Part I

1. In a large class, the correlation between midterm scores and final scores is found to be nearly .65, every term. The scores are normally distributed. Predict the percentile rank on the final for a student whose percentile rank on the midterm is:

   (a) 10%
   (b) 70%
   (c) 50%

2. A legislative chamber is split 53%/47% between the majority and minority parties. If members are assigned at random to a nine-member committee, what is the probability that the minority party makes up a majority on the committee?

3. Suppose you are interested in examining the factors that explain whether industrialized countries adopt a particular type of environmental regulation. Since countries adopt their policies at different points in time, you collect data for each country over a 20 year period of time. You also collect data on exogenous variables that will help you explain whether or not a given country adopts the policy in a given year. What estimation problems are you likely to encounter with these data, and how would you deal with them?

4. All statistical models require assumptions in order to identify and estimate parameters. Think about the problem of getting good estimates of the parameters of a model. What assumptions are most important for doing so? DO NOT go through a list of all assumptions underlying regression or other multivariate models. State specifically which assumptions you think are most important for making valid inferences and why.

5. Suppose you are interested in studying vote choice in a country with three major parties: a left-wing party, a right-wing party and a center party. What options are there to estimate models of vote choice in a situation like this? What are their advantages and disadvantages? Suppose a researcher were only interested in examining predictors of the choice between the left and right parties. To simplify the analysis she excludes all people who voted for the center party and runs a dichotomous logit equation. Is this a valid approach? Why?

6. Over the last five years, statistical matching has become much more commonly applied in political science. What problem is matching intended to solve? What are the advantages and disadvantages of matching compared to alternative solutions to the problem? Under what conditions would you expect matching to work well? Under what conditions would you expect matching to not work well?
Part II

Either submit an empirical research paper along with the exam or schedule an oral exam after the written exam.