STAT 440: Statistical Computing
Spring 2018

Instructor: Matthew Reimherr
Department of Statistics
411 Thomas Building
mreimherr@psu.edu
Office Hours: Thursdays 10:30-11:30am and by Appointment

TA: Ilias Moysidis
ium19@psu.edu
Office Hours: Friday 11-12 and by Appointment, held in 301 Thomas
The TA will hold a weekly office hour for general questions. In addition, the TA will hold more detailed R sessions during the first few weeks for students that want some additional information regarding the basics of R.

Prerequisites: STAT 415 (Math Stat)
MATH 220 (Matrices)

Class Schedule: TH 9:05-10:20, Thomas 207

Textbook: *Statistical Computing with R*, by Maria L. Rizzo. This is not a required text, but can be a useful resource. Class notes and code will be provided.

Webpage: Available through CANVAS, check regularly for updates.

Course Topics:

1. Introduction to R, RMarkdown, and coding
2. Random number generation and simulation
3. Numerical and Monte Carlo integration
4. Markov Chain Monte Carlo
5. Resampling and permutation methods
6. Numerical linear algebra for fitting large linear models
7. Basics of nonlinear optimization
8. EM algorithm

Learning Objectives:

The major goal of this course is for students to become comfortable using R to carry out various statistical procedures learned in other courses. Students will code some methods themselves as well as learning details behind built in R functions or packages.

1. The first few weeks will be spent familiarizing students with R. Students will learn how to navigate RStudio and to learn about basic coding R, including how to install packages and create custom functions.
2. While learning R we will begin learning the basics of random number generation and simulation in R.
3. Students will learn about how to compute complex integrals, found in many statistical procedures, using either numeric integration or Monte Carlo in R.
4. Students will then learn how to use Markov Chain Monte Carlo for even more complex integrals and statistical calculations.
5. We will then shift to learning about resampling methods, such as the Bootstrap and the Jackknife, as well as permutation techniques.
6. We will finish the class by discussing various problems in optimization, especially as they relate to likelihoods. Basic convex optimization, the EM algorithm, numerical linear algebra, and other methods will be explored, as time permits.

Grading:

Homework: 50%
Midterm Project: 25%
Final Project: 25%

Homework:

Homework will be assigned most Fridays and will be due a week later. All homework must be turned in by **Friday** before the end of the day through CANVAS. **No late homework will be accepted and no scores dropped.** You are encouraged to work together on homework, but each student must turn in their own write up and answers. A homework that is complete with all questions seriously attempted will receive full credit. Missing answers will lose points at the discretion of the grader. Solutions to homework will be posted online and all students should carefully examine the solutions and compare with their own work.

Projects:

In place of exams, we will have two projects. Students will have 48 hours to complete these projects and then turn them in via canvas. While students are encouraged to work together on homework, they are not allowed to do so for these projects and are not allowed to discuss the projects in any way until after the 48 hour period. These projects must be the students' own work, though class notes, books, etc., may still be used.
Midterm Project: Feb 22-23
Final Project: TBA

R and RMarkdown:

All homework must be done in R and solutions prepared as an RMarkdown file. Students should turn in the markdown file as well as its html/pdf output via Canvas. Students will have access to the department's RStudio server that will have all of the

required R packages installed. It can be easily accessed via a web browser. Students can also install RStudio on their personal machines if they prefer.

Academic Integrity:

Academic integrity is the pursuit of scholarly activity free from fraud and deception and is an educational objective of this institution. All University policies regarding academic integrity apply to this course. Academic dishonesty includes, but is not limited to, cheating, plagiarizing, fabricating of information or citations, facilitating acts of academic dishonesty by others, having unauthorized possession of examinations, submitting work of another person or work previously used without informing the instructor, or tampering with the academic work of other students. All exam answers must be your own, and you must not provide any assistance to other students during exams.

ECOS Code of Mutual Respect:

The Eberly College of Science Code of Mutual Respect and Cooperation (http://science.psu. edu/climate/code-of-mutual-respect-and-cooperation) embodies the values that we hope our faculty, staff, and students possess and will endorse to make the Eberly College of Science a place where every individual feels respected and valued, as well as challenged and rewarded.

Disability Services:

Penn State welcomes students with disabilities into the University's educational programs. If you have a disability-related need for reasonable academic adjustments in this course, contact the Office for Disability Services (ODS) at 814-863-1807 (V/TTY). For further information regarding ODS, please visit the Office for Disability Services Web site at http://equity.psu.edu/ods/. In order to receive consideration for course accommodations, you must contact ODS and provide documentation. If the documentation supports the need for academic adjustments, ODS will provide a letter identifying appropriate academic adjustments. Please share this letter and discuss the adjustments with your instructor as early in the course as possible.