Methodology Minor Field Exam

Fall 2017

For the minor field exam, you must answer two questions, one in the morning session and one in the afternoon session. In the afternoon session, you may use the software of your choice. You are free to use whatever word processing or typesetting software you like to write your answers. The questions must be answered in the allotted time.

Morning Session: Statistical Theory and Modeling Decisions

Answer one of the following two questions:

1. *Bayesian Statistics:* Bayesian statistics require the analyst to set a prior on each parameter that is estimated. There are several ways in which an analyst can set a prior. Name three different techniques for choosing a prior for a parameter in a Bayesian model. For each technique, describe how the prior is formulated. For each technique, what are some reasons that someone might find that kind of prior useful?

   Now consider an example. Suppose you were studying a small population like the 50 American states. 26 of the states allow ballot initiatives, meaning that citizens can put their own policy questions on the public ballot with enough signatures. The other 24 do not have this institution. Imagine you wanted to evaluate a conditional relationship: Is the strength of the relationship between public ideology and public policy stronger in initiative states than non-initiative states? To test that, your dependent variable is per capita public spending by the state (higher values are assumed more liberal), your independent variable is the percentage of the public that identifies as liberal, and your conditioning variable is an indicator for an initiative state. You want to know the probability that an interaction term between *percent of the public that is liberal* and the *initiative state dummy* is positive.

   How would you choose the priors for a model like this? Explain how you would apply one of the three techniques you described at the beginning of the answer for this specific problem. What would this choice mean for each of the parameters in the model you are estimating? Why is the strategy that you chose better in this specific case than the other two that you named? Your answer may or may not refer to what you *believe* would be the case about the quality of the measures, whether you think there are good outside sources of information, how much you want the data to speak for themselves, or the fact that $N = 50$ in this case.

2. *Qualitative Analysis:* Consider the following puzzle: Many international borders have a dyad of cities directly opposite each other on the border. On the U.S.-Mexico border, El Paso and Juárez are adjacent. On the Sweden-Finland Border there are Haparanda and Tornio. On the Estonia-Russia border there are Narva and Ivangorod. Brazzaville and Kinshasa are opposite each other on the Republic of the Congo-Democratic Republic of the Congo border. In many cases these city dyads have developed policies to cooperate with each other in the local economy, with regard to emergency services, and in the construction of infrastructure. Suppose you wanted to answer the following research question: Why do some border-sharing city dyads cooperate more than others do? One hypothesis that has been raised as a possible answer is that when a city is in a more decentralized country, it is better able to forge an alliance with an international partner. Therefore, we ought to see more cooperation when both of the cities are in decentralized countries than when only one of them or neither of them are in a decentralized country.

   How would you construct a qualitative research design to test this hypothesis? Be sure to explicitly address how you would select the cases of city dyads that you would study. What kinds of information would you need to gather for each case? Would you examine the historical record? Conduct interviews with officials, business elites, or ordinary citizens? Would you examine records of laws and agreements? Whatever type of information you would gather, explain what specifically you would need to study (e.g., what kinds of historical details you would need, or whom you would need to interview). Explain how it would help you test this hypothesis.

   Based on your proposed research design, in what circumstance would you conclude that the hypothesis is correct? That is, what evidence would convince you that national decentralization does lead to international city dyad cooperation? In what circumstance would you conclude that the hypothesis is wrong and there is no evidence that national decentralization influences city cooperation? Why does each set of circumstances lead to the respective conclusion?

Afternoon Session: Analyzing Data

Answer one of the following two questions:

3. *Ordered Logit Resgression:* Please analyze the data set *ches2014* using an ordered regression model. The outcome of interest in *EUposition*, which is a 3 category odered variable with 1 representing anti-EU positions, 2 moderate positions, and 3 pro-EU positions for poltical parties in the EU member states in 2014 The data can be found here: http://spia.uga.edu/faculty_pages/rbakker/ches2014.dta

   There are several different arguments in the scholalry literature about what types of parties are more or less likely to be in favor of futher European integration. Some argue that economic position is the most important predictor of a party's stance toward the EU. Additionally, the literature finds that parties in government tend to be more pro-EU than opposition parties. In more recent work, scholars have argued that more socially progressive parties tend to be more pro-EU. Estimate a model using all of the variables in the data set and discuss whether or not your results support the above hypotheses.

   The variables in the data set are:

   - vote–the percentage of vote won by the party in the most recent national election
   - govt–a dummy variable for whether (1) or not (0) the party was in government
   - eu salience–a measure of how salient the EU is for a party
   - eu dissent–a mesure of how much internal dissent there is within a party over the EU
   - lrecon–the economic left-right position of a party (higher values = more right-wing
   - galtan–the social left-rigth position of a party (higher values = more socially traditional

   Present the results of the model in a table including the coefficients, standard errors, and any additional information you think is relevant. What can you conclude from the results you present? Please also provide predicted probabilty graphs for the galtan and the govt variables and describe what the graphs illustrat in terms of the probabilty of a party being pro-EU.

4. *Poisson Regression:* Please analyze the data set *couart2.dta* using a poisson regression model. The data set contains information on the number of articles published by PhD students during the last 3 years of their education.
The data can be found here: http://spia.uga.edu/faculty_pages/rbakker/couart2.dta
The variables are as follows (you must use them all):

   art Outcome variable–number of articles published in last 3 years of PhD.

   fem Dummy for gender. 1 = female.

   mar Dummy for marital status. 1 = married.

   kid5 Number of children

   phd Prestige score of PhD granting institution (higher = more prestigious)

   ment Number of articles published by student's mentor in last 3 years

   Present the results of this model in a table including the coefficients, the standard errors, and any additional information you would like. What can you conclude from the t-ratios associated with each coefficient? What can you conclude from the model fit?

   Also, please present graphs of predicted counts against all covariates.

   Finally, discuss whether or not you think Poisson regression is the appropriate technique for these data and justify your answer. If no, then discuss other options and why they may be more appropriate.