

Quant II - Problem Set VII

Models for Count Data and Dealing with Uncertainty in Inferential Models

Kenneth Benoit

Assigned: Wednesday, March 11, 2009

Due: Friday, March 20, 2009

1. The STATA-dataset AFFAIRS includes among others the following variables for 601 survey respondents:
 - naffairs - the number of extramarital affairs within the last year
 - male - equals one if respondent is male
 - age - the age of the respondent
 - yrsmarr - the number of years married
 - kids - equals one if respondent has kids
 - ratemarr - a rating of the marriage: 5 = very happy, 4 = happier than average, 3 = average, 2 = somewhat unhappy, 1 = very unhappy
 - relig - degree of religiosity: 5 = very religious, 4 = somewhat religious, 3 = slightly religious, 2 = not at all religious, 1 = anti-religious
- (a) Convert the variables ratemarr and relig into factors and add appropriate levels.
- (b) Run a negative binomial model regressing the number of affairs on the other variables mentioned above. Interpret the findings.
- (c) Perform an appropriate test of the hypothesis that the coefficients referring to the rating of the marriage and religiosity are jointly insignificant. Interpret.
- (d) Plot expected values of the number of affairs against duration of marriage, holding other variables at their medians. What are the characteristics of the person the plot refers to? (Notes: Use the sample range of marriage for the x-axis and a range of 0 to 3 for the y-axis. There is no need to use simulation-based methods here.)

- (e) What is the *expected* number of marital affairs for a man of age 38, who is married for 6 years, has kids, is not at all religious and reports that the happiness of his marriage is average? First answer the questions by using `predict()`, then employ a simulation-based approach using Zelig. Compare the results and report the confidence intervals Zelig gives.
 - (f) Using Zelig, what is the first difference in the *expected* number of marital affairs if the man described in the previous question is compared to the same man five years later in his life? Assume that all variables not referring to time remain constant. Give confidence intervals.
 - (g) Same question as the previous one, but what if, in addition to the passage of 5 years, the marriage of the man becomes somewhat unhappy while he also turns somewhat religious?
 - (h) Extra credit: Using Zelig, what is the *predicted* probability of having at least one affair for a 42-year old man, who is married 10 years, has no kids, is somewhat religious, and whose marriage is somewhat unhappy?
2. Use the STATA dataset `habeas.dta`, which gives information on all cases of Habeas Corpus decided by the US Supreme Court between 1953 and 1996.

The variable we want to explain is *propetit*, which equals one if the court ruled in favour of the petitioner. The main independent variable of interest is *liberal*, a measure of the Court's liberalism ranging from 0 (most conservative) to 1 (most liberal). Other variables are *ineffcou* (= 1 if the habeas claim was based on assertion of ineffective counsel), *tcterror* (= 1 if the claim was based on assertion of trial court error), *multpet* (= 1 if the case was a second or subsequent petition for habeas), and *usparty* (= 1 if the US federal government was a party to the litigation).

Run a logit model and plot simulation-based expected probabilities of a pro-petitioner ruling associated with different levels of Supreme Court liberalism, holding all other variables constant at their medians. Include confidence intervals. (Hint: Have a look at the examples discussed in class.)