

PS206 Homework #6 (Due Tuesday, May 19th)

Listed under Homework #6 on our class page is a tab-delimited text dataset called “ps206hw6.txt”. This data is from a survey conducted just before the 1992 U.S. presidential election, and contains 1650 observations on seven variables – vote (1 for Bush, 2 for Clinton, 3 for Perot), opinions on changes in personal finances over the last year (-1 for got worse, 0 for no change, 1 for got better), opinions on changes in the national economy over the last year (-1 for got worse, 0 for no change, 1 for got better), income (a 23-point scale with higher numbers indicating greater income), age (in years), a 7-point scale for education (with higher numbers indicating more education), and dummy variables for race (0 for white, 1 for minority) and gender (0 for men, 1 for women).

Download the data and do the following five problems:

(1) Estimate a multinomial logit to determine what effect opinions on personal finances, opinions on the national economy, age, gender, education, and race had on vote choice. Discuss what the estimated coefficients tell you about the effect of the independent variables on the dependent variable. In **Stata** you simply use the `mlogit` command. In **R**, first install the “nnet” library, if it’s not already installed, and then load it as before. Then set all categorical variables to factors as usual. You can then estimate a multinomial model using the command `multinom`. For example, if your dependent variable is “y”, your independent variables are “x” and “z”, and your data set is “mydata”, type: `multinom(y ~ x + z, data = mydata)`. You’ll want to save the output of the model as usual.

Interpret the results of this multinomial logit as far as you can.

(2) Now we will look at some predicted probabilities. As in previous homeworks, we will first calculate these predicted probabilities the long way. Come up with a hypothetical individual, and calculate the probability that this hypothetical individual would vote for each of the three candidates. That is, you will need to calculate:

$$Pr(y = m) = \exp(X\beta_i) / \sum_j \exp(X\beta_j)$$

for each of the three candidates for this hypothetical person. Report the probabilities you have calculated.

(3) Now change this hypothetical person’s opinion on the national economy, and recalculate the probabilities from problem 2. Describe how the probabilities have changed from those calculated problem 2, and discuss whether this makes sense.

(4) Now we will calculate predicted probabilities with standard errors. For **Stata** you should use *Clarify*, while for **R** you will want to modify the code I provided for Homework #3 (the multiple draws of coefficients will be the same – you then just plug them into a different formula to calculate the probabilities). Calculate the probabilities of voting for each of the three candidates for a hypothetical individual profile you did not consider in problems 2 or 3. Report what you have found.

(5) Finally, estimate a likelihood-ratio test to determine if constraining the coefficients on the economic opinion variables to 0 (that is, omitting these variables from the model) is a valid constraint. The commands to do this are covered in Homework #4.