Autumn 2018

STAT 6730 Introduction to Computational Statistics

meetings:	MW 1:50–2:45 in Enarson 211 (Mondays) and EA 295 (Wednesdays)
instructor:	Vincent Q. Vu (vqv at stat osu edu)
office:	Cockins Hall 325
office hours:	Tuesdays 1:50–2:45, or by appointment
web:	Class schedule, assignments, and course announcements will be posted on Carmen (carmen.osu.edu)
prerequisites:	Stat 6410 and 6450, or Stat 6910 and 6950, or permission of the instruction

1 Overview

Computational statistics is an area within statistics that encompasses computational and graphical approaches to solving statistical problems. Students will learn how to manipulate data, design and perform simple Monte Carlo experiments, and be able to use resampling methods such as the Bootstrap. They will be introduced to technologies that are useful for statistical computing. Through creating customized graphical and numerical summaries students will be able to discuss the results obtained from their analyses. The topics of the course include:

1. Introduction to R
2. Dynamic and reproducible reports with R Markdown
3. Data manipulation in R
4. Visualization of data
5. Smoothing and density estimation
6. Generating random variables
7. Monte Carlo simulation
8. The Bootstrap
9. Permutation methods
10. Cross-validation

2 Course materials

The primary resource for reading will be slides and additional references assigned for reading by the instructor. There is one required book for the course that will be used for the parts of course dealing with data manipulation and visualization in R:

- Wickham, H. & Grolemund, G.: *R for Data Science*. O'Reilly Media.
  - Web version: `http://r4ds.had.co.nz`
  - Print version: `http://shop.oreilly.com/product/0636920034407.do`

  - Online access to print version: http://proxy.lib.ohio-state.edu/login?url=http://proquest.safaribooksonline.com/?uiCode=ohlink&xmlId=9781491910382 **Note that the web and print version have different chapter numbering.** I will list reading for both versions. You can also access the print version online via Safari Books using your OSU login.

The following software is required:

- R www.r-project.org
- RStudio www.rstudio.com

Students are expected to be able to access working installations of **current versions** of the required software on either their personal computer or a campus computer.

3 Prerequisites

This course is intended to be taking during the second year of the MAS program. It is expected that students will have exposure at a mathematical level to foundational concepts in probability and statistics including random variables, estimation, hypothesis testing, and linear regression. The formal prerequisites for this course are: Stat 6301 and 6302 or equivalent; Stat 6410 and 6450, or Stat 6910 and 6950, or permission of the instruction. Previous programming experience is not required, but familiarity with computer systems is expected.

4 Coursework & Grading

There will be homework, labs, two in-class exams, and a final exam:

- 15% Homework (due at the beginning of class on due date)
- 15% Labs (Wednesday meetings)
- 20% Exam 1 (October 3)
- 20% Exam 2 (November 7)
- 30% Final project

Inform the instructor of any scheduling conflicts at least two weeks in advance.

4.1 Homework and labs

Homeworks will generally be assigned on a biweekly basis. Students are required to use `R Markdown` for their homeworks. They should be written in a style that smoothly integrates prose, code, tables and graphics. **It should be human-readable**. Submit both the **source code** in Carmen Dropbox and **hard copy** of the generated document in class. Late homework will not be accepted. Homework will be graded on a 3 point scale: 1 point for good-faith effort, 1 point for technically-correct working solutions, 1 point for well-formated and easily-readable code.

Labs will consist of in-class exercises. Like homework, students are required to use `R Markdown` and the lab report should be written in a style that smoothly integrates prose, code, tables and graphics. **It should be human-readable**. Submit the **source code** (Rmd) to Carmen within 24 hours of the end of the class. Labs will be graded on the same 3 point scale as in homework.

Homework and lab assignments in the wrong format (e.g. `.pdf`, plain text, `.doc`) will receive 0 points automatically, no exceptions.

4.2 Exams

Both exams are open book/internet access, but absolutely no communicating with other humans. Each exam is cumulative.

4.3 Final project

Students will be assigned to small groups to work on a final data analysis project. The instructor will provide a list of topics. Each group will cooperate on the data analysis, report writing, and making a presentation on the project in class.

4.4 Fine Print

4.4.1 Academic Misconduct

It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term "academic misconduct" includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct `http://studentlife.osu.edu/csc/`.

4.4.2 Accommodations for Students with Disabilities

Students with disabilities (including mental health, chronic or temporary medical conditions) that have been certified by the Office of Student Life Disability Services will be appropriately accommodated and should inform the instructor as soon as possible of their needs. The Office of Student Life Disability Services is located in 098 Baker Hall, 113 W. 12th Avenue; telephone 614-292-3307, `slds@osu.edu`; `slds.osu.edu`.

4.4.3 Disclaimer

This syllabus is a approximate guide to the course content and dates, however the instructor reserves the right to deviate from the syllabus. An updated version of the syllabus will be maintained on the course webpage.