Exercise - Week 7

- Load the wtp.dta data set.
- Construt the log of income variable: **gen lny = ln(y)**

Part 1: Multinomial Logit Model

- We are interested in predicting the probability of living (1) in a house, (2) in an apartment and (3) in a low-cost flat. We would like to predict these probabilities given individual characteristics (age, sex, income, education and size of household). We then do some policy analysis, varying the education level and see how this affects the choices of individuals.

1 Construct a variable **accom**, which takes values, 1 (house), 2 (apartment) or 3 (low-cost flat). To this end, type **gen accom = house**, and then **recode accom 1/3 = 1 4 = 2 5 = 3 6/max =.**.

2 Construct a variable **educ** which takes the value 1 if g_6 = 7 (high education). (type **gen educ = g_6 == 7**)

3 Estimate the multinomial logit by typing **mlogit accom age sex educ lny g_4, b(3)**, where g_4 is the size of the household and b(3) represents that we take the choice of low-cost flat as the baseline choice. The estimated coefficient of X for, say, living in a house is to be interpreted as the marginal effect of X on the odds-ratio of living in a house relative to living in a low-cost flat.

4 Predict the probabilities, pHouse, pApart, pFlat for each case. (type **predict pHouse, outcome(1)**, and similarly for the other probabilities). Graph these probabilties as a function of log income.

5 Compute the marginal effects of log income on the choice probabilities of House, Apartment and low-cost Flat. What do you find?

6 We now look at a policy experiment which would change individual's education, while holding constant all other variables such as age and sex. Predict the probabilities assuming that everybody are highly educated. To this end, do the following:

   - **replace educ = 1**, then calculate the predicted probabilities pHouseH, pApartH and pFlatH (**predict pHouseH, outcome(1)**). Note that here we imagine that everyone in the sample has a high education and predict the probabilities for each individual, while holding constant age, sex, income and the size of household.
   - **replace educ = 0**, then calculate the predicted probabilities pHouseL, pApartL and pFlatL (**predict pHouseL, outcome(1)**).
   - compare the adjusted probabilities with the unadjusted ones. What is the effect of education?

Part 2: Ordered Probit Model

- We want to evaluate the determinant of the attitudes toward the environment. The variable a_2 is a positive integer variable which takes the values between 1 (not concerned at all) and 5 (very concerned). Explanatory variables are log of income, age, sex and smell.

1 Tabulate the variable a_2 indicating the environmental concern. What are the frequecies of different values of the environmental concern? And tabulate jointly the variable a_2 and sex (**tab a_2 sex, col**). What is the effect of sex on the relative frequency of choosing outcome 5?

2 Estimate the ordered probit with log income, sex, age and smell as independent variables by typing **oprobit a_2 lny sex age smell**. Which variables are significant and insignificant? What can you say about the effects of income on the probability of choosing a given outcome?

3 Predict the probabilities associated with each outcome (type **predict p1 p2 p3 p4 p5**). Compare the predicted probabilities with the observed frequencies obtained through **tab a_2**.

4 Calculate the marginal effect of log income on the probabilities of each outcome. Remember the marginal effect of X_1(log income) on each probability is

$$\begin{aligned}
\frac{\partial \Pr(Y=1)}{\partial X_1} &= -\beta_1 \phi(\mu_1 - \beta X) \\
\frac{\partial \Pr(Y=2)}{\partial X_1} &= \beta_1 [\phi(\mu_1 - \beta X) - \phi(\mu_2 - \beta X)] \\
\frac{\partial \Pr(Y=3)}{\partial X_1} &= \beta_1 [\phi(\mu_2 - \beta X) - \phi(\mu_3 - \beta X)] \\
\frac{\partial \Pr(Y=4)}{\partial X_1} &= \beta_1 [\phi(\mu_3 - \beta X) - \phi(\mu_4 - \beta X)] \\
\frac{\partial \Pr(Y=5)}{\partial X_1} &= \beta_1 \phi(\mu_4 - \beta X),
\end{aligned}$$

where βX represents $\beta_1 \log income + \beta_2 sex + \beta_3 age + \beta_4 smell$. To this end, type mfx compute, predict (outcome(1)) to get the marginal effects for choice 1 (not concerned at all). What is the effect of a poor quality neighborhood (smell = 1) on the probability of choosing options 1 to 5? What is the effect of sex on choosing outcome 5? Compare with what you got in question 1.