Boston University
Department of Mathematics and Statistics

MA 578
Bayesian Statistics

Time: TR 12:30-1:45 PM
Location: PHO Building, Room 211
Website: Available through http://learn.bu.edu

Professor: Luis Carvalho
Department of Mathematics and Statistics
111 Cummington Mall Room 230
lecarval@math.bu.edu
(617) 358-5875
Office Hours: Mon 11:00AM-12:00 PM, Wed 10:00AM-12:00PM

Course Description:

This course will introduce students to the principles and methods of Bayesian statistics and provide practical experience analyzing data under this paradigm. Bayesian statistics offer a different interpretation of statistics based on subjective probability; parameters of interest are not just unknown as in classical, frequentist statistics – as seen in courses such as MA 213 / 214 – but are regarded as random and having initially a prior probability distribution that gets updated with the data likelihood according to Bayes's rule to yield a posterior distribution. Bayesian methodology and computational methods related to modeling, estimation, and simulation will be discussed. Topics include: inference using Monte Carlo integration and Markov chain Monte Carlo simulation, hierarchical models, mixture models, regression, model checking, and model selection.

Prerequisite: MA 214 Applied Statistics and MA 581 Probability, or their equivalents.

Textbook References:

Primary Textbook:

1. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B., *Bayesian Data Analysis*, 3rd Edition. Chapman & Hall/CRC

Supplementary Texts:

Methodology:

2. Sivia, D. and Skilling, J., *Data Analysis: A Bayesian Tutorial*, 2nd Edition. Oxford University Press.
3. Hoff, P. D., *A First Course in Bayesian Statistical Methods*, 1st Edition. Springer.

Practice:

3. Albert, J., *Bayesian Computation with R*, 2nd Edition. Springer.
4. Kruschke, J., *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. 2nd Edition. Academic Press.

Assignments:

Problem sets (40% of grade):

Problem sets will be assigned regularly, and due two weeks from when they are assigned. These assignments will typically involve some combination of theoretical and practical components. The practical components will involve downloading data sets from the course website, and analyzing them using the methods discussed in class. The purpose of these exercises is to gain experience applying Bayesian methods to real data. Please feel free to collaborate on these problem sets.

Data Analysis Software:

The preferred language for performing these analyses is **R**, but you are free to use any other statistical software or programming language you like (e.g. MATLAB/STATA/SAS/Python).

The R software package is now becoming a standard for statisticians. The software is free and compatible with Windows, Mac, and Linux/Unix. It may be downloaded from http://www.r-project.org. An alternative interface to R is Rstudio, available at http://www.rstudio.com.

Exams (20% of grade for each):

There will be a midterm around the 8th week of class and a final exam that will cover mostly Bayesian methodology. The exams will be open book and open notes for any notes you take in class, but you may not collaborate with anyone or use any other external references.

Class Paper (20% of grade):

The final project for the course will be a research or review paper, to be submitted by the last day of class. This paper should be in the format of a journal research manuscript or review article. Please choose one of the following topics for your paper:

1. I will provide a selection of data sets and a description of a Bayesian modeling problem for each. Choose one and use the methods discussed in class to explore the problem, analyze the the data, and make conclusions.

2. Use data collected in your lab or publicly available to formulate an analysis problem that uses Bayesian methods, explore the problem using the tools developed in class and draw conclusions about the data.

3. Identify and read the literature surrounding the application of Bayesian statistics to a particular field of study, and write a review article discussing the methods used and conclusions drawn in these analyses.

Tentative Schedule:

Week 1: Introduction: probability and inference
Week 2: Single parameter models
Week 3: Multiple parameter models
Week 4: Asymptotics and connections to frequentist approaches
Week 5: Hierarchical models
Week 6: Regression models
Week 7: Introduction to Bayesian computation: basic methods and tools
Week 8: Approximate Bayesian computation
Week 9: Advanced Bayesian computation: Markov chain Monte Carlo simulation
Week 10: Model checking and model selection
Week 11: Models for missing data
Week 12: Finite-mixture models
Week 13: Bayesian non-parametrics

Please Note:
Your conduct in this course, as with all BU courses, is governed by the *BU Academic Conduct Code*. The code is available at:
http://www.bu.edu/academics/policies/academic-conduct-code/