GOV 2003: Topics in Quantitative Methodology

Kosuke Imai Santiago Olivella

Spring 2020

Abstract

This course introduces social science students to applied Bayesian statistics. We will begin by introducing Bayes' rule, which allows us to learn from data in an intuitive and coherent way. We then cover a set of simple probabilistic models and cover powerful computational tools that will be used for the remainder of the course. Finally, we will learn about social science applications of Bayesian models including regression models, topic models, social network models, and dynamic models. The course will build everything up from the basic principles, only requiring the knowledge of elementary probability and statistics along with the familiarity with R programming.

1 Contact Information

Instructors

	Kosuke Imai	Santiago Olivella
Office:	CGIS K306	CGIS S313
Email:	imai@harvard.edu	olivella@fas.harvard.edu
URL:	https://imai.fas.harvard.edu/	http://www.santiagoolivella.info
Office Hours:	Tuesdays 1:20pm – 2:50pm (sign up at http://bit.ly/ImaiOfficeHours), email me for an appointment, or stop by any time!	Wednesdays 3:30pm – 5:30pm (sign up at http://bit.ly/OlivellaOfficeHours) or email me for an appointment at a different time.

Teaching Fellow

	Soichiro Yamauchi
Email:	syamauchi@g.harvard.edu
URL:	https://soichiroy.github.io/
Office Hours:	TBD

2 Logistics

- Lectures: Monday and Wednesday, 10:30am–11:45am, Location Sever Hall 308.
- Section: Thursdays, 3:00pm – 4:15pm, Location CGIS K354.

Use of laptops and cell phones during the lectures is not allowed unless you are given a permission from the instructor. Research has shown that the use of these electronic devices often interferes with your learning and has a negative impact on those around you. We will bring copies of lecture slides to class so that you can take notes directly on the slides. Lecture slides will be posted at Piazza by the end of the lecture day.

3 Prerequisites

Students should have taken Gov 2000 or the equivalent course in basic probability and applied statistics (e.g., Stat 110 and Stat 111).

4 Questions, Announcements, and Submissions

- In addition to TF sections and office hours, please use the Piazza Discussion Board at https://piazza.com/ when asking questions about lectures, problem sets, and other course materials. This allows all students to benefit from the discussion and to help each other understand the materials. Both students and instructors are encouraged to participate in discussions and answer any questions that are posted. To join the Gov 2003 Piazza site, click https://piazza.com/harvard/spring2020/gov2003. You will then be prompted to enter your harvard.edu email address to confirm your registration. All class announcements as well as electronic submissions of assignments will be made through Piazza.
- We will use Canvas for the submission of all assignments including the final exam and projects. The Canvas site for this course is at https://canvas.harvard.edu/courses/65389
- A Google calendar that contains the information about the course logistics is available at https://bit.ly/2RBC1Ir

5 Class Requirements

For each requirement, no late submission is allowed without a prior approval of the instructors.

- **Problem sets (20% of final grade):** There will be four problem sets throughout the semester, one for each section of the course. The exact dates they will be available and due are listed below in the schedule. Each problem set will be equally weighted.
  - Collaboration: You may collaborate with other classmates (with the exception of bonus questions) and receive help from course staff on problem sets. However, you must not copy anybody else's code or answers and are required to submit your own answer for each problem set.
- **Weekly labs (10% of final grade)** Every week you will work on a short applied task during TA section. You will be required to submit a completed task no later than noon Sunday of the corresponding week. The same collaboration rules that apply to problem sets apply to these weekly labs.
- **Two quizzes (30% of final grade):** There will be two in-class, closed-book quizzes. The exact dates when they will take place are available on the schedule below. Each quiz is equally weighted.
- **Final project (40% of final grade):** The final project will be completed in collaboration with another student in the class. All projects must use a Bayesian method. Ideal projects will either (a) apply a technique discussed in class (or a related technique) to answer a substantive question or (b) extend such techniques in useful ways. To help keep you on track, there will be multiple deliverables throughout the semester.
  - **February 17 (Project and collaborator identification)** By this date, pairs should submit a one-page project proposal with a brief statement of the problem to be solved or the question to be answered and your strategy for tackling it.
  - **February 26 (Project approval)** Pairs must schedule a time to meet with instructors between 2/17 and 2/26 to discuss your proposal, **and receive their approval on your project no later than 2/26**. Pairs also have to resubmit the revised proposal by this date.
  - **March 23 (First deliverable)** By this date, pairs must submit a first deliverable (maximum of 5 pages) including a well crafted abstract, an elaboration on the problem statement, descriptive analysis of the data to be analyzed, and a preliminary model formulation.
  - **April 29 (Preliminary result)** By this date (the final day of classes), pairs must submit a PDF slide-deck (maximum of 10 slides) with the preliminary results of your analysis.
  - **May 16 (Final report)** By this date (last day of the exam period), pairs must submit the final report (no longer than 20 doublespaced pages) focusing on methods and results.

  Although these are fixed milestones, we are happy to supervise your project throughout the semester and beyond!

6 Textbook

The textbook for the course is Gelman et al. (2013) *Bayesian Data Analysis* (3rd. Edition). Other readings are available online on the course Canvas website.

7 Class Schedule

Section	Dates	Week	Topic	Readings	Assignments
Introduction to Bayesian inference	1/27, 1/29	0	Bayesian primer and probability refresher	BDA: 1	
	2/3, 2/5	1	Inference in single parameter models	BDA: 2	PS1 out on 2/5
	2/10, 2/12	2	Inference in multi-parameter models	BDA: 3	PS1 due on 2/15
	2/19	3	Hierarchical models	BDA: 3–4	Proposal due on 2/17
	2/24	4	Quiz 1		
Fundamentals of Bayesian computation	2/26	4	Classical simulation	BDA 10; Bishop 9	PS2 out on 2/26, project approval by 2/26
	3/2, 3/4	5	MCMC, Gibbs	BDA 11	PS2 due on 3/7
	3/9, 3/11	6	MH and slice sampler		
Regression models	3/23, 3/25	7	HMC and diagnostics	BDA 12, 13	First deliverable due on 3/23, PS3 out on 3/25
	3/30, 4/1	8	Bayesian regression models	BDA 14, 16	PS3 due on 4/4
	4/6, 4/8	9	Topics in regression	BDA 15, 17, 18	
Latent variable modeling	4/13	10	Quiz 2		
	4/15	10	Ideal Point Estimation	Jackman 9; Martin and Quinn 2002; Imai et al. 2016	PS4 out on 4/15
	4/20, 4/22	11	Topic Models	Blei et al. 2003; Roberts et al. 2014 (w/ online appendix); Gerrish and Blei 2012	PS 4 due on 4/25
	4/27, 4/29	12	Network Models	Hoff et al. 2002; Airoldi et al. 2006; McCormick and Zheng 2012	Preliminary results due on 4/29