POL 571: Quantitative Analysis I

Spring 2006

Kosuke Imai (Instructor)
Office: Corwin Hall 041
Office hours: anytime
Phone: 609–258–6601, Fax: 973-556-1929
Email: kimai@Princeton.Edu
URL: http://imai.princeton.edu

Andrew Owen (Preceptor)
Office: Robertson 322
Office hours: TBA
Phone: 609–258–0091
Email: aowen@Princeton.Edu

1 Logistics

Lectures: Tuesdays and Thursday from 10:00 to 11:20 in 006 Robertson.
Precepts: TBA.

2 Course Mailing List

Send questions about lectures and problem sets to POL571_S2006@Princeton.Edu Andrew also has weekly office hours and precepts where he can answer any question about problem sets.

3 Description

This course is the first course in applied statistical methods for social scientists. Students will learn how statistical methods can be used to conduct causal inferences, exploratory data analysis, forecasting, and hypothesis testing. The first half of the course will be devoted to probability theory, which serves as a foundation of statistical theory. The second half covers statistical inference in general, the linear model in some depth, and if time permits also introduces generalized linear models. An emphasis of the course is given to practical data analysis, and students will learn statistical programming as well as basic principles of probability theory and statistical inference. This course assumes the mathematical knowledge taught in POL 502, and prepares students for the next course in the sequence, POL 572.

4 Textbooks

The course will be based on my lecture notes. I will also occasionally assign journal articles for you to read. However, you may be interested in purchasing the following main reference books for this course.

DeGroot, M. H. and Schervish, M. J. (2002). *Probability and Statistics*, Addison-Wesley.

Fox, J. (2002). *An R and S-PLUS Companion to Applied Regression*. Sage Publications.

You may also be interested in acquiring the following textbooks, which are more advanced than DeGroot and Schervish (2002):

Casella, G., and Berger, R. L. (2002). *Statistical Inference*. Duxbury Press.

Wooldridge, J. M. (2002) *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

5 Outline of the Course

I plan to cover the following topics in the order that they are listed below with the corresponding chapters in DeGroot and Schervish (2002). The teaching principle I follow is that we proceed as fast as possible given that *everyone* in the class is understanding the materials. So, never hesitate to ask questions during the class! Weekly sections should also provide us with the opportunity to make sure that everyone is keeping up with the course.

5.1 Probability Theory

1. Probability and Independence (Chapters 1 and 2)
2. Random Variables and Their Distributions (Chapters 3 and 5)
3. Expectation and Moment Generating Functions (Chapter 4 and 5)
4. Convergence of Random Variables and Limit Theorems (Chapter 5)

5.2 Statistical Inference

1. Statistical Inference: An Overview
2. Point Estimation (Chapter 6)
3. Interval Estimation (Chapter 7)
4. Hypothesis Testing (Chapter 8)
5. Regression Models (Chapter 10)
6. Simulation Methods (Chapter 11)

6 Course Requirements

The final course grade is the weighted average of the following three components.

1. Problem Sets (40%). There will be approximately weekly problem set. It will include analytical questions, statistical programming exercises, and data analysis problems. Although you are strongly encouraged to form a study group and discuss the materials together, you should not simply copy someone else's computer code or write-ups. Each problem set will be equally weighted.
2. Mid-term Exam (30%). The midterm exam on probability will be held right after the spring break. It is a closed book exam except that you are allowed to bring a letter-size paper on which contains whatever information you think is useful.

3. Final Project or Final Exam (30%). You have an option of conducting a final project or taking a final exam.

   (a) The final project should consist of the following three steps: (1) come up with a research question of interest (one page report due the end of February), (2) acquire a data set (one page report due the end of March), and (3) analyze the data set as an effort to answer the question. Since our knowledge of statistical analysis is still limited at the end of this course, the analysis can be largely exploratory by fitting simple regression models as well as effectively summarizing the data numerically and graphically. The brief report summarizing the results of your final project is due in my Corwin mailbox at 4pm, May 22. No late submission will be accepted.

   (b) The final exam is scheduled to be held on May 22. It will be based on the materials that are covered after the midterm exam.

7 Computing and Typesetting

In this course, we use a statistical computing environment, called R. R is available for any platform and without charge at http://www.r-project.org/; it is a clone of commercial software S-plus. Many of the models we cover in this class are also included in easy-to-use software *Zelig: Everyone's Statistical Software* available at http://gking.harvard.edu/zelig.

To write up the final paper, I recommend using LaTeXrather than MS Word. It is freely available for any platform (Mac OS X and Unix users can install tetex, and Windows users can download MikTeX). For programming (and if you are writing papers in LaTeX), you need a good text editor: emacs and Xemacs are well suited for this purpose and freely available. They are readily available for Linux and Max OS X users. An alternative option for Windows users is to use WinEdt (available at http://www.winedt.com/), which is commercial software and works well with both R and LaTeX. OIT and Politics clusters have WinEdt, MikTeX, and R installed. The following websites give you the installation instructions of these software for Windows machines,

- Xemacs, ESS (Emacs Speaks Statistics), and R: http://socserv.mcmaster.ca/jfox/Books/Companion/ESS/
- emacs, ESS, LaTeX, etc.: http://www.math.aau.dk/~dethlef/Tips/introduction.html
- WinEdt and LaTeX: http://www.winedt.com/tex.html)

There are many other useful online resources available.