ICPSR Workshop: Bayesian Modeling for the Social Sciences I
Introduction and Application

Workshop Meeting: M-F 9:00a—11:00a

Dr. Justin Esarey
Associate Professor of Political Science
Wake Forest University

E-mail: justin@justinesarey.com
Phone: 678-383-9629

Syllabus Version: 6/25/2022

Schedule an Office Hours Appointment: https://calendly.com/esareyje/office-hours

Teaching Assistants

Giulia Venturini
Florida State University
Office Hours: https://bit.ly/3y4roo1
gventurini@fsu.edu

Madelyn Quirk
University of Michigan
Office Hours: https://tinyurl.com/MaddieBayesOH
quirkm@umich.edu

Course Objectives and Learning Outcomes

In this course, students will learn how to apply and interpret Bayesian models to the study of social scientific questions. The course will focus on practical Bayesian implementations of the (hierarchical) linear model. Students will learn to use R for programming, data management, and visualization with RStudio as an IDE. Students will also learn to use JAGS and STAN as engines for posterior sampling.

Grading Policies and Assignment Details

Grading will be based on problem sets. Two problem sets will be distributed each week. Include all code you use to complete your assignments with your submission. Each submitted assignment will be graded using the following rubric:

- ✓+ (3 points): Results are correct. They are presented and discussed in a clear manner. Figures and tables are properly labeled. All code is included in the document.
- ✓ (2 points): Results are on the right track, although there may be some errors. They are presented and discussed in a clear manner. Figures and tables are properly labeled. All code is included in the document.
- ✓- (1 point): There are substantial omissions or errors in the results. The presentation is confusing. Code is missing from the document.

All assignments must be typed in LaTeX or RMarkdown.

Final Grades: Submitting all assignments (8-12 points) earns a B. Submitting all assignments and earning a ✓ on most of them (13-16 points) earns an A-. Submitting all assignments with a ✓+ on some assignments (17-21 points) earns an A. A very strong performance on problem sets (22-24 points) earns an A+.

Attendance: Attendance is mandatory in this class. However, attendance will not be formally recorded or factored into the final grade.

COURSE MATERIALS

Required Texts:

- McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in R and STAN, Second Edition.* Boca Raton: CRC Press. Available in Hardcover or as a Kindle e-book.

Other readings are available on the web or the Canvas workshop website.

Software: This course will teach material primarily through R and RStudio. R is free and available from http://cran.r-project.org/. The RStudio IDE is available at https://www.rstudio.com/products/rstudio/download/. If you are an advanced user and wish to take advantage of multithreaded math libraries, you might install the Microsoft R Open distribution available for Windows and Linux at https://mran.microsoft.com/open. We will also study Hamiltonian Monte Carlo and its implementation in STAN; STAN is available at https://mc-stan.org/users/interfaces/rstan.html.

Students may install these programs on their personal computer. They may also choose to use the RStudio Cloud instance that I have reserved for this class. The advantage of RStudio Cloud is that all necessary software is pre-installed on the server and can be accessed with any device that has a browser and an internet connection. A link to join the RStudio Cloud space for this class is available in Canvas.

All assignments must be typed in LaTeX or RMarkdown. If you wish, you may use LyX (http://www.lyx.org/), a WYSIWYG LaTeX editor, in combination with MiKTeX on Windows (http://miktex.org/), MacTeX on Macintosh (http://www.tug.org/mactex/) or TeXLive on Linux (http://www.tug.org/texlive/). RMarkdown PDF output can be produced through RStudio.

All students must have a valid University of Michigan e-mail address and login (and access to the Canvas website) to participate in this course.

COURSE OUTLINE AND ASSIGNED READINGS

June 20: LATEX Introduction and Administrative Activities (no workshop)

June 21: Software and Preliminaries

Readings (25 pages)

- McElreath, Preface and Chapter 1

Software Installation

Choose one of the following options:

- **Local Installation**
  - Install R on your laptop from http://cran.r-project.org/.
  - Install RStudio on your laptop from https://www.rstudio.com/products/rstudio/download/.
  - Install the C++ Toolchain, STAN and RStan on your computer by following the instructions at "RStan Getting Started" and page xvi of McElreath.
  - Install the `rethinking` R library by following the instructions on page xvi of McElreath and (optional) check out its RDocumentation file: https://www.rdocumentation.org/packages/rethinking/versions/2.13.
- **RStudio Cloud Access**
  - Create an account on https://rstudio.cloud and click the link in Canvas to join the workspace for this class.

June 22: A Few Tools from Probability Theory

Readings (18 pages)

- McElreath, Chapter 2: Sections 2.1-2.3

June 23: Bayes' Rule

Readings (10 pages)

- McElreath, Chapter 2: Sections 2.4-2.5

June 24 and 27: Sampling from Distributions and Summarizing the Results

Readings (20 pages)

- McElreath, Chapter 3

June 28: The Gaussian Normal Distribution and Univariate Bayesian Modeling

Readings (21 pages)

- McElreath, Chapter 4: Sections 4.1-4.3

June 29: Quadratic/Laplace Approximation of the Posterior

Readings (25 pages)

- Peng, *Advanced Statistical Computing*, Section 5.1
- Yurko, "Bayesian Baby Steps: Normal Next Steps"

June 30: Bivariate Bayesian Linear Models - Introduction

Readings (10 pages)

- McElreath, Chapter 4: Section 4.4

July 1: Multivariate Bayesian Linear Models - Introduction

Readings (37 pages)

- McElreath, Chapter 5

July 4: Independence Day (no workshop)

July 5: Multivariate Bayesian Linear Models - Model Specification

Readings (29 pages)

- McElreath, Chapter 6

July 6: Fit Diagnostics and Model Comparison

Readings (29 pages)

- McElreath, Chapter 7

July 7: Markov Chain Monte Carlo - Metropolis-Hastings and Gibbs Sampling

Readings (31 pages)

- McElreath, Chapter 9: Sections 9.1-9.2
- Section 4 onward (pp. 492-505) of Jackman, Simon. 2004. "Bayesian Analysis for Political Research." *Annual Review of Political Science* 7: 483-505. DOI: https://doi.org/10.1146/annurev.polisci.7.012003.104706.

July 8: BUGS and JAGS and BUGS and JAGS and BUGS and JAGS

Readings (41 pages)

- "Working with BUGS/JAGS from R." Chapter 6, Section 6.3 (pp. 256-261) in Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. Chichester: John Wiley & Sons. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470686621
- Kevin Ross, Section 10.1 (Introduction to Posterior Simulation and JAGS) of *An Introduction to Bayesian Reasoning and Methods*
- Johannes Karreth, "Using JAGS via R."

Software Installation **(**If you are not using RStudio Cloud):

- Install JAGS on your computer; consult http://www.stat.yale.edu/~jtc5/238/materials/JAGS-quick-start.htm.
- Install the `R2jags`, `coda`, `lattice`, `rjags`, `runjags`, and `MCMCpack` R libraries on your computer as described in Steps 5-7 of Karreth pp. 2-3.

July 11: Hamiltonian Monte Carlo

Readings (37 pages)

- McElreath, Chapter 9: Section 9.3
- Sections 1-3 (pp. 1-26) of Michael Betancourt. 2017. "A Conceptual Introduction to Hamiltonian Monte Carlo." arXiv:1701.02434.

July 12: STAN, STAN, STAN, egg, and STAN

Readings (17 pages)

- McElreath, Chapter 9: Section 9.4
- Optional: STAN User's Guide, v. 2.29.

July 13: Markov Chain Assessment and Diagnostics

Readings (48 pages)

- McElreath, Chapter 9: Section 9.5
- Chapter 6, Sections 6.1-6.2 (pp. 251-256) and Section 6.4 (pp. 261-292) in Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. Chichester: John Wiley & Sons. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470686621.

July 14: Multilevel Models – Varying Intercepts

Readings (33 pages)

- McElreath, Chapter 13

July 15: Multilevel Models – Varying Slopes

Readings (21 pages)

- McElreath, Chapter 14, Sections 14.1-14.2