UNH, DURHAM FALL 2019

SYLLABUS CRN: 16770

MATH 941: BAYESIAN AND COMPUTATIONAL STATISTICS

Class Times: Mondays and Wednesdays 2:10 – 3:30 pm Eastern Time (USA)

Location: Kingsbury S 320, or anywhere live on-line

Instructor: Ernst Linder, N321B Kingsbury Hall, 603-862 2687, elinder@unh.edu

Office Hours: (*tentative*) M/W 3:45 – 4:45 + , or by appointment (this may change…)

COURSE OUTLINE

In this course, we will be discussing current approaches to Bayesian modeling and data analysis, and computation in statistics and related science fields. The course will be introductory in the beginning but will advance quickly to more current and especially computationally sophisticated methods. The emphasis will be on data, and on methodologies and their application in practice.

Prerequisites: *(Please contact the instructor if you are not sure if you have sufficient background)*

Knowledge of intermediate statistics: Distributions, discrete and continuous random variables, transformation of variables, bivariate and multivariate normal distributions.

Working knowledge of linear regression and analysis of variance.

Some very basic linear algebra: Vectors and matrices, linear spaces, matrix multiplication, inverse of a matrix, positive definiteness. Matrix-vector notation for linear regression and ANOVA

Familiarity with the basic ideas of calculus, in particular conceptual understanding of notation and meaning of integrals and derivatives in one or more dimensions.

R programming. *If you never used R before, **you are required** to work through the first 5 weeks of materials (including quizzes) of MATH 759/859 (1 credit) - Introduction to R.* You should complete this within the first 3 weeks of the semester. You may either register for the course, or ask me for access to the materials.

COURSE ORGANIZATION

Required Text: *(I spotted this text at the Joint Statistical Meetings in Denver, August 2019)*

- Reich, Brian,J and Ghosh, Sujit, K. **Bayesian Statistical Methods** (Chapman and Hall - CRC Press, 2019).
  Visit the text's website (Brian Reich): https://bayessm.org/

Recommended Text (an excellent resource, very comprehensive):

- Gelman, Carlin, Stern, Rubin, Dunson, and Vehtari, ***Bayesian Data Analysis*** (Chapman and Hall - CRC Press, third edition: 2014).

Visit the text website (by Andrew Gelman): http://www.stat.columbia.edu/~gelman/book/

Software: We will be using the following software
- R (open source, freeware)
- RStudio (a convenient platform for R)
- Stan (freeware)

Distance Learning: Course Delivery:

This course is delivered synchronously on-line using ZOOM. Try it by going to http://zoom.unh.edu ; join a "meeting" and play with the software. Try to set up your audio (microphone and loudspeakers) functions. Make sure that you are not producing echoes and extraneous noises when you have both, mic and loudspeakers on at the same time. Sometimes we recommend an earphone or a headset. Links to our class meetings will be posted on Canvas (myCourses.unh.edu) and class materials, including recordings will also be posted there.

ZOOM provides a fully interactive platform for course delivery and student participation. ***It is therefore expected that students attend the class either in person in the classroom, or, virtually from a computer during the scheduled class time.*** While the classes are recorded for replay at a later point in time, it is not acceptable for students to never attend class when it is scheduled and simply peruse the recording later. The recordings are to be used for re-play in cases where parts of the lecture were not clearly understood the first time. ***Students who anticipate having time conflicts with the scheduled classes need to inform the instructor of their planned absences from the live class.***

It is also expected that students who attend the class from a distance have the necessary equipment that allows them to fully participate in the course, in particular a microphone (a laptop mic tends to be sufficient).

Homework: Homework assignments will be given on a regular basis. Most homework will require computer-based calculations.

Quizzes (tentative): I may use the MyCourses on-line quiz menu occasionally if the material is suitable for that. Online quizzes can be useful to review material in a question-and-answer format. We may also make use of the Discussion Forums in MyCourses.

Final Exam: A comprehensive final take-home exam will be assigned at the end of the semester.

Grade Calculation: Homework, Quizzes and Discussions: 75 %, Final Exam: 25 %

Tentative List of Topics:

The aim is to cover most of the Reich & Ghosh text book.

1 Basics of Bayesian inference
1.1 Probability background
1.1.1 Univariate distributions
1.1.2 Multivariate distributions
1.1.3 Marginal and conditional distributions
1.2 Bayes' rule
1.3 Introduction to Bayesian inference
1.4 Summarizing the posterior
1.5 The posterior predictive distribution

2 From prior information to posterior inference
2.1 Conjugate priors
2.2 Improper priors
2.3 Objective priors

3 Computational approaches
3.1 Deterministic methods
3.2 Markov chain Monte Carlo (MCMC) methods
3.3 MCMC software options in R
3.4 Diagnosing and improving convergence

4 Linear models
4.1 Analysis of normal means
4.2 Linear regression
4.3 Generalized linear models
4.4 Random effects
4.5 Flexible linear models

5 Model selection and diagnostics
5.1 Cross validation
5.2 Hypothesis testing and Bayes factors
5.3 Stochastic search variable selection
5.4 Bayesian model averaging
5.5 Model selection criteria
5.6 Goodness-of-fit checks

6 Case studies using hierarchical modeling
6.1 Overview of hierarchical modeling
Case Studies

7 Statistical properties of Bayesian methods
7.1 Decision theory
7.2 Frequentist properties
7.3 Simulation studies