Bayesian Methods for Political Science Data Analysis

PS 919, Spring, 2001
Instructor:
Charles Franklin, University of Wisconsin, Madison
franklin@polisci.wisc.edu
608-263-2022 (Office)
608-265-2663 (Fax)

This course introduces Bayesian methods for data analysis in political science. Bayesian methods provide a flexible and powerful approach to complex statistical models and have a theoretical elegance and clarity that is impressive. Bayesian models inherently recognize and incorporate subjective judgements of the researcher, which the source of both their great power the controversy surrounding their use. We will discuss some of the epistemological issues raised by Bayesian methods as well as their application.

We will cover the basic concepts of Bayesian statistical inference. There are several sets of tools needed to do applied Bayesian modeling. First we need to review some probability theory and learn some integral calculus to allow us to deal with continuous distributions. Second, we will develop the fundamental notion of Bayes theorem as a foundation for statistical inference. We'll also see how likelihood is incorporated within the Bayesian framework (brush off the MLE notes!) Finally, we'll explore the world of applied Bayesian modeling. This will include learning some new software tools using `WinBugs` and `S-Plus/R`.

The main focus of the course will be application of Bayesian models to cutting edge issues in political science. This will include a number of applied readings. We'll take time to discuss these applications in class in order to develop a feel for what research that takes a Bayesian approach "feels" like.

The goal of the class is to develop the necessary theoretical understanding to correctly apply Bayesian models using modern software. A second goal is to reach a level of understanding that will support further reading and learning on your own. The syllabus points the way to further reading throughout in the "Advanced Topics" sections.

Text

There are several good contemporary textbooks on Bayesian data analysis as well as several older classics. Each of these has strengths and weaknesses. I've chosen two for the class in the hopes that they will complement one another. In addition, the syllabus lists a number of books for alternative treatments.

The primary texts for the course are

- Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 1995. *Bayesian Data Analysis*, New York: Chapman & Hall.
- Thomas Leonard and John S. J. Hsu. 1999. *Bayesian Methods*, New York: Cambridge University Press.

A good alternative to the Gelman text is

- John B. Carlin and Thomas A. Louis. 2000. *Bayes and Empirical Bayes Methods for Data Analysis*, second edition. New York: Chapman & Hall.

This book is a bit stronger on the technical issues of estimation and Markov-Chain Monte-Carlo methods, but is perhaps a little less good on the applications. We'll read several chapters.

The Gelman et. al. text is now widely used in statistics departments. It is the more demanding of the two texts, and it can be criticized for not providing enough worked examples and derivations. On the other hand, it covers both theory and application well. There are even political science examples!

The Leonard and Hsu book is, in contrast, full of excellent examples and complete derivations. This is definitely the book to read *first* before tackling Gelman. The potential drawback is that Leonard and Hsu tend to treat most topics a bit lightly rather than devoting a lot of effort to getting to the details. I believe the two complement one another nicely.

We will also read all or most of several papers and manuscripts by Simon Jackman (Political Science, Stanford University). Jackman is the current leading advocate for Bayesian methods in political science and his papers make every effort to remain accessible and substantively interesting.

In addition most weeks will include articles applying the methods we are discussing. The syllabus includes sections of historical or advanced topics which can be read at leisure.

Course Requirements

There will be frequent homework exercises. Some of these will be theoretical but most will focus on applying Bayesian models to actual data analysis. There will be between 6 and 10 such exercises in all.

There is also a required term paper in which the student will develop a Bayesian model of a substantive topic of their choosing. You should consult with the instructor about your interests and the specific application you will develop.

I will also expect everyone to be prepared to discuss readings and to collectively discuss problems in class.

The sum of the homework and the term paper will be equally weighted in the final grade. Lack of consistent participation in class will lower the final grade.

Prerequisites

I assume everyone has completed PS 818, Maximum Likelihood Estimation (or the equivalent), and therefore has a good grasp of likelihood functions, differential calculus, and discrete probability distributions. And, of course, matrix algebra.

The level of mathematics required is not extremely high, but we will need integral calculus for an understanding of the continuous distributions which are most common in Bayesian models. I assume students once had a one-semester course in calculus that covered univariate integrals. We will spend time in extra sessions to cover the necessary background for integration, though I assume everyone is familiar with differential calculus.

We also must make extensive use of continuous probability distributions. I assume most students have limited familiarity with these topics, and will again spend time in extra sessions to cover these topics.

Computation

The vast majority of our computation will be done using the program `WinBugs` (version 1.3) under the Windows operating system. this software is free, and can be downloaded from `http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml`.

It is also possible to use the `R` (or `S-Plus`) statistical software system for Bayesian analysis. `R` or `S-Plus` is particularly useful for post-estimation diagnostics of Bayesian models using either `Coda` or `Boa`. The `R` system is also freely available and can be downloaded from `http://cran.r-project.org/` and `http://www.r-project.org/`. The computers in the political science lab all have the `S-Plus` system installed. Time spent mastering `R/S-Plus` will pay long-term benefits as well because these extremely powerful programs are becoming increasingly used for advanced statistical applications, whether Bayesian or not.

I have put several relevant manuals in the computer lab. These include the `Bugs`, `Coda` and `Boa` manuals, the two volumes of `Bugs` examples (bound together), and two sets of manuals introducing `R`. Further, the `S-Plus` manuals in the lab can be used for both `S-Plus` and for `R`. I've made two copies of each of these for the Lab. PLEASE do not take them further than the Reading Room, and do NOT take them from the building. You can download and print all of these manuals for your personal use if you wish. Links at the `Bugs` and `R` home pages will lead you to the files.

Background Readings

For those interested in some guidance concerning review materials, the most important resources are

Kleppner, Daniel and Norman Ramsey. 1985. *Quick Calculus,* 2nd edition. New York: John Wiley. Chapter 3: Integral Calculus, pp. 151-207.

Any good upper level probability text, for example Morris H. DeGroot. 1986. *Probability and Statistics*, second edition. Reading MA: Addison-Wesley.

Thomas Leonard and John S. J. Hsu. 1999. *Bayesian Methods*, New York: Cambridge University Press. Chapter 1. (This is a review of likelihood.)

Evans, Merran, Nicholas Hastings, and Brian Peacock. 1999. *Statistical Distributions,* Third Edition. New York:John Wiley. (An excellent and (relatively) inexpensive handbook and guide to a variety of probability distributions. Because we'll be dealing with a number of new distributions, you'll want to familiarize yourself with them.)

Syllabus

1 Introduction to Bayesian Models

1.1 Core Reading

Read Western and Jackman 1994 for an interesting example of Bayesian models applied to substantive questions. Read Jackman 2000 very lightly. This is a technical introduction to topics we will spend the entire semester on. Browse it now to get a feel for what lies ahead.

Western, Bruce and Simon Jackman. 1994. "Bayesian Inference in Comparative Research," *American Journal of Political Science* 88: 412-423.

Jackman, Simon. 2000. "Estimation and Inference via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo." *American Journal of Political Science* 44:375-404.

1.2 Alternate Sources

For a nice overview of Bayesian models, see:

John B. Carlin and Thomas A. Louis. 2000. *Bayes and Empirical Bayes Methods for Data Analysis*, second edition. New York: Chapman & Hall. Chapter 1.

To see the original arguments that Western and Jackman address see:

Wallerstein, Michael. 1989. "Union Organization in Advanced Industrial Democracies." *American Political Science Review* 83:481-501.

Stephens, John and Michael Wallerstein. 1991. "Industrial Concentration, Country Size and Trade Union Membership." *American Political Science Review* 85:941-53.

1.3 Historical Background

Barnard, G. A. 1958. "Studies in the History of Probability and Statistics: IX. Thomas Bayes's Essay Towards Solving a Problem in the Doctrine of Chances." *Biometrika* 45: 293-295.

Bayes, Thomas. 1763/1958. "Studies in the History of Probability and Statistics: IX. Thomas Bayes's Essay Towards Solving a Problem in the Doctrine of Chances." *Biometrika* 45: 296-315.

2 Basic Probability and Bayes Theorem

2.1 Core Reading

Read Wonacott & Wonnacott for review. Be sure you completely understand the basic concepts. Then read Greene. You will need to understand Greene.

Greene, William H. 2000. *Econometric Analysis*, fourth edition. Upper Saddle River, NJ: Prentice Hall. Chapter 3. Pay special attention to 3.2, 3.3, 3.5, 3.6 and 3.7.

Thomas H. Wonnacott and Ronald J. Wonnacott, *Introductory Statistics*, 5th edition. Chapters 3, 4, 5.1, 5.2.

2.2 Alternate Sources

In order of increasing sophistication.

Berry, Donald A. 1996. *Statistics: A Bayesian Perspective*. New York: Duxbury Press. Chapters 4 & 5.

DeGroot, Morris H. 1986. *Probability and Statistics*, second edition. Reading MA: Addison-Wesley. Chapters 1 & 2.1-2.2.

DeGroot, Morris H. 1970. *Optimal Statistical Decisions.* New York: McGraw Hill. Chapters 2 & 3.

3 Basic Bayes

3.1 Core Reading

This section introduces the fundamentals of Bayesian modeling.

Thomas Leonard and John S. J. Hsu. 1999. *Bayesian Methods*, New York: Cambridge University Press. Chapter 2.

Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 1995. *Bayesian Data Analysis*, New York: Chapman & Hall. Chapter 1.

3.2 Alternate Sources

Berry, Donald A. 1996. *Statistics: A Bayesian Perspective*. New York: Duxbury Press. Chapters 6-11.

John B. Carlin and Thomas A. Louis. 2000. *Bayes and Empirical Bayes Methods for Data Analysis*, second edition. New York: Chapman & Hall. Chapter 2.

3.3 Application

Western, Bruce. 1994. "Unionization and Labor Market Institutions in Advanced Capitalism, 1950-1985." *American Journal of Sociology* 99:1314-1341.

4 Single Parameter Models

4.1 Core Reading

Thomas Leonard and John S. J. Hsu. 1999. *Bayesian Methods*, New York: Cambridge University Press. Chapter 3.

Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 1995. *Bayesian Data Analysis*, New York: Chapman & Hall. Chapters 2 & 18.

5 Multi-parameter Models

5.1 Core Reading

Thomas Leonard and John S. J. Hsu. 1999. *Bayesian Methods*, New York: Cambridge University Press. Chapter 5.

Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 1995. *Bayesian Data Analysis*, New York: Chapman & Hall. Chapter 3.

6 Empirical Bayes

6.1 Core Reading

John B. Carlin and Thomas A. Louis. 2000. *Bayes and Empirical Bayes Methods for Data Analysis*, second edition. New York: Chapman & Hall. Chapter 3.

7 Computation: MCMC, Gibbs Sampling, Metropolis-Hastings and all that

7.1 Core Reading

Jackman, Simon. 2000. "Estimation and Inference via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo." *American Journal of Political Science* 44:xxx-zzz.

John B. Carlin and Thomas A. Louis. 2000. *Bayes and Empirical Bayes Methods for Data Analysis*, second edition. New York: Chapman & Hall. Chapter 5.

8 Application

Quinn KM. Martin AD. Whitford AB. 1999. "Voter choice in multi-party democracies: A test of competing theories and models." *American Journal of Political Science* 43:1231-1247.

8.1 Alternate Sources

Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 1995. *Bayesian Data Analysis*, New York: Chapman & Hall. Chapters 9-11.

8.2 Advanced Topics

Chen, Ming-Hui, Qi-Man Shao and Joseph G. Ibrahim. 2000. *Monte Carlo Methods in Bayesian Computation.* New York: Springer. Especially chapters 2-3.

Robert, Christian P. and George Casella. 1999. *Monte Carlo Statistical Methods.* New York: Springer. Especially chapters 4-7 (Oh, hell, just read the whole book!)

9 Regression Models

9.1 Core Reading

Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 1995. *Bayesian Data Analysis*, New York: Chapman & Hall. Chapter 8.

10 Hierarchical Models

10.1 Core Reading

Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 1995. *Bayesian Data Analysis*, New York: Chapman & Hall. Chapter 5.

10.2 Application

Western, Bruce. 1998. "Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modeling Approach." *American Journal of Political Science* 42:1233-59.

11 Model Checking and Sensitivity Analysis

11.1 Core Reading

Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 1995. *Bayesian Data Analysis*, New York: Chapman & Hall. Chapter 6.

11.2 Alternate Sources

John B. Carlin and Thomas A. Louis. 2000. *Bayes and Empirical Bayes Methods for Data Analysis*, second edition. New York: Chapman & Hall. Chapter 6.

12 Robust Inference and Sensitivity Analysis

12.1 Core Reading

Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 1995. *Bayesian Data Analysis*, New York: Chapman & Hall. Chapter 12

13 Hierarchical Linear Models

13.1 Core Reading

Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 1995. *Bayesian Data Analysis*, New York: Chapman & Hall. Chapter 13.

14 Generalized Linear Models

14.1 Core Reading

Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 1995. *Bayesian Data Analysis*, New York: Chapman & Hall. Chapter 14.

15 Mixture Models

15.1 Core Reading

Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 1995. *Bayesian Data Analysis*, New York: Chapman & Hall. Chapter 16.

16 Models for Missing Data

16.1 Core Reading

Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin. 1995. *Bayesian Data Analysis*, New York: Chapman & Hall. Chapter 17.

17 Subjective Probability and Decision Theory

17.1 Core Reading

DeGroot, Morris H. 1970. *Optimal Statistical Decisions.* New York: McGraw Hill. Chapter 6.

Thomas Leonard and John S. J. Hsu. 1999. *Bayesian Methods*, New York: Cambridge University Press. Chapter 4.

17.2 Alternate Sources

DeGroot, Morris H. 1970. *Optimal Statistical Decisions.* New York: McGraw Hill. Chapters 7 & 8.

17.3 Really Advanced Topics

Bernardo, José M. and Adrian F. M. Smith. *Bayesian Theory.* New York: John Wiley & Sons. Chapters 1-3.

Raiffa, Howard and Robert Schlaifer. 1960/2000. *Applied Statistical Decision Theory.* New York: Wiley Classics. (A reprinting of an early classic.)