MACS 30301 - Introduction to Bayesian Statistics

Computational Social Science
Division of the Social Sciences
University of Chicago
Autumn/2019

Course information

Instructor: Diogo Ferrari
e-mail: diogoferrari@uchicago.edu
Office: 5730 S. Woodlawn, Room 406

Office-hours: Monday 15:00-17:00 (or by appointment)

- **Meeting day/time** MWF 10:30-11:20
- **Location**: Stuart Hall 105

Description

The goal of this course is to give students an overview of the theory and methods for data analyses using the Bayesian paradigm. Topics include: (1) foundations of Bayesian inference; (2) development of Bayesian models and prior choices; (3) analytical and simulation techniques for posterior estimation; (4) model choice and diagnostics; (5) sensitivity analysis, and; (6) introduction to Monte Carlo Markov Chain (MCMC) simulations. Students will also learn how to estimate and summarize Bayesian models using Bayesian statistical packages (R/JAGS/Bugs). The course will use working examples with real application of Bayesian analysis in social sciences. Prerequisites: Basic knowledge of probability (e.g., joint and conditional distributions, expectation, variance) and introductory-level experience with R or Python (Note: Open to Advanced Undergraduates with Instructor Permission)

Evaluation

	Total	Total weight
Problem sets	4	80%
Final exam	1	20%

Schedule

Calendar

#	Week	Day	Topic	Assignment
1	1	W10/02	Introduction	
2	1	F10/04	Motivation for Bayesian Statistics	
			Subjective Bayes	
3	1	M10/07	The prior distribution: Preliminaries	
4	2	W10/09	Prior Distribution (conjugate priors)	
5	2	F10/11	Prior Distribution (conjugate priors)	
			Objective Bayes and **Sensitivity**	
6	2	M10/14	Prior Distribution	
7	3	W10/16	Prior Distribution	
8	3	F10/18	Sensitivity and Prior perturbation analysis	PS1 (ho)
			Sensitivity	
9	3	M10/21	Sensitivity and Prior perturbation analysis	
10	4	W10/23	Hierarchical Bayesian Models	
11	4	F10/25	MC methods	
			Monte Carlo Methods	
12	4	M10/28	MC indirect sampling (RS)	
13	5	W10/30	MC indirect sampling (ARS, IS)	
14	5	F11/01	MC indirect sampling (IS, AIS, wrap-up)	PS2 (ho); **PS1 (d)**
			Monte Carlo Markov Chain Methods	
15	5	M11/04	MCMC (Intuition, GBC, DBC)	
16	6	W11/06	MCMC (DBC and Metropolis-Hasting update)	
17	6	F11/08	MCMC (Foundations of MCMC method)	
			Various MCMC and Diagnostics	
18	6	M11/11	MCMC (MH Variants: Random Walk)	
19	7	W11/13	MCMC (MH Variants: Gibbs, Slice Sampler)	
20	7	F11/15	MCMC (Into Implementation)	PS3 (ho); **PS2 (d)**
			Implementation	
21	7	M11/18	MCMC (Implementation)	
22	8	W11/20	MCMC (Implementation)	
23	8	F11/22	MCMC (Implementation)	
			Predictive and Summaries	
24	8	M11/25	MCMC (diagnostics)	
25	9	W11/27	Summaries of the posterior	PS4 (ho); **PS3 (d)**
26	9	**F11/29**	No class: **Thanksgiving Break**	
27	9	M12/02	Model Choice and Predictive Analytics	
28	10	W12/04	Wrap-up	**PS4 (d)**
29	10	**F12/06**	No class: **College Reading Period**	
30	10	M12/09	**Final Exam**	**Final Exam**

- ho: hand out
- d: due date

Textbooks

Required

- Gill, J. (2014) Bayesian methods: a social and behavioral sciences approach: CRC press.
  Practical introduction of the concepts of Bayesian statistics with examples of application in social sciences.

Recommended

- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2014) Bayesian Data Analysis, Chapman & Hall/CRC Boca Raton, FL, USA.
  Called for many people uThe Bibleu, this book contains a comprehensive overview of many topics in Bayesian statistics.

- Jackman, S. (2009) Bayesian analysis for the social sciences , John Wiley & Sons.
  The book contains more advanced treatment of some topics in Bayesian analysis.

- Kruschke, J. (2015) Doing bayesian data analysis: a tutorial with r, jags, and stan, Academic Press.
  This is an introduction-level book with many examples and illustrations of Bayesian analysis. It covers the computational side, and it is a useful tutorial-like approach for practical implementation of Bayesian models in JAGS.

Software

- Team, R. C. (2018) R: a language and environment for statistical computing.
  Free statistical software available for download at this website. The software is available for all main operation system (Linux, OS, Windows)

- Plummer, M., Stukalov, A., & Denwood, M. (2018) Rjags: r api to jags.
  Free software for Bayesian analysis available for download at this website. The software is available for UNIX based operation system (Linux, OS).

- Lunn, D., Jackson, C., Best, N., Thomas, A., & Spiegelhalter, D. (2012) The bugs book: a practical introduction to bayesian analysis, CRC press.
  Windows users can install the free windows based version called WinBUGS instead of JAGS, which is available for download at this website

- Plummer, M. (2017) Jags version 4.3.0 user manual.
  R package used to estimate Bayesian models using JAGS.

Diversity, Inclusion, and Disability Statement

This course is open to all students who meet the academic requirements for participation. Any student who has a documented need for accommodation should contact Student Disability Services (773-702-6000 or disabilities@uchicago.edu) and the instructor as soon as possible.

It is my intent that students from all diverse backgrounds and perspectives be well-served by this course, that students' learning needs be addressed both in and out of class, and that the diversity students bring to this class be viewed as a resource, strength, and benefit. Please let me know ways to improve the effectiveness of the course for you personally, or for other students or student groups. Your suggestions are encouraged and appreciated.

It is my intent to present materials and activities that are respectful of diversity: gender identity, sexuality, disability, age, socioeconomic status, ethnicity, race, nationality, religion, and culture. I will attempt to foster an environment in which each class member is able to hear and respect one another. It is my intent to maintain an atmosphere of trust and safety in the classroom. Please let me know if something said or done in the classroom, by either myself or other students, is particularly troubling or causes discomfort or offense. While our intention may not be to cause discomfort or offense, the impact of what happens throughout the course is not to be ignored and is something that I consider to be very important and deserving of attention. If and when this occurs, there are several ways to alleviate some of the discomfort or hurt you may experience:

1. Discuss the situation privately with me. I am always open to listening to students' experiences and want to work with students to find acceptable ways to process and address the issue.

2. Discuss the situation with the class. Chances are there is at least one other student in the class who had a similar response to the material. Discussion enhances the ability for all class participants to have a fuller understanding of context and impact of course material and class discussions.

3. Notify me of the issue through another source such as your preceptor, a trusted faculty member, or a peer. If for any reason you do not feel comfortable discussing the issue directly with me, I encourage you to contact your preceptor and/or your program's Diversity and Inclusion representative: Darcy Heuring (MAPSS), Matthias Staisch (CIR), and Chad Cyrenne (Computation). You are also welcome and encouraged to contact the Faculty Director of your program.

The University of Chicago is committed to diversity and rigorous inquiry from multiple perspectives. The MAPSS, CIR, and Computation programs share this commitment and seek to foster productive learning environments based upon inclusion, open communication, and mutual respect for a diverse range of identities, experiences, and positions. Any suggestions for how we might further such objectives both in and outside the classroom are appreciated and will be given serious consideration. Please share your suggestions or concerns with your instructor, your preceptor, or your program's Diversity and Inclusion representatives: Darcy Heuring (MAPSS), Matthias Staisch (CIR), and Chad Cyrenne (Computation). You are also welcome and encouraged to contact the Faculty Director of your program.

Academic Integrity

The University of Chicago has a formal policy on academic honesty that you are expected to adhere to. Here are some guidelines we expect you to follow:

1. Courtesy, honesty, and respect should be shown by students toward faculty members, guest lecturers, administrative support staff, and fellow students. Similarly, students should expect

faculty to treat them fairly, showing respect for their ideas and opinions and striving to help them achieve maximum benefits from their experience in the School.

2. Academic dishonesty can encompass many activities, which includes plagiarism, cheating, fabrication, falsification of records or official documents, intentional misuse of equipment or materials (including library materials), and aiding and abetting the perpetration of such acts. One of the gravest academic dishonesty is plagiarism: knowingly handing in someone else's work as your own, whether it be work done by another student in the class or available publicly on the Internet. This class has a zero tolerance policy for plagiarism.

3. The preparation of solutions for problem sets, papers, and examinations, assigned on an individual basis, must represent each students own effort. Therefore:

   - You MUST NOT copy or use someone else's work (with or without their permission) in your own solution. You have to write your own.
   - DO NOT post your solutions to problem sets or exams in publicly-accessible websites, like pastebin, a public GitHub repository, GitHub gists, etc. While these tools may seem like convenient mechanisms for sharing code with an instructor/TA or with a project partner, they can also expose your code to other students in the class. If you do post your solution in a publicly-accessible location, and we find out about it outside of a plagiarism incident, you will just get a warning. However, if another student in the class uses code that you posted on such a site (even if you did not intend for that code to be used by someone else), you be considered an equally guilty party in a plagiarism offense, and will receive the exact same penalty as the student who used your code.

Scheduled readings

Lecture 1: Introduction

Suggested

- Fienberg, S. E. et al. (2006). When did bayesian inference become "bayesian"? *Bayesian analysis*, 1(1):1–40

Lecture 2: Motivation for Bayesian Statistics

Required

- Gill, J. (2014). *Bayesian methods: A social and behavioral sciences approach*, volume 20. CRC press
  - Ch 2.1: Purpose
  - Ch 2.3: The Basic Bayesian Framework (only subsection 2.3.1)
  - Ch 2.4: Bayesian "Learning"
  - Ch 2.6: Bayesian versus Non-Bayesian Approaches

Lecture 3: Prior Distribution: Preliminaries

Required

- Gill, J. (2014). *Bayesian methods: A social and behavioral sciences approach*, volume 20. CRC press
  - Ch 2.5: Comments on Prior Distributions
  - Ch 4.1: A Prior Discussion of Priors

Suggested

- Wasserman, L. et al. (2006). Frequentist bayes is objective (comment on articles by berger and by goldstein). *Bayesian Analysis*, 1(3):451–456

Lecture 4 and 5: Prior Distribution (conjugate priors)

Required

- Gill, J. (2014). *Bayesian methods: A social and behavioral sciences approach*, volume 20. CRC press
  - ch 4.3: Conjugate Priors (subsections 4.3.1-4.3.3)

Lecture 6 and 7:

Required

- Gill, J. (2014). *Bayesian methods: A social and behavioral sciences approach*, volume 20. CRC press
  - Ch 4.3.3: Limitations of Conjugacy
  - Ch 4.4.2: Jeffreys Prior (all subsections)

Lecture 8, 9:

Required

- Gill, J. (2014). *Bayesian methods: A social and behavioral sciences approach*, volume 20. CRC press
  - Ch 6.2.1 Global Sensitivity Analysis
  - Ch 6.2.2 Local Sensitivity Analysis
  - Ch 6.2.3 Global and Local Sensitivity Analysis with Recidivism Data

Lecture 10:

Required

- Gill, J. (2014). *Bayesian methods: A social and behavioral sciences approach*, volume 20. CRC press
  - Ch 12.5 Essential Structure of the Bayesian Hierarchical Model
  - Ch 12.6 The General Role of Priors and Hyperpriors

Lecture 11: MC methods

Required

- Gill, J. (2014). *Bayesian methods: A social and behavioral sciences approach*, volume 20. CRC press
  - Ch 9.1 Background
  - Ch 9.2 Basic Monte Carlo Integration

Suggested

- Robert, C. and Casella, G. (2011). A short history of markov chain monte carlo: Subjective recollections from incomplete data. *Statistical Science*, 26(1):102–115

Lecture 12: MC indirect sampling

Required

- Gill, J. (2014). *Bayesian methods: A social and behavioral sciences approach*, volume 20. CRC press
  - Ch 9.3 Rejection Sampling

Lecture 13-14

Required

- TBA

Lecture 15

Required

- Robert, C. and Casella, G. (2011). A short history of markov chain monte carlo: Subjective recollections from incomplete data. *Statistical Science*, 26(1):102–115

Lecture 16

Required

- Chib, S. and Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The American Statistician*, 49(4):327

Suggested

- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092
- Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications
- Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6):721–741
- Gelfand, A. E. and Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409

Lecture 17

Suggested

- Mengersen, K. L., Tweedie, R. L., et al. (1996). Rates of convergence of the hastings and metropolis algorithms. *The annals of Statistics*, 24(1):101–121

Lecture 18-29

Required

- TBA

Lecture 25: Summaries of the Posterior

Required

- Gill, J. (2014). *Bayesian methods: A social and behavioral sciences approach*, volume 20. CRC press
  - Ch 2.3.2 Summarizing Posterior Distributions with Intervals
  - Ch 2.3.3 Quantile Posterior Summaries

Lecture 27:

Required

- TBA

1 :ignore:

References

Chib, S. and Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The American Statistician*, 49(4):327.

Fienberg, S. E. et al. (2006). When did bayesian inference become "bayesian"? *Bayesian analysis*, 1(1):1–40.

Gelfand, A. E. and Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409.

Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6):721–741.

Gill, J. (2014). *Bayesian methods: A social and behavioral sciences approach*, volume 20. CRC press.

Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications.

Mengersen, K. L., Tweedie, R. L., et al. (1996). Rates of convergence of the hastings and metropolis algorithms. *The annals of Statistics*, 24(1):101–121.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092.

Robert, C. and Casella, G. (2011). A short history of markov chain monte carlo: Subjective recollections from incomplete data. *Statistical Science*, 26(1):102–115.

Wasserman, L. et al. (2006). Frequentist bayes is objective (comment on articles by berger and by goldstein). *Bayesian Analysis*, 1(3):451–456.