

Theoretical and Methodological Foundations

Empirical Implications of Theoretical Models

Saturday, June 24, through Wednesday, June 28, 2017

James Lo

University of Southern California

Course Description

The basic motivation of EITM is that it pays to be more conscious and purposeful about how game theoretic models and statistical tests are connected. Unfortunately, game theory is often taught in a manner that submerges its relationship to testing, and statistical training typically begins with the hypothesis for testing having already been developed. The Foundations Seminar presents statistical methods that lend themselves well to testing predictions from formal models, such as probability models, general maximum-likelihood methods, and the tools of Bayesian statistics.

Logistics

The R refresher will take place Saturday and Sunday afternoon, while the main course runs both in the morning and afternoon on Monday to Wednesday. Morning classes will run from 9:30 am - 12 pm, while afternoon classes run from 2:30 pm - 5 pm. Readings and homework assignments to illustrate each session's techniques, to be discussed on the following day, will be distributed over email.

I am available during the Foundations seminar week for out-of-class discussion, and my email address is lojames@usc.edu.

Background Readings

To make the best use of our brief time in the Foundations course, we will assume all participants are familiar with certain basics of statistics and game theory. The background readings

listed below are designed as a brush-up for any who may need it, and should definitely be examined prior to participants' arrival at EITM.

Moreover, students should read in advance the material covered in each unit of the course schedule. Time during the course can thus be spent re-reading important materials, and implementing some of the techniques developed in the class. "Further readings," however, can be postponed indefinitely.

For both background and covered readings, I have provided guidance to what is important. We will draw from these readings throughout the sessions. In the course schedule below, note the particularly salient readings for the session.

Background Readings: Mathematical Statistics

DeGroot, Morris H., and Mark J. Schervish. 2002. *Probability and Statistics, Third Edition*. Reading, MA: Addison-Wesley.

This is an excellent introduction to mathematical statistics. It covers the basics of probability and distribution theory, statistical inference (from both frequentist and Bayesian perspectives), and simulation. Students should focus on Chapters 1-4 (probability and distribution theory), Chapter 6 (on Bayesian inference and maximum likelihood), Chapters 7 and 8 (on confidence intervals and hypothesis testing), and Chapter 11 (on simulation). Except for Section 11.4, the sections noted as optional can be skimmed.

R Programming Language

R Development Core Team. 2012. *An Introduction to R*. <http://www.r-project.org>.

We will be using the R Language for Statistical Computing throughout the Institute. R is an extremely flexible environment which can be used flexibly for many types of analyses, including simulation, model fitting, and graphing. R is also open-source free software, and is available for Windows, MacOS, and various versions of unix, and can be downloaded from the web. R is not a menu-driven canned statistical package, but rather an environment that is easily used to perform all kinds of analysis. Of course many commonly used tools have been implemented in R, but the real utility of the language is its ability to easily implement non-standard models and simulations by hand. All students should work through this document to familiarize themselves with R before arriving at EITM. However, as part of this course, we will be providing an optional refresher in R for students who are less familiar with it.

Course Schedule

Saturday afternoon: Introduction to R (Optional)

- Operations on different data types (scalars, vectors, matrices)
- Control statements (if(), for())
- Basic graphics
- Writing functions

Sunday afternoon: Introduction to R, part II (Optional)

- Data analysis
- Distribution functions
- Optimization

Monday morning: Probability and probability models

- Random Variables and Distributions
- Joint and Conditional Distributions
- Bayes Theorem
- Probability Models

Readings: DeGroot and Schervish (2000), Chapter 3, all; Chapter 5, Sections 1, 2, 4, 6, 9, and 12. Skim Chapters 1, 2, and 4.

Monday afternoon: Maximum likelihood methods

- The logic of maximum likelihood
- Statistical inference and maximum likelihood
- Maximum likelihood for a battle of the sexes game

Readings: DeGroot and Schervish (2000), Chapter 6, Sections 1, 5, and 6; Chapter 10, Sections 1-3. Skim remainder of Chapter 6, and Chapters 7 and 8.

Tuesday morning: Bayesian statistics

- Introduction to Bayesian statistics
- Model adequacy and non-nested model testing

Readings:

- DeGroot and Schervish (2000), remainder of Chapter 6
- Andrew D. Martin. 2008. “Bayesian Analysis.” *The Oxford Handbook of Political Methodology* (Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, eds.). Oxford: Oxford University Press.
- Edward Greenberg. 2008. *Introduction to Bayesian Econometrics*. Cambridge University Press, Chapters 2-4.
- Raftery, Adrian. 1995. “Bayesian model selection in social research (with Discussion).” In *Sociological Methodology 1995*, (Peter V. Marsden, ed.), Cambridge, MA: Blackwells, pp. 111-196.
- Robert E. Kass and Adrian E. Raftery. 1995. “Bayes factors.” *Journal of the American Statistical Association*. 90: 773-795.

Tuesday afternoon: Methods for Bayesian estimation

- Markov Chain Monte Carlo simulation
- The Gibbs sampler
- Metropolis-Hastings algorithm

Readings:

- Greenberg (2008), Chapters 4, 5, 7, and 8.
- Casella and George. 1992. “Explaining the Gibbs Sampler.” *The American Statistician*. 46: 167-174.
- Chib and Greenberg. 1995. “Understanding the Metropolis-Hastings Algorithm.” *The American Statistician*. 49: 327-335.

Wednesday morning: Applied Bayesian statistics

- Using JAGS
- Analysis and replication of Shipan's 2004 article

Readings:

- Charles Shipan. 2004. "Regulatory Regimes, Agency Actions, and the Conditional Nature of Congressional Influence." *American Political Science Review* 98(3): 467-480. Special attention to "Estimating Political Influence of the FDA," pp. 472-478.
- Martyn Plummer. 2012. *JAGS Version 3.2.0 user manual*.
- Andrew D. Martin, Kevin M. Quinn, and Jong Hee Park. 2011. "MCMCpack: Markov chain Monte Carlo in R." *Journal of Statistical Software*. 42(9): 1-21.

Wednesday morning: Introduction to Bayesian Measurement

- An intuitive introduction to measurement applications in political science

Readings:

- Keith Poole and Howard Rosenthal. 1991. "Patterns of Congressional Voting." *American Journal of Political Science* 35: 228-78.
- Simon Hix, Abdul Noury, and Gerard Roland. 2007. *Democratic Politics in the European Parliament*. Cambridge: Cambridge University Press.