Political Science 100a/200a
Statistical Analysis for Political Science, part I

Fall Quarter 2001

Instructor: James D. Fearon
Encina West, Rm. 441; 5-1314; jfearon@stanford.edu
Office Hours: Wed 2-4pm

TAs: Juliana Bambaci
Junga Kim

This course is a graduate-level introduction to data analysis for political scientists. Why not just take statistics from statisticians, you might ask? There are two main reasons. First, this course is designed for students with a broad range of prior exposures to the subject, which is typically the case for incoming political science graduate students. The course assumes *no* prior mathematical or statistical knowledge, except for some recollection of algebra. Nonetheless, we will cover certain topics unlikely to be stressed in a really basic introductory statistics course, and I will often "show you the math" in the lectures even if you will not be required to reproduce or extend it in problem sets or on the final. The second difference between this course and a "Statistics 100" is that this course is political science flavored. It will rely on political science examples, data sets, and preferred techniques, and is informed overall by the methodological styles and implicit philosophies typical of current practice in political science.

Prerequisites: None, as noted above, except for algebra. Some memory of basic calculus will be helpful, since I will cover some material that uses it. However this is certainly not necessary (for example, the main text uses no calculus and barely any algebra).

Requirements: Four problem sets and one take-home final. I strongly encourage the formation of 2-4 person study groups for mutual support on the problem sets, as this seems to facilitate learning.

Course website

`http:\\www.stanford.edu\class\polisci100a`

Course materials

The books listed below are available at the Stanford Bookstore and should also be on reserve at Green. There is no reader for the course; however, there will be a few articles assigned that you will need either to download from JSTOR or make copies of at the library.

- Achen, Christopher. 1982. *Interpreting and Using Regression*. Sage.
- Bulmer, M.J. 1967. *Principles of Statistics*. Dover.
- Fiorina, Morris. 1989. *Congress: Keystone of the Washington Establishment*. Yale University Press.
- Freedman, David, Robert Pisani, and Roger Purves. 1998 (3rd edition). *Statistics*. Norton.
- King, Gary, Robert O. Keohane and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press.

Computer-related stuff

All of the assignments will involve using Stata 7, a statistics package loaded onto the several machines in the political science computer cluster. (Stata is also available on the unix machines, although the interface is less attractive and many of nice graphics are not available unless you set up an X-Windows client, which we will tell you to do if this is your preferred means of access.) We will teach basic Stata as we go, sometimes using sections expressly for this purpose. The course website (link) will have a section for data sets to be used in the problem sets, as well as links to sites with downloadable data sets.

Topics and readings

Week 1: Sept. 26

1. Course Introduction and conceptual overview
    - Some useful background to a problem discussed in the lectures is Henry Farber and Joanne Gowa, "Polities and Peace," *International Security*, Fall 1995.

Week 2: Oct. 1 and 3

1. Conceptual overview continued, plus examples and exemplars
    - FPP (Freedman et al.), chaps. 1 and 2.

- Durkheim, *Suicide*, Book I, chap. 3, Book II, chaps. 1-3.
- Fiorina, *Keystone*, chaps. 1-6, and 10.

2. cont.

   - Edward Muller and Mitchell Seligson. 1994. Civic Culture and Democracy: The Question of Causal Relationships. *American Political Science Review* 88, 3 (September), 635-652. Download from JSTOR.
   - Alan S. Gerber and Donald P. Green, "The Effects of Personal Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment," *American Political Science Review* 94, 3 (September 2000), 653-665.

Week 3: Oct. 8 and 10

1. Describing data: graphs and summary statistics
2. cont., plus some math review

   - FPP, chaps. 3-9.
   - Bulmer, chap. 4.
   - KKV (King, Keohane and Verba), chaps. 1-2.
   - Achen, Christopher. 1977. Measuring Representation: Perils of the Correlation Coefficient. *American Journal of Political Science* 21, 4 (November), 805-815.

Week 4: Oct. 15 and 17

1st problem set due on Monday.

1. Probability: set theory, axioms, and counting

   - FPP, chaps. 13-15.
   - Bulmer, chap. 1-2.

2. continued: conditional probability, Bayes' rule

Week 5: Oct. 22 and 24

1. More probability, random variables, utility theory as an example.

2. cont., density functions, cdfs, and some probability distributions
    - Bulmer, chaps. 3, 5, 6.
    - Handout reading on distributions

Week 6: Oct. 29 and Oct. 31

2nd problem set due Monday.

1. The central limit theorem
    - FPP, chaps. 16-18.
    - Bulmer, chap. 7.
2. Sampling, surveys, and case selection
    - FPP, chaps. 19-23.
    - KKV, chap. 4.

Week 7: Nov. 5 and 7

1. Hypothesis testing
    - FPP, chaps. 26-27.
    - Bulmer, chaps. 8-9.
2. continued
    - FPP, chaps. 28-29.

Week 8: Nov. 12 and 14

3rd problem set due Monday.

1. Ordinary Least Squares
2. cont.
    - FPP, chaps. 10-12.
    - Bulmer, chap. 12.

- Achen, pages 1-37.
- KKV, chap. 3.

Week 9: Nov. 19 and 21

1. Multivariate regression
2. cont.
    - Achen, pages 37-68.
    - KKV, chaps. 4 and 5.

Week 10: Nov. 26 and 28

4rth problem set due Monday.

1. cont.
2. Review