Applied Bayesian Modeling for the Social Sciences

POLS 607
Spring 2017

INSTRUCTOR: Florian M. Hollenbach
Email: fhollenbach@tamu.edu; **Web:** fhollenbach.org
Office: 2061 Allen Building; **Phone:** 979-845-5021
Office Hours: Wednesdays 3:00pm – 4:00 pm or by appointment

Class Meeting Time: Thursday: 1:00 pm – 3:50 pm
Class Location: Bush Academic Building West 2064

COURSE DESCRIPTION: This class will cover applied Bayesian statistics for social scientists. In the first half of the semester we will cover fundamental topics of Bayesian inference, both mathematically, as well as computational strategies of inference. The second half of the semester we will cover specific problems that students may encounter in their research and discuss Bayesian strategies to deal with these problems.

LEARNING OUTCOMES: At the end of the semester, after completing this course students are expected to:

- Have gained an understanding of basic concepts of Bayesian inference
- Be able to explain, understand, program, and implement several Bayesian statistical models
- Be able to understand and program MCMC techniques, as well as analyze, interpret, and diagnose problems
- Be able to do simple programming of Bayesian models in R and Stan

COURSE STRUCTURE & REQUIREMENTS:

The class will meet once a week on Thursday afternoons. Generally, we will spend the first 45–60 min of class reviewing and discussing the homework that is due that week. This involves students presenting solutions to the homework problems and discussing them. I expect each of you to be able to present the solution to at least some of the problems and to have an attempt of a solution for all problems. The second part of the class will be an interactive lecture. I will lead the class and present the topic that we are covering that week, but I expect that you participate and especially that you ask questions when you have trouble understanding something. This class will cover a variety of (complicated) concepts, which will first be covered in the readings. I expect you to have *all* of the required reading done before each class period.

For some weeks, the readings & topics covered will be quite technical and challenging, thus it is even more important that you try to understand the material before coming to class. If you do not understand part of the readings or lecture, please raise questions in class. I guarantee you will not be the only one that has trouble with the material and by asking questions, you are providing a service to your classmates.

There will be regular homework assignments. These are for you to deepen your understanding of the material. Some of the assignments will be quite hard. It is, however, important for your own progress that you at least attempt to solve it on your own first, before seeking help. If you really are stuck, I encourage you to seek help from your classmates or myself.

Grading & Your Responsibilities: Your grade will be based on the take home exam (27 %) at the end of the semester, class participation (10%), 13 homework assignments (48% combined – your lowest score will be dropped), and a class presentation (15%).

I expect you to do all of the assigned readings for each class before the lecture, participate in class discussions, and come prepared with questions. You will be graded on:

- 13 home work assignments - 4% each – your lowest score will be dropped – 48% total
- Class participation (10 %)
- Presentation (15 %)
- Take Home Exam (27 %)

Your class participation grade will be based on participating in the class discussion, homework solutions, and asking questions. There are no stupid questions in this class. Rather any question about the material will reflect positively on you.

The grading scale (in %) used in this class for all assignments and exams will be the following:

- A = ≥ 89.5
- B = ≥ 79.5 – < 89.5
- C = ≥ 69.5 – < 79.5
- D = ≥ 59.5 – < 69.5
- F = < 59.6

Homeworks: There will be (13) weekly homework assignments handed out in class. The assignment will be due in class the following week. We will spend the first 45–60 min of each class session discussing the homework problems due that day. This part of the class will be based on your participation, i.e. you will be presenting solutions to the individual problems. I expect each of you to be prepared to present your solutions. It is okay if you struggled with some of the problems, as long as you can show a serious attempt at solving the problem.
The homework assignments will be one of the most important mechanisms for your learning progress in this class and as such also a major basis of your grade. I encourage you to spend a lot of time on the homework and try to understand each of the problems.

All homework assignments must be typed in LaTeXand submitted as hardcopies in class.

Presentations: Each of you will choose one of the topics in weeks 9–15. You are then to find one published scientific article (preferably in the social sciences) that applies the general method discussed that week. Prepare a short ($\sim$ 30 min) presentation about the article. What is the application and how do the authors motivate the use of Bayesian statistics? How is it executed? What are the model specifications? What would you do differently? In addition, replicate the article and add one reasonable modification/extension to the model. Show how the results change or do not change. Please meet with me in the week leading up to your presentation, so we can make sure you are on the right track.

ACADEMIC HONESTY: All students should follow the highest standards of academic integrity. Cheating or plagiarism will not be tolerated in any way. If you are unsure what entails plagiarism, come talk to me. For more info, see: `http://student-rules.tamu.edu/aggiecode` & `http://aggiehonor.tamu.edu`. **"An Aggie does not lie, cheat or steal, or tolerate those who do."**

Regarding group work: You are allowed to work together on homework problems as a last resort. Yet, I want to encourage each of you to attempt the homework by yourself and only work together once you get stuck on a problem. Also, note that passively attending group work sessions will not help you understand the material. Lastly, all final homework assignments must be based your **own** work. Copying another students homework will be treated as cheating.

READINGS & SOFTWARE: We will primarily use three books. The main book is available in general book stores or on Amazon. It is entitled: "Statistical Rethinking" by Richard McElreath. This book will be referenced as **McElreath** below. The other two main books we use are available in electronic form through the TAMU library or you can buy them in print. They are: " A first course in Bayesian Statistical Methods" by Peter D. Hoff and "Bayesian Analysis for the Social Sciences" by Simon Jackman. These books will be denoted **Hoff** and **Jackman** below.
All other readings will either be available electronically through the library or made available as scans on the class GoogleDrive.

Required Books

- McElreath, Richard. 2016. "**Statistical Rethinking: A Bayesian Course with Examples in R and Stan**". Chapman & Hall/CRC Texts in Statistical Science. (denoted McElreath below)
- Hoff, Peter D. 2009. "**A First Course in Bayesian Statistical Methods**". Springer Texts in Statistics. (denoted Hoff below)
- Jackman, Simon. 2009. "**Bayesian Analysis for the Social Sciences**". John Wiley & Sons, Ltd

Recommended Books

- Gelman, Andrew, Carlin, John B., Stern, Hal S. Dunson, David B., Vehtari, Aki, and Rubin, Donald B. 2013. "**Bayesian Data Analysis, Third Edition**". Chapman & Hall/CRC Texts in Statistical Science. (denoted BDA below) – required Chapters will be provided

Software: In part of this class we will be working with statistical software. We will use the statistical programming language $\mathcal{R}$. Please make sure your computer has R installed. You are also welcome to use R-Studio. Later in the class we will work with both STAN and JAGS and their respective integration in R.

CLASSROOM BEHAVIOR, ATTENDANCE, PARTICIPATION, & ELECTRONIC DEVICES: We will usually meet once a week during the semester. This is a graduate level class, so I expect you to do all of the required readings, homework, and come to class prepared. Some of the homework assignments will be quite tough and I expect that you will struggle with some of them. That is okay, however, even if you are unable to find the right solution I expect to you to try as hard as you can. We can then discuss solution strategies in class. The class will be interactive and I expect each of you to participate in class discussions and raise questions. It is pertinent to the success of this class that you ask questions when you do not fully understand something.

Attendance is mandatory. You are allowed to miss one class unexcused. Each additional unexcused absence will lead to a 2.5 percentage point deduction on your final grade. Given that this is a graduate class, I do not expect this to happen. Excused absences are those that are university excused, as well as attendance of professional conferences.

We will use laptops for parts of this class, so please bring your laptop. I ask you, however, to keep your laptops closed during times that we are not actively working with the computer. It has been shown repeatedly that students learn better when taking notes by hand and it will lead to a more open class environment. Please turn off or mute your cell phones and arrive to class on time.

EXAM ABSENCES & LATE POLICY: Make-up exams will be permitted only in the case of university-excused absences. To be eligible for a make-up exam, you will have to present original written documentation of legitimate circumstances that prevented you from taking the exam on time. Except in the case of observance of a religious holiday, to be excused, the student must notify his or her instructor in writing (acknowledged e-mail message is acceptable) **prior** to the date of absence. In cases where advance notification is not feasible (e.g. accident or emergency) the student must provide notification by the end of the second working day after the absence. This notification should include an explanation of why notice could not be sent prior to the class. Accommodations sought for absences due to the observance of a religious holiday can be sought either prior or after the absence, but not later than two working days after the absence. Legitimate circumstances include religious holidays, illness (verified by a doctor), serious family emergencies and participation in group activities sponsored by the University, etc. See `http://student-rules.tamu.edu/rule07` for additional information. Please note that I do not accept Xeroxed copies of medical excuses from students. Unexcused absences from either exam will result in a score of 0 for the exam.

Unexcused late work will be penalized by a 7.5 **percentage point deduction for each day your work is late.** For example, if you hand in the a homework assignment on the same day it is due, but after class, your maximum score will be 92.5%. If you hand in your assignment the next day, your maximum score will be 85%, and so on. Late work will be excused only in the case of university-excused absences.

There will be no exceptions to these rules.

RE-GRADING POLICY: Students that want to appeal a grade received on an exam or assignment must submit a regrading request in written form (no email). This request has to be turned in within five working days after the graded exams or assignments are returned to the class. The written statement must explain exactly why the student believes the current grade is incorrect. I will then regrade the entire assignment or exam extra carefully. NOTE, as a consequence your grade may go up or down.

COMMUNICATION:

The best place to ask questions is in the class room. If your question is not related to class material or relevant to other students, we can discuss it after class. If you have any difficulties with the readings or homework, I encourage you to first ask questions in class. You are also welcome to come to office hours. Again, however, you should at least attempt to solve the problem on your own first.

You can expect me to reply to emails within 24 hours during the work week. I will not reply to emails on the weekend, except for urgent matters.

DISABILITY: All discussions will remain confidential. University policy is in accordance with the Americans with Disabilities Act Policy Statement.

The Americans with Disabilities Act (ADA) is a federal anti-discrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a

learning environment that provides for reasonable accommodation of their disabilities. If you believe you have a disability requiring an accommodation, please contact Disability Services, currently located in the Disability Services building at the Student Services at White Creek complex on west campus or call 979-845-1637. For additional information, visit `http://disability.tamu.edu`.

Reasonable accommodations will be made for all students with disabilities, but it is the student's responsibility to inform the instructor early in the term. Do not wait until just before an exam to decide you want to inform the instructor of a learning disability; any accommodations for disabilities must be arranged well in advance.

DIVERSITY POLICY: The Department of Political Science supports the Texas A&M University commitment to diversity, and welcomes individuals from any racial, ethnic, religious, age, gender, sexual orientation, class, disability, and nationality. (See `http://diversity.tamu.edu/`). In the spirit of this vital commitment, in this course each voice in the classroom has something of value to contribute to all discussions. Everyone is expected to respect the different experiences, beliefs and values expressed by fellow students and the instructor, and will engage in reasoned discussion that refrains from derogatory comments about other people, cultures, groups, or viewpoints.

CLASS SCHEDULE:
(Subject to Change)

Week 1 (01/19
Introduction, Syllabus, Logistics
Reading:

- Hoff: Chapter 1 & 2,
- Jackman Chapter 1
- McElreath Chapter 1 & 2

Week 2 (1/26)
One parameter models & Sampling
Reading:

- McElreath Chapter 3
- Hoff: Chapter 3 & 4

Week 3 (02/02)
The normal model & linear model
Reading:

- Hoff Chapter 5
- McElreath Chapter 4

Week 4 (02/09)
Multivariate Linear Model
Reading:

- McElreath Chapter 5 & 6
- Jackman Chapter 2.5

Week 5 (02/16)
Markov Chain Monte Carlo Methods & Gibbs Sampling
Reading:

- Hoff Chapter 6
- Jackman Chapters 3

Week 6 (02/23)
Markov Chain Monte Carlo Methods II, Metropolis Hastings, & HMC
Reading:

- McElreath Chapter 8
- Jackman Chapter 4 & 5

Week 7 (03/02)
Methods for MCMC estimation: R, STAN, JAGS Reading:

- Betancourt, Michael. 2017. "A Conceptual Introduction to Hamiltonian Monte Carlo". ArXiv.org. Available here: `https://arxiv.org/abs/1701.02434`
- Savage, Jim. "A quick-start introduction to Stan for economists". Available here: `https://nbviewer.jupyter.org/github/QuantEcon/QuantEcon.notebooks/blob/master/IntroToStan_basics_workflow.ipynb`
- STAN Language modeling. Skim Sections I and II. Available here: `https://github.com/stan-dev/stan/releases/download/v2.14.0/stan-reference-2.14.0.pdf`

Week 8 (03/09)
Generalized Linear Models in a Bayesian framework
Reading:

- McElreath Chapter 9 & 10
- Jackman Chapter 8

Week 9 (03/23)
Mixture Models
Reading:

- McElreath Chapter 11
- BDA Chapter 22

Week 10 (03/30)
Hierarchical Models
Reading:

- McElreath Chapter 12
- BDA Chapter 15

Week 11 (04/06)
Bayesian Approaches to Measurement
Reading:

- Jackman Chapter 9
- Clinton, Joshua, Jackman, Simon, & Rivers, Douglas. 2004. "The Statistical Analysis of Roll Call Data". *American Political Science Review.* 98:2. 355–370.

Week 12 (04/13)
Variable Selection & Bayesian Model Averaging
Reading:

- Montgomery, Jacob M. & Nyhan, Brendan. 2010. Bayesian Model Averaging: Theoretical Developments and Practical Applications. *Political Analysis.* 18(2):245-270
- O'Hara, R.B. & Sillanpää, M. J. 2009. A Review of Bayesian Variable Selection Methods: What, How and Which. *Bayesian Analysis.* 4(1): 85–118

Week 13 (04/20)
Missing Data
Reading:

- BDA Chapter 18
- McElreath Chapter 14

Week 14 (04/27)
Causal Inference
Reading:

- Chapter 1 in: Imbens, Guido W. & Rubin, Donald B. 2015. "Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction". Cambridge University Press. New York, NY. (Provided on GoogleDrive)
- Kraay, Aart. 2012. "Instrumental variables regressions with uncertain exclusion restrictions: a Bayesian approach". *Journal of Applied Econometrics*. 27:1.108–128.

Week 15 (05/04)
Take Home Exam Due