STAT 789-01: Foundations of Bayesian Data Analysis

Syllabus and Course Information – Spring 2020

2020 January 6

1 Overview

1.1 Course Description from RIT

This course is an introductory graduate-level course on the principles and techniques of indference data analysis using Bayesian methods. Topics covered include parameter estimation, hypothesis testing and model selection, marginalization over nuisance parameters, and numerical methods. Bayesian methods are compared and contrasted with classical frequentist techniques where appropriate. Demonstrations and exercises will make use of a high-level scripting language such as R or Python.

1.2 Computing Environment

We will make extensive use in this course of Jupyter notebooks running R for lessons, homework and exams. You should install the Python 3.7 version of the Anaconda distribution from `https://www.anaconda.com/distribution/` and enable an environment running Python 3.7 and R.

1.3 Textbooks

- **Required:** Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B., *Bayesian Data Analysis*, 3rd edition (CRC, 2013)
- **Required:** McElreath, R., *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (CRC, 2017)
- **Recommended:** Kruschke, J., *Doing Bayesian Data Analysis: A Tutorial with R, JAGS and Stan*, 2nd edition (Academic Press, 2014)
- **Possibly useful:** Sivia, D. S. with Skilling, J., *Data Analysis: A Bayesian Tutorial*, 2nd edition (Oxford, 2006)
- **Possibly useful:** Jaynes, E. T., *Probability Theory: The Logic of Science* (Cambridge, 2003)
- **Possibly useful:** Bolstad, W. M. and Curran, J. M., *Introduction to Bayesian Statistics*, 3rd edition (Wiley, 2017)
- **Possibly useful:** Bolstad, W. M., *Understanding Computational Bayesian Statistics* (Wiley, 2009)

1.4 Instructor

Dr. John T. Whelan; `jtwsma@rit.edu`
Office Hours: by appointment (slack or BlueJeans).

1.5 Prerequisites

This course is restricted to students in the MS or Advanced Certificate program in Applied Statistics. Students should also have taken an introductory statistics course.

2 Course Structure

This is a 3-credit online course; you should expect to spend at least nine hours per week on activities for this class.

2.1 List of Topics

1 Bayesian interpretation of probability
   - 1.1 Probability as extended logic
   - 1.2 Sum and product rules
   - 1.3 Bayes's theorem and conditional probabilities

2 Bayesian parameter estimation
   - 2.1 Calculation of posterior probabilities
   - 2.2 Marginalization over nuisance parameters
   - 2.3 Choice of prior probability distributions
   - 2.4 Point estimates and plausible intervals
   - 2.5 Hierarchical models, hyperparameters, and hyperpriors

3 Bayesian hypothesis testing
   - 3.1 Model comparison using the Bayes factor
   - 3.2 Bayesian decision theory

4 Numerical methods
   - 4.1 Markov chain Monte Carlo
   - 4.2 Gibbs sampler
   - 4.3 Hamiltonian Monte Carlo (time permitting)

2.2 Timetable for the Course

See the course outline at `outline.html`

2.3 Getting Started

There is a "week zero" of preliminary material consisting of an introduction and overview, and a brief R tutorial. This material has no due dates, but will be available 2020 January 6, one week before the semester starts. It is availble in mycourses and listed in the outline above.

2.4 Lessons

The course is divided into weeks. Each week includes a brief video introduction and one to three lessons, presented as Jupyter notebooks, available in mycourses. (The notebooks need to be saved to your computer and run within Jupyter.) These notebooks provide explanations interspersed

with R demonstrations, to be executed one step at a time. You are encouraged to tinker with the commands to explore how the demo changes if parameters, procedures, etc are modified.

2.5 Homework

Each week has a problem set, due the following Wednesday at 9am Eastern Time. The problem set is in the form of a Jupyter notebook, and is to be completed by including notebook cells with LaTeX/markdown (for explanations and formal calculations) and R commands (for numerical computations). Problem sets should be turned in with all of the cells excecuted. They should be zipped and submitted via mycourses, since mycourses does not accept `.ipynb` files for submissions. Solutions in the form of executed notebooks will be made avalable after the problem set is due. Problem sets will not be accepted after the solutions have been released. Note that homeworks will be checked for completeness, and a subset of the problems checked for correctness, so it's important to go through the solutions yourself. We will also have an area on the discussion forum for followup questions about the homework.

2.6 Exams

There will be two preliminary exams, one covering weeks 1-4 and one covering weeks 5-8. The exams will be open book, open notes, in Jupyter notebook format, and available for a 24-hour period. Once you download the exam, you have two hours to submit it. The final exam will be taken under similar conditions, but with a three-hour timespan.

2.7 Discussion Board

There is a discussion board in mycourses, on which you are encouraged to ask about and discuss both conceptual and practical aspects of the week's materials with me and your peers.

3 Course Policies

3.1 Student Identity Verification

As with all RIT Online courses, students must complete the Student Identity Verification Checklist.

3.2 Collaboration

It is acceptable and encouraged to discuss and brainstorm with your peers while doing the homework, but each student should turn in their own work. Collaborating on exams with people in or outside the course is of course cheating and will not be tolerated.

3.3 Grades

Grades will be based on the following components:

- Homework, Including Project [25%]
- First Prelim Exam [20%]
- Second Prelim Exam [20%]
- Final Exam [35%]

Your score on each component of the course (each prelim, the final, and all the homeworks together) will be converted to a numerical "grade point" score, and the weighted average of those five scores will be your final grade, converted to a letter grade according to the scale below.

3.4 Grading Scale

$3.8\overline{3}$ to 4.5	3.5 to $3.8\overline{3}$	$3.1\overline{6}$ to 3.5	$2.8\overline{3}$ to $3.1\overline{6}$	2.5 to $2.8\overline{3}$
A	A-	B+	B	B-
$2.1\overline{6}$ to 2.5	$1.8\overline{3}$ to $2.1\overline{6}$	1.5 to $1.8\overline{3}$	0.5 to 1.5	-0.5 to 0.5
C+	C	C-	D	F

3.5 Graded Feedback

I will check homeworks for completeness and give feedback on the correctness of a subset of the problems. Solutions will be made available, and **you are responsible** for going through your own homework submissions and learning from any mistakes. You will receive updates on your grades to date (a grade for each exam and a preliminary composite grade for the homeworks so far) three times during the semester: after each preliminary exam, and before the final exam. You are welcome to discuss with me your progress in between these milestones.

3.6 University Policies

3.6.1 Academic Integrity

As an institution of higher learning, RIT expects students to behave honestly and ethically at all times, especially when submitting work for evaluation in conjunction with any course or degree requirement. RIT Online encourages all students to become familiar with the RIT Honor Code and with RIT's Academic Integrity Policy.

3.6.2 Reasonable Accommodations

RIT is committed to providing reasonable accommodations to students with disabilities. If you would like to request accommodations such as special seating or testing modifications due to a disability, please contact the Disability Services Office. It is located in the Student Alumni Union, Room 1150; the Web site is `www.rit.edu/dso` . After you receive accommodation approval, it is imperative that you contact me so that we can work out whatever arrangement is necessary.

3.6.3 Use of copyrighted material

Certain materials used in this course are protected by copyright and may not be copied or distributed by students. You can find more information at `http://www.rit.edu/academicaffairs/policiesmanual/sectionC/C3_2.html`

3.6.4 Emergencies

In the event of a University-wide emergency course requirements, classes, deadlines and grading schemes are subject to changes that may include alternative delivery methods, alternative methods of interaction with the instructor, class materials, and/or classmates, a revised attendance policy, and a revised semester calendar and/or grading scheme.

3.6.5 Student support availability

Student Learning, Support & Assessment offers a wide range of programs and services to support student success including the Academic Support Center, College Restoration Program, Disabilities Services, English Language Center, Higher Education Opportunity Program, Spectrum Support program, and TRiO Support Services. Students can find out about specific services and programs at `www.rit.edu/slsa`