Bayesian Modeling for the Social Sciences I: Introduction and Application
ICPSR Summer Program 2018

Ryan Bakker (Weeks 1–2)
Associate Professor
Department of Political Science
University of Georgia
rbakker@uga.edu

Johannes Karreth (Weeks 3–4)
Assistant Professor
Department of Politics & International Relations
Ursinus College
jkarreth@ursinus.edu

This syllabus will be updated prior to the course. Please visit www.jkarreth.net/bayes-icpsr.html for the most recent information on topics, labs, rooms, and assigned readings.

Workshop: 9am–11am / Room: See the Summer Program handbook.
Office: Helen Newberry Building
Office hours: 11am-12pm and 1pm-2pm, M–F. Feel free to stop by the office any time and come in if the door is open. We’re also happy to schedule meetings at most other times during the day.

Teaching assistants
Herrissa Lamothe
Princeton University
hlamothe@princeton.edu
Office: Helen Newberry Building
Office hours: TBD

Neil Williams
University of Georgia
snwill@uga.edu
Office: Helen Newberry Building
Office hours: TBD
Course description

This workshop introduces the basic theoretical and applied principles of Bayesian statistical analysis in a manner geared toward students and researchers in the social sciences. The Bayesian paradigm is particularly useful for the type of data that social scientists encounter given its recognition of the mobility of population parameters, its ability to incorporate information from prior research, and its ability to update estimates as new data are observed. The workshop begins with a discussion of the strengths of the Bayesian approach for social science data and the philosophical differences between Bayesian and frequentist analyses. Next, the workshop covers the theoretical underpinnings of Bayesian modeling and provides a brief introduction to the primary estimation algorithms. The bulk of the workshop focuses on estimating and interpreting Bayesian models from an applied perspective. Participants are introduced to the Bayesian forms of the standard statistical models taught in regression and MLE courses (i.e., linear, logit/probit, poisson, etc.). Additional topics include measurement models, model comparison, and an in-depth treatment of multilevel modeling.

Participants should have a solid understanding of the linear model and matrix algebra and some exposure to models with limited dependent variables. The Mathematics for Social Scientists III lecture with Don Eckford is an ideal companion lecture for this workshop.

The workshop relies mostly on R and WinBUGS/JAGS for estimation, with a short segment on Stan, a newer, but rapidly growing tool for Bayesian inference. Prior experience with R is preferred but not assumed or necessary. We offer several lab sessions to familiarize participants with R, WinBUGS, JAGS, and Stan.

Goals. Upon conclusion of this workshop, we aim for participants to be able to:

- appreciate the fundamental differences and similarities between frequentist and Bayesian approaches to inference
- apply Bayes’ rule to statistical inference
- formulate linear and generalized linear models in the Bayesian framework
- estimate linear and generalized linear models in the Bayesian framework using customizeable code
- exploit the advantages of Bayesian estimation with regard to
  - incorporating prior information
  - estimating uncertainty in parameter estimates
  - dealing with missing data
  - measuring latent concepts
  - incorporating variance at multiple levels of observation
- present and communicate results from Bayesian (and frequentist) estimation in an effective manner
- have fun learning new methods and better understanding familiar ones!

A note on computing. This workshop mostly uses JAGS and WinBUGS (and its open-source twin OpenBUGS) as the primary software options to fit Bayesian models, with one unit toward the end dedicated to Stan. We access JAGS and Win/OpenBUGS through R. Most lectures build on JAGS and Win/OpenBUGS. The languages of these two programs are nearly identical. Win/OpenBUGS run on Macs only with Windows emulation software, but can be a bit buggy. JAGS runs on all platforms, including Macs. We offer special Mac-friendly lab sessions and support both JAGS and Win/OpenBUGS. JAGS code for all models encountered in this workshop and other JAGS-specific code and examples are provided.

Workshop resources

M-Box: All slides, code used in workshop sessions, and problem sets will be posted on M-Box. Details will be provided on the first day of the workshop. The TAs are also always available to point you to these materials.
Google Group: We will ask you to sign you up for a Google Group for this workshop. We will use this Google Group to share information related to the workshop during the ICPSR Summer Program.

Website with additional materials: Additional code, tutorials, and other materials for weeks 3–4 are posted on Johannes' website: http://www.jkarreth.net/bayes-icpsr.html.

Reading materials

Books

The main required texts for this workshop are:


You may also find the following titles useful for many of the topics discussed in this workshop. They are available in the ICPSR Summer Program Library for borrowing:


As a general primer for R, we recommend:


Articles

All articles listed in the syllabus are made available to you on M-Box and through the University of Michigan library website from the campus network. Please contact the TAs if you have trouble locating any of the articles listed in the syllabus.

Software

This workshop relies mostly on R, JAGS/WinBUGS/OpenBUGS, and Stan. We may also briefly discuss Stata as an alternative for some applications. We provide assistance installing R and JAGS/WinBUGS/OpenBUGS on your computers during labs in the first week of the workshop. There is a learning curve for these programs, but you need not have any computer programming background to learn them rather easily—just patience and desire. Our goal is to make you as comfortable as possible with these programs by the end of this workshop so that you will be able to use them with ease at your home institutions and in your own work.

Mac and JAGS users: See Johannes’ website for more information on installing JAGS.
**Problem sets**

We assign problem sets in class. Our goal is to make sure participants receive sufficient feedback to complete all assignments successfully. We distribute between 2 and 4 assignments per week. They are mostly computer-based with the exception of the first assignment. We ask that you include all code you used to complete your assignments; the TAs will show you how to best do this. The TAs will aim to return graded assignments to you within 5 days with comments via email. We (the instructors and TA) are more than happy to provide help with assignments during office hours: don't be afraid to come by and ask.

**Labs**

We offer several labs with guided hands-on exercises. Lab sessions will be held at the times in the room(s) listed in the schedule below (watch for updates). Likely topics:

1. Installing and using R
2. Using R and RMarkdown for an integrated and reproducible workflow for Bayesian (and frequentist) statistics
3. Installing and accessing JAGS/BUGS from R
4. Obtaining convergence diagnostics using R
5. Model presentation
6. Using Stan

**Preparing for each workshop meeting**

To get the most out of this workshop, we recommend that you read the assigned background & textbook readings for each day in depth and skim at least one of the applied studies if any are listed. Particularly in the second half of the workshop, we provide a larger number of applied works—pick one that is closest to your area of interest.

You should also feel strongly encouraged to come to our office hours on any day of the workshop to follow up on topics discussed during workshop meetings and to discuss how any topic we discussed might relate to your own work.
Workshop content and schedule

The following dates and topics may be modified as the course proceeds. The most recent version of the syllabus will always be at www.jkarreth.net/bayes-icpsr.html.

Monday, June 25
No course meeting

Recommended: Introduction to the LATEX Text Processing System, 5:30pm–7:30pm.

Day 1: Tuesday, June 26
Introduction: Background and Basics of Bayesian Inference

Please read:

- Gill: Chapter 1.

Day 2: Wednesday, June 27
Review of Generalized Linear Models

Refresher:

- Gill: Section 2.2.
- Gelman & Hill: Chapter 6.

Day 3: Thursday, June 28
Probability and Bayes’ Rule

Please read:

- Gill: Chapter 2.
Day 4: Friday, June 29

Priors

Please read:

- Gill: Chapter 4.

**HW 1 assigned: Prior and posterior distributions.**

Day 5: Monday, July 2

Sampling Methods and Introduction to the BUGS/JAGS Language

Please read:

- Gill: Chapters 9 & 10.

Day 6: Tuesday, July 3

Convergence Diagnostics

Please read:


Background on specific convergence diagnostics:


R implementations of different convergence diagnostics:


**HW 2 assigned: Becoming familiar with WinBUGS/JAGS.**
Wednesday, July 4
No lecture
Instructor & TAs are available for office hours by appointment.

Day 7: Thursday, July 5
The Normal Distribution; Priors (ctd.)
Please read:

- Gill: Chapter 3

Day 8: Friday, July 6
The Bayesian Linear Model
Please read:

- Gill: Chapter 5.

Sample application:


HW 3 assigned: Linear model.
HW 4 assigned: Debugging BUGS/JAGS code.

Day 9: Monday, July 9
Binary Outcomes
If you'd like a refresher for generalized linear models and their interpretation, please read:

- Gelman & Hill, Chapter 5.
Sample application:


**HW 5 assigned: Logistic regression model.**

---

**Day 10: Tuesday, July 10**

**Ordered and Categorical Outcomes**

If you'd like a refresher on today's models, please read:

- Gelman & Hill, section 6.5.

Sample applications (please read one):


**HW 6 assigned: Ordered or multinomial logit model.**

---

**Day 11: Wednesday, July 11**

**Count Outcomes**

If you'd like a refresher on today's models, please read one of the following:

- Gelman & Hill, section 6.2.
- Ntzoufras, sections 7.4 and 8.3

Sample applications (please read one):


**HW 7 assigned: Poisson model.**
Day 12: Thursday, July 12
Measurement Models

For background, please read:

- On factor models:
- On IRT models:

Sample applications (please read one):


**HW 8 assigned: Factor or IRT model.**

---

**Day 13: Friday, July 13**

**Bayes Factors and Bayesian Model Averaging**

Please read:

· Raftery, Adrian E. 1995. “Bayesian Model Selection in Social Research.” *Sociological Methodology* 25:111–163 (Background on BMA, read if you’re interested)
· Gelman, Andrew and Rubin, Donald B. 1995. “Avoiding Model Selection in Bayesian Social Research.” *Sociological Methodology* 25:165–173 (Background on BMA, read if you’re interested)

---

**Day 14: Monday, July 16**

**Model Checking and Model Presentation**

**Multilevel Models (Intro)**

Please read:

· Gill: Chapters 6 & 7.
HW 9 assigned: Model checking for linear regression.

Day 15: Tuesday, July 17
Multilevel Models (Fundamentals)
Please read:
- Gelman & Hill: Chapter 16 or/and Gill: Chapter 10
- Gelman & Hill: Chapter 11 (for a refresher on multilevel models).

HW 10 assigned: Multilevel model.

Day 16: Wednesday, July 18
Multilevel Models (non-continuous outcomes; time-series cross-sectional data as multilevel data)
Multilevel regression with poststratification (MRP)
Please continue to read:
- Gelman & Hill: Chapter 17 (Chapter 15 for a refresher).

as well as any of these empirical articles using MLMs that is/are in your area of interest:
• Mummolo, Jonathan and Peterson, Erik. 2018. “Improving the Interpretation of Fixed Effects Regression Results.” *Political Science Research and Methods*.

Overview and applications of multilevel regression with poststratification (MRP):

Day 17: Thursday, July 19  
Bayesian Analysis of Spatial Data  

If interested, please read the following for background and applications of spatial modeling using Bayesian inference:

- Lunn et al.: Section 11.3.

Day 18: Friday, July 20: Using Bayesian Modeling in Your Applied Work  

Please read:


Also have a look at:


Last updated: March 21, 2018 • http://www.jkarreth.net/bayes-icpsr.html