Math459: Bayesian Statistics

Spring 2018

Instructor:	Professor Nan LIN
	Office: Room 205, Cupples I
	Email: nlin@wustl.edu
Time and location:	11:30am-1pm TuThur and Sever 102
Office hours:	1-2pm Tuesday
TA:	Wei Wang(wang.w@wustl.edu)
TA office hour:	TBA

General information

Textbook: *Bayesian Data Analysis*, by Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. CRC Press/Taylor & Francis, 2013, 3rd Edition. ISBN: 9781439840955

Reference 1: *Bayesian Computation with R*, by Jim Albert. Springer, 2009, 2nd Edition. ISBN: 0387922970

Class webpage: All homework assignments, handouts, and other information will be available on Blackboard (http://bb.wustl.edu/). Students should check the class webpage frequently for updates.

Course Description

This is an advanced undergraduate/master level course that introduces the Bayesian approach to statistical inference for data analysis in a variety of applications. This course teaches the theory of Bayesian inference, and data analysis using statistical software (mostly R and some WinBUGS) will also be emphasized. Topics include: comparison of Bayesian and frequentist methods, Bayesian model specification, prior specification, basics of decision theory, Markov chain Monte Carlo, Bayes factor, empirical Bayes, Bayesian linear regression and generalized linear models, hierarchical models. The coverage of this class includes Chapters 1-6, 10-11, 13-16 of the Textbook and some additional topics selected by the instructor.

Prerequisite

- **Math and statistics**: Calculus; linear algebra; probability and statistics at the level of Math493 is required. Math494 is preferred. Students should have also taken at least one course on statistical models such as math322, math420, math434, math439, and math475, or equivalent. More specifically, this class assumes familiarity with the following topics:
  - Calculus
  - Matrix algebra, such as vectors, matrices, matrix inverse, determinants, linear transformation
  - Multivariate normal distribution
  - Joint, conditional, marginal distributions
  - Change-of-variable formula
  - Estimation
  - Bias, variance, covariance
  - Hypothesis testing

- **Programming**: Programming skills are essential for this class. Familiarity with R is required. If students do not know R but are familiar with languages such as Matlab, Stata and Python, you should be able to learn the basics of R by studying the online book "An Introduction to R" at `http://cran.r-project.org/doc/contrib/Lam-IntroductionToR_LHL.pdf` on Chapters 1-5, Chapter 7 (7.1-7.3) and Chapter 8 (8.1-8.3). Another useful online book on R is "simpleR - Using R for Introductory Statistics" at `http://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf`.

Course schedule

The following course schedule is tentative and subject to changes.

Week	Topic
1	Introduction, Comparison of frequentist and Bayesian statistics
2	One-parameter models
3	One-parameter models, **Simulation**
4	Prior specification: Conjugate prior, Jeffereys prior
5	Empirical Bayes, **brute-force posterior simulation**
6	Basics of decision theory, multi-parameter models
7	Multivariate models
8	Linear regression, asymptotic approximation to posterior distributions, Midterm
9	Spring break
10	Hierarchical model
11	Gibbs sampling, Metropolis-Hastings algorithm
12	Theory of MCMC, Convergence diagnostics for MCMC
13	**WinBUGS tutorial**, GLMM
14	latent variable model for ordinal data, Bayes factor
15	DIC, SSVS

- Items in bold require intensive computer work.

Computing

Students are **required** to use R (or WinBUGS) to complete all assignments. Both R and WinBugs are free software. R can be downloaded from http://cran.r-project.org/, and it works under major operating systems, including Windows, Linux and Mac OS. WinBUGS can be downloaded from http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml, but be aware that it works only under Windows. There are other versions of BUGS that work under Linux or Mac OS, but this class only covers WinBUGS.

For SAS users: R is very different from SAS, so knowledge of SAS does not help for this class. On the other hand, previous experience with Matlab or Python can be very helpful for those new to R.

Homework

There will be 6-7 homework assignments that will be assigned approximately every other week. Students will have one week to finish the homework and then submit in on Blackboard. If the homework is hand-written, it must be uploaded to Blackboard as a scanned pdf file. The grader will grade homework and assign a score for each homework set. Late homework submitted within 2 days of due date will receive 25% penalty for each day late. Any homework late by more than 2 days will not be graded and receive zero point.

Exams

There will be one midterm exam and one cumulative final exam (1-3pm, Monday, May 7). The midterm exam will be held in the regular class time on Thursday, March 8. Both exams will be closed book and closed notes. Students can bring a calculator to the exams, but sharing calculators is not allowed.

Make-up exams will **NOT** be given under any circumstances. If verifiable documentation is given for a legitimate absence, then your final exam grade will be reweighted. If a student misses the midterm exam, her/his final will count 60% of the final grade, instead of the usual 40%. However, an 'Incomplete' will be given if the final exam is missed no matter for what reason.

Grading

Grades will be based on the homework sets (40%), the midterm (20%), and the final exam (40%). Cr means D or better if you elect "Credit/No Credit." The final letter grade is given according to the following scale.

$[95, 100]$	$A+$	$[83, 85)$	$B+$	$[65, 75)$	C
$[87, 95)$	A	$[77, 83)$	B	$[60, 65)$	D
$[85, 87)$	$A-$	$[75, 77)$	$B-$	< 60	Fail

Learning Tips

1. Try to show up in all the lectures. Make good notes.
2. Ask questions in class. Your questions may be others' as well. No questions are too elementary, and all deserve to be answered.
3. Discuss with your classmates about your questions. It is perfectly acceptable to work together on homework assignments.
4. Finish homework in time.

Class Policies

1. <u>No auditing is allowed.</u> Students must formally register for this class.
2. Late homework: Late homework submitted **within 2 days of due date will receive 25% penalty** for each day late. Any homework late by more than 2 days will not be graded and receive zero point.
3. Exam conflicts: Prior permission and arrangement only. Students need to inform the instructor at least one week in advance.
4. Collaboration: I encourage discussion of homework in broad, conceptual terms where one student is trying to educate another without giving away the answer, but **all work turned in must be your own**. For example, each student must write his/her own programs in entirety.
5. Academic Integrity: All students are expected to adhere to the university's academic integrity policy. Any student who is found to have cheated on an assignment or exam will receive a zero score for that work, regardless of the extent of the offense.