}_i + \beta_2 \texttt{Republican}_i + \beta_3 \texttt{Income}_i + \\ &\quad \beta_4 \texttt{Gender}_i + u_i] \end{aligned}$$

1. Recode party identification into dummy variables for `Democrat (1=Democrat, 0 otherwise` and `Republican (1=Republican, 0 otherwise)`.

2. Leave the income variable on its 25-point scale. Do not center it. Do not divide by two standard deviations.

3. Estimate a linear regression model and either an ordered logit model or ordered probit model for each dependent variable. Produce a table of your results showing the estimates and standard errors for the effect of each independent variable. The table is best constructed with independent variable labels as rows with separate columns for each of the linear regression and ordered logit or probit models for each of the two economic evaluations. Report the likelihood ratio for the ordered logit or probit model.

4. Briefly (in a few sentences) evaluate and compare the statistical and substantive significance of the variables between the ordered logit/probit models of retrospective national evaluations and retrospective personal evaluations. Do you notice any significant differences in the results across the linear probability, logit and probit models?

5. For each ordered logit/probit model, test whether each successive cut point is statistically different from the previous cut point.

6. Calculate the predicted value on each of the two dependent variables for a middle income woman who is a political independent.

7. Graph the predicted probabilities for each dependent variable as a function of income (on its 25 category scale).

8. Graph the cumulative probabilities for each dependent variable as a function of income (on its 25 category scale).

This assignment is due *in hard copy* at 9:00 a.m. EST on Tuesday, July 3, 2012.