Quantitative Methods II
Assignment 11
November 6, 2011
Solutions

PROBLEM: BAC TO THE FUTURE   ⟦5⟧

- RQ: Are the reaction times significantly different for people with a Blood Alcohol Content level of 0.08% versus a BAC level of 0.10%?

- Unit of analysis: Individual person

- Data: The dependent variable is a measure of reaction times for the individual. These range from 0.05 to 0.12 seconds. The independent variable is the BAC level for the individual. These are either 0.08% or 0.10%.

Solution: The key is to first look at the dependent variable. Here, the dependent variable is a continuous variable. There may be some concern that they need to be log-transformed; however, as none of them get (relatively) close to zero, I would merely double check to see if this is a factor. It should not be.

Now, as the independent variable is categorical, we could use either regression or a t-test to compare the two means. Either is acceptable. Both make the same assumptions (Normality, etc.).   ◇

Problem: Duck, Duck, Goose   ⟦5⟧

- RQ: To what extent does ambient temperature affect the departure times for the Lesser Snow Goose?
- Unit of analysis: Goose flock
- Data: The dependent variable is the time the flock leaves its roost (in minutes before or after sunrise). The research variable is the ambient temperature, measured in degrees Celsius. The control variables used are relative humidity (values between 51 and 100%), light intensity (values between 5.0 and 14.2), and cloud percentage (values between 0 and 100%).

Solution: The key is to first look at the dependent variable. Here, the dependent variable can take on positive and negative values, is not bounded, and is continuous (or close enough to it). As such, GLM using a Gaussian distribution and identity link is sufficient.   ⋄

PROBLEM: THE JAWBONE OF A ... RABBIT?   ⟦5⟧

- RQ: What is the effect of age on the length of a jackrabbit's jawbone?
- Unit of analysis: Jackrabbit
- Data: The dependent variable is the length of the jawbone (in mm). In this data, it ranges from 15.5 to 50.7mm. The independent variable is the age of the jackrabbit (in years). In this data, the age ranges from 0.01 to 3.65 years.

Solution: The key is to first look at the dependent variable. Here, the dependent variable is a length, which is continuous and bounded below. Is the lower bound important? It may be. As such, I would log-transform the dependent variable and do my regression just to be safe—it would not hurt. So, a GLM using the Gaussian distribution and a log link. ◇

- RQ: What is the effect of the level of democracy on the probability that a politically violent group will use terrorism?
- Unit of analysis: Politically violent group
- Data: The dependent variable is whether or not the group used terrorism that year (2002). The research variable is the democracy level of the State (ranging from -10 to 10). The control variables are the State's stability (in years since last coup), the economic strength of the State (in percent change of GDP from last year), and number of ethnic groups in the State (ranging from 1 to 15 for this dataset).

Solution: The key is to first look at the dependent variable. Here, the dependent variable is a binary variable (0/1). As such, we need to use one of the binary dependent variable links. As terrorism is a rare event, you will want to compare the results under logit, log-log, and complementary log-log to determine which model best fits the data.   ◇

Problem: Taking the Initiative is a Crime?   ⟦5⟧

...What is the predicted number of times this state uses the initiative process?

Solution: I predict that the state will use the initiative process 9 times in this decade.

As the dependent variable is a count, I performed count regression using the Poisson, quasi-Poisson, and negative binomial distributions. Of course, there was overdispersion, so the Poisson was not applicable. Unfortunately, the results of the quasi-Poisson and the negative binomial models were quite different.

As the unified government variable was not statistically significant, I removed it from the analysis and used the reduced model. The quasi-Poisson and the negative binomial still gave difference estimates. However, the differences between the two were much smaller: 8.9 versus 9.9. As such, I used the prediction from the quasi-Poisson model.   ◇