Bayesian Statistics for the Social Sciences

Political Science 919 Spring 2017
Time 1:30–3:30pm (Wednesdays)
Dates Jan. 18, Jan. 25, Feb. 8, Feb. 15, March 1, March 8, March 29, April 5, April 19 & April 26

Instructor Alex Tahk
Email atahk@wisc.edu
Office 212 North Hall
Office Hours Mondays, 2–4pm, and by appointment
Note: Meetings can be held in person or via Google Hangouts (atahk@wisc.edu).
Please sign up for meetings during office hours in advance at: http://go.wisc.edu/07300x

Overview

This is a course on Bayesian statistics with applications to the social sciences. Bayesian statistics offers an alternative framework to maximum likelihood and other frequentist approaches to estimation and inference. With the huge increases in computational power over the last two decades, Bayesian statistics has become increasingly important and common in political science and other fields. It offers both philosophical and theoretical advantages, but its current popularity has primarily been due to practical advantages, as modern Markov-chain Monte Carlo methods and associated software can make estimation of many complex or otherwise difficult-to-estimate models quite practical.

Statistical computing

Computational components of the problem sets will make use of JAGS and R. No prior knowledge of JAGS is needed. Students may make use of other software, such as Stan or WinBUGS, in their projects.

JAGS, which performs automated Gibbs sampling of Bayesian models, can be downloaded from: http://mcmc-jags.sourceforge.net/

R, an implementation of the S statistical programming language, can be downloaded for free from: http://www.r-project.org/

Prerequisites

This course assumes you have taken classes on an introductory statistics class and a class on linear models (POLI SCI 812 and 813 or the equivalent) and have experience with statistical computing in R.

Textbooks

The primary textbook for this course is:

Much of the material is also covered in the following books:


For Markov-chain Monte Carlo methods:


**Grading**

Grading will be divided between problem sets (25%), a replication project (25%), partner comments (10%) and a final project (40%).

**Problem sets**

There will be short problem sets handed out in class, typically one every one to two weeks and due on the Wednesday of the following week (unless otherwise noted). All homework must be typewritten and submitted in PDF format through the course website. These will be graded on a check-plus/check/check-minus/zero basis. Late assignments are strongly discouraged. A pattern of late assignments will result in a grade penalty. Assignments more than one week late will not be accepted.

The problem sets will cover both theory and application. You are welcome to discuss the problem sets with each other, but the final write-ups, results and coding should be your own.

**Replication project**

The replication project consists of estimating the model of a recently published paper in a Bayesian framework. The paper you are replicating should not have used Bayesian methods. You will reimplement the same model, but do so in a Bayesian framework using JAGS, Stan, WinBUGS or NIMBLE. Please do not use canned estimating routines even if they exist for the model you are replicating.

The paper should include your JAGS, Stan, WinBUGS or NIMBLE code in an appendix. No introduction, literature review or theory section is needed. The paper should be approximately 5–10 pages in length, not including the appendix, and is due **March 5**.
**Final project**

The final project consists of original research using Bayesian methods. It can be an extension of the models in a published article, including the article used in the replication project, but cannot simply replicate these models. The project must be approved **Friday, March 31**. The final paper is due **Wednesday, May 3**.

Students will be assigned a partner, who will function as a co-author and provide written comments on one draft of the paper. This draft is due to both me and your partner on **Friday, April 21, at 5pm**. You should have your comments back to your partner by **the following Monday, April 24, at 5pm**.

The paper is expected to look like a draft of an empirical article to be submitted to a social science except that it should contain only the data analysis section of the paper describing the statistical model, data, results and conclusions. Do not include an introduction, literature review or theory section. A good summary of how to construct a paper is available here: [http://gking.harvard.edu/files/paperspub.pdf](http://gking.harvard.edu/files/paperspub.pdf)

The paper should follow the style of an article in the *American Political Science Review*. It should be no longer than 20 pages double spaced, but a suitable paper can be shorter. Do not include references in the page count, but do include tables and figures. Note that graphs often communicate information better than tables. A good discussion of this point is available here: [http://eduardoleoni.com/published/graphs.pdf](http://eduardoleoni.com/published/graphs.pdf)

While it will not affect your grade in this class, you are encouraged to apply to present a poster on your project at the Annual Meeting of the Society for Political Methodology, held July 13–15 in Madison. Because applications will be due before project approval is needed for this course, you should decide on your project early if you wish to apply. More information is available here: [http://polmeth.polisci.wisc.edu/](http://polmeth.polisci.wisc.edu/)

**Topics and readings**

The syllabus is organized around topics rather than by day. Topics marked with an asterisk are additional topics that we will discuss if time permits.

**Introduction to Bayesian statistics**

Jackman, chapter 1

**Introduction to JAGS**

Jackman, chapter 6.3

Karreth, “Using JAGS via R”

**Bayesian analysis for simple models**

Jackman, chapter 2

**Markov-chain Monte Carlo**

Jackman, chapter 3–6

**Hierarchical models**

Jackman, chapter 7
Choice models
Jackman, chapter 8

Measurement models
Jackman, chapters 9.1–9.2

Ideal-point estimation and item-response theory
Jackman, chapter 9.3
Clinton, Jackman and Rivers, “The statistical analysis of roll-call votes”

Dynamic measurement models
Jackman, chapter 9.4
Linzer, “Dynamic Bayesian forecasting of presidential elections in the states”

Change-point models
Park, “Change-point analysis of binary and ordinal probit models: An application to bank rate policy under the interwar gold standard”

Topic models
Grimmer and Stewart, “Text as data: The promise and pitfalls of automatic content analysis methods for political texts”
Quinn, Monroe, Colaresi, Crespin and Radev, “How to analyze political attention with minimal assumptions and costs”

Multilevel regression and poststratification*

Dirichlet process and nonparametric Bayes*
Gill and Casella, “Nonparametric priors for ordinal Bayesian social science models”
Spirling and Quinn, “Identifying intraparty voting blocs in the U.K. House of Commons”

Variational Bayes*
Grimmer, “An Introduction to Bayesian Inference via Variational Approximations”

Approximate Bayesian computation*
Marin, Pudlo, Robert and Ryder, “Approximate Bayesian computational methods”