Simon Jackman
Department of Political Science
Encina Hall West, Room 315
e-mail: jackman@stanford.edu
http://jackman.stanford.edu

James Holland Jones
Department of Anthropological Sciences
Building 360 (Main Quad)
e-mail: jhj1@stanford.edu
http://www.stanford.edu/~jhj1

AnthSci 254/PoliSci 354F: Applied Bayesian Analysis

This class provides a practical introduction to Bayesian statistical inference, with an emphasis on applications in the social sciences. We will begin slowly, with a consideration of how Bayesian statistical inference differs from classical or frequentist inference. We will examine these differences in the context of simple statistical procedures and models: e.g., one and two-sample t-tests, the analysis of a two-by-two tables, one-way ANOVA, and regression.

We then show how a the relatively recent explosion in desktop computing power has made Bayesian approaches attractive for more complex models. Specifically, the set of algorithms known as Markov chain Monte Carlo (MCMC) allow researchers to tackle classes of problems that used to fall in the "too hard" basket. Today, MCMC is well and truly part of the statistical computing toolkit available to social scientists, and implemented in various forms in many different software packages (we will survey some of these, see below). We will examine how these algorithms make Bayesian inference feasible, their strengths and weaknesses, and some of the pitfalls to avoid when deploying MCMC algorithms. The applications to be considered in this part of the course include

- generalized linear models for binary and ordinal data
- models for latent variables, e.g., factor-analytical models, including structural equation model; item-response models;
- classification and clustering, both cross-sectionally and dynamically (i.e., change-point or "structural breaks")
- dynamic latent state models (e.g., tracking public opinion over time),
- hierarchical models of various flavors (appropriate to many forms of data in the social sciences),
- models for spatial data
- modes for lifetimes or events (possibly with latent variables such as correlated frailties between relatives)

The class presumes that students have had only an intermediate level of exposure to probability and statistical modeling; i.e., knowing what a likelihood function is and its role in classical inference would be a useful prerequisite.

Texts

There is no prescribed text for this class. One of us (Jackman) is completing a book-length treatment of the material covered in this survey. Selected draft chapters will be circulated to the class throughout the quarter. Until Jackman's book appears in print, we recommend

Gelman, Andrew, John B. Carlin, Hal S. Stern and Donald B. Rubin. 2004. *Bayesian Data Analysis*. 2nd edition. Chapman and Hall/CRC Press: Boca Raton, Florida.

We will be adding readings and chapters specific to the applications as we move into that part of the class.

The following books are more compact than the Gelman et al. text, and give emphasis to different aspects of Bayesian statistics, but are written at a fairly accessible level:

- Gill, Jeff. 2002. *Bayesian Methods: A social and behavioral sciences approach*. Chapman and Hall: New York
- Koop, Gary. 2003. *Bayesian Econometrics*. Wiley: Chichester.
- Lancaster, Tony. 2004. *An Introduction to Modern Bayesian Econometrics*. Blackwell: London.

Books at quite simple level of exposition include

- Bolstad, William M. 2004. *Introduction to Bayesian Statistics*. Wiley: Hoboken, New Jersey.
- Lee, Peter. 2004. *Bayesian Statistics: an introduction*. 3rd edition. Hodder Arnold: London.

Peter Congdon has written a couple of nice recipe books at an accessible level

- Congdon, Peter. 2001. *Bayesian Statistical Modelling*. Wiley: Chichester.
- Congdon, Peter. 2005. *Bayesian Models for Categorical Data*. Wiley: Chichester.

More advanced texts include

- Bernardo, José and Adrian F.M. Smith. 1994. *Bayesian Theory*. Wiley: Chichester.
- Geweke, John. 2005. *Contemporary Bayesian Econometrics and Statistics*. Wiley: Hoboken, New Jersey.
- Liu, Jun S. 2002. *Monte Carlo Strategies in Scientific Computing*. Springer: New York.

- Robert, Christian P. 2001. *The Bayesian Choice*. Springer: New York.
- Rossi, Peter, Greg Allenby and Rob McCulloch. 2006. *Bayesian Statistics and Marketing*. Wiley: Hoboken, New Jersey.
- Ibrahim, Joseph G., Ming-Hui Chen and Debajyoti Sinha. 2004. *Bayesian Survival Analysis*. Springer: New York.
- Banerjee, Sudipto, Bradley. P. Carlin and Alan E. Gelfand. 2003. *Hierarchical Modeling and Analysis for Spatial Data*. Chapman & Hall/CRC: Boca Raton, FL.
- Lawson, Andrew B., William J. Browne and Carmen L. Vidal Rodeiro. 2003. *Disease Mapping with WinBUGS and MLwiN*. Wiley: New York.

Software

Both Jackman and Jones use and recommend R; we strongly suggest that you familiarize yourself with the R Bayesian Inference taskview, describing facilities within R for Bayesian inference. We will also examine the free, general purpose Bayesian inference program OpenBugs (formerly `WinBugs`). We will also consider JAGS, another free, open-source, general purpose Bayesian analysis program. Both `OpenBugs` and `JAGS` require some programming.

Assessment

We will have

1. three or four homeworks, comprising data analysis and writeups
2. a final research paper (15-25pp), where you deploy Bayesian statistical methods in an interesting application

with roughly a 40-60 weighting given to the homeworks and the final paper, respectively.