

STATISTICAL COMPUTING
POLITICAL SCIENCE 493B

This course will focus on aspects of statistical computing. Modern statistical packages provide many tools for analyzing data. One purpose of this course is to take a behind-the-scenes look at some of these tools in an effort to move away from a simple point-and-click environment to a better understanding of some of these powerful tools. In addition, we will examine how advances in computing have led to advances in statistical analysis through tools such as simulation. A variety of statistical methods will be discussed in the context of specific data analysis problems. Topics may include distributions and random data, numerical methods, Monte Carlo experiments, and optimization (via e.g., maximum likelihood).

Although there are no formal prerequisites for this course, students should be comfortable using computers and should be conversant with some statistical software package. Although I prefer to do my work in a UNIX environment, you are welcome to work in either a UNIX environment or a windows environment. This course will emphasize the S-plus statistical package primarily through its free counterpart, R. Although I will not spend much time discussing implementation in other statistical package, you should feel free to use and complete assignments using any package that is appropriate for the tasks.

This syllabus as well as some class notes, announcements, and problem sets will be available at the web site <http://cho.pol.uiuc.edu/ps493/>. This web site will be updated regularly as the course progresses. Data sets for the computer assignments and various other items of interest will also be available at the class web site.

Course Requirements.

Grades will be based on computer projects. You are highly encouraged to work in groups. There will be no exams. Assignments should be typeset, preferably in \LaTeX , and e-mailed or otherwise electronically submitted to me. Handwritten assignments will not be accepted.

Useful Texts. Although none of these books is required for the course, you may find these books to be helpful as references.

- Lamport, Leslie. 1994. *TEX: A Document Preparation System: User's Guide and Reference Manual, 2d Edition*. Addison Wesley Longman, Inc.
- Venables, W.N., Brian D. Ripley, and William N. Venables. 1999. *Modern Applied Statistics with S-PLUS*. New York: Springer Verlag.
- Hamilton, Lawrence C. 1997. *Statistics with Stata 5*. Wadsworth.

Course Outline

Readings and assignments are listed below. I would like you to e-mail me your assignments. While I will accept these in any electronic format, I highly encourage you to produce your assignments in \LaTeX . We will spend the first class going over \LaTeX basics, and the web site has various \LaTeX example files that you are free to copy and modify.

There are no assigned readings from the texts. Rather, the texts are to be used as references. In general, I will cover the necessary material from the texts during the lectures, but the texts will have more detailed and additional material that you may want to consult. In addition, there is a Texlive CD that should be available at each site. The CD has freeware software on it that will allow you to compose documents in \LaTeX format on a windows platform.

Week 1. \TeX , Postscript, Pdf.

October 26. Creating Pretty Technical Documents. Texlive CD.

Example: \LaTeX Template Files (available online).

Week 2. S-plus and R.

November 2. S-plus and R Basics. Reading Data. Functions.

Example: Florida Voting Data.

Week 3. S-plus and R.

November 9. Branching and Loops. Distributions. Generating Random Data. Matrix Regression.

Example: Splus handout.

Week 4. Monte Carlos.

November 16. Simulations. Data Generating Processes.

Example: Monte Carlos from "Iff the Assumption Fits..."

Week 5. Off for Thanksgiving, November 23

Week 6. Bootstrapping and Jackknifing.

November 30. Bootstrap and Jackknife

Example: Simulated Data.

Week 7. Optimization

December 7. Maximum Likelihood.

Week 8. Review.

December 14. Open-ended (Review or topics suggested by students).