Spring 2023 - PM 520: Advanced Statistical Computing

COURSE INFORMATION:

Instructor:	Nicholas Mancuso, Ph.D.
E-mail:	nmancuso@usc.edu
Office hours:	TBD
Office location:	NRT G517G
TA:	Menglin Wang
E-mail:	menglinw@usc.edu
Office hours:	TBD
Office location	TBD
Course Schedule:	Mondays, 10am—1pm
Course Location:	SSB (Soto) 116

Course Objective:

The objective of the course is to give you the skills to enable you to both be able to program, and better understand, statistical machinery using the R programming language. Note that this is not really a course about R, as such. Instead, we use R as a way of illustrating how to write code to perform a number of probabilistic and statistical procedures. We will introduce the R details as we go, but students will be expected to invest some time of their own each week, outside class, in which they improve their R programming skills if necessary.

Each week's lecture (60-90 minutes) will focus on one or more areas in probability and statistics, describing how those procedures are performed. The rest of the class (90-120 minutes) will be a lab in which students will sit down and start to code those procedures in class. Students will be encouraged to show their work "in progress". For example, you've written 30 lines of code and it isn't working properly: we will put the code on the screen, and we will all try to work out what is wrong. In other words, we will learn by doing, and by learning from our mistakes, rather than relying upon the traditional method of listening to a member of faculty drone on for 3 hours. In doing so we hope to make the course more interactive. As the saying goes: turn-about is fair play. So, as the instructor, I will also display bugged code of my own to the rest of the class, so that we can learn from my mistakes as well.

The course will be examined by several projects during the course. These projects will involve writing R code to perform a particular task (optimize a function, fit a regression, simulate a stochastic process, …). That code will be turned in to me and I will run it on some test problems to see how it does. In addition, the final exam will consist of your giving small group (2-3 people) presentations regarding statistical computing methods.

Recommended Textbook: "Scientific Programming and Simulation Using R" – Owen Jones, Robert Maillardet and Andrew Robinson, CRC Press.
Available at **USC Library (*online for free!*)**, USC Bookstore (?), and Amazon.

> **PLEASE NOTE** that ***each student will also be required to bring a laptop to each class*** so that they can actively write code during the lab.

Other Material: Copies of slides will be posted, as will example bits of code. But there are a host of free online resources for R:

- Mastering Software Development in R, 2017. Roger Peng, Sean Kross, Brooke Anderson. https://bookdown.org/rdpeng/RProgDA/
- R for Data Science, 2017 Garrett Grolemund and Hadley Wickham. http://r4ds.had.co.nz/
- Advanced R, 2014, Hadley Wickham, CRC press. http://adv-r.had.co.nz/ (Second edition is available in print on June 28 https://www.amazon.com/gp/product/0815384572/, online here: https://adv-r.hadley.nz/)
- R Packages, Hadley Wickham, O'Reilly. https://r-pkgs.org/
- Advanced Statistical Computing, 2018 Roger Peng https://bookdown.org/rdpeng/advstatcomp/

Grading: 70% projects + 20% final presentation + 10% participation.

Grading Scale: Final grades will be determined on the percentage of points earned by each of the assignments and exams described above. Course final grades will be determined using the following scale.

A	95-100
A-	90-94
B+	87-89
B	83-86
B-	80-82
C+	77-79
C	73-76
C-	70-72
D+	67-69
D	63-66
D-	60-62
F	59 and below

Academic integrity:

A grade of zero will be applied to submitted work that does not comply with the USC standards of academic conduct. Such work may not be resubmitted for a new grade. Academic integrity info is included at the end of the syllabus.

Pre-requisites:

None, but you will need to have some basic statistics and probability. If you are not sure, come and see me to talk about it.

COURSE OUTLINE:

Most weeks will be based upon one of the chapters from the course text, but some of the latter weeks will digress from the text and use material that will be distributed or described in class [labeled "(external material)"].

Week 1: Introduction – Using Github; Random number generation and Monte Carlo Estimation. How to estimate things that you cannot calculate. Coin-tossing. Occupancy problems. Hypercubes. Golf balls.

Week 2: More Monte Carlo Estimation: Estimating pi, Random variable simulation, Likelihood Estimation. Bayesian methods. Accept/Reject Algorithms (external material).

Week 3: Methods for finding function roots and fixed points. Math as art. (Chapter 10 of course text).

Week 4: Probability and Stochastic Simulation – Urn Models and Chinese Restaurants (Chapter 18 in 1st edition; or *Chapter 20, in the 2nd ed.).*

Week 5: Optimization and Regression (chapter 12)

Weeks 6-8: Markov Chain Monte Carlo [MCMC] Methods. Adaptive MCMC, Parallel Tempered-MCMC, Code-breaking (external material).

Week 9: Gibbs Sampling, Accept/Reject Algorithms (external material).

Week 10: Permutation tests, Numerical Integration and Importance Sampling (Chapters 19 and 22; *chaps. 21 and 24 in the 2nd ed.*).

Week 11: Approximate Bayesian Computation, ABC-Rejection, ABC-MCMC (external material).

Week 12: Sequential Monte Carlo Methods, Regression-adjusted ABC (external material)

Week 13: E-M algorithm; Bootstrapping. (external material).

Closing Project: Short presentation from a topic we haven't covered in the course: e.g., fancy MCMC versions, (e.g., Hamiltonian MC, Langevin MC, Stochastic Gradient-based MC), or a related project from your own research.

STATEMENT FOR STUDENTS WITH DISABILITIES:
Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me (or to TA) as early in the semester as possible. DSP is located in STU 301 and is open 8:30 a.m.–5:00 p.m., Monday through Friday. The phone number for DSP is (213) 740-0776.

STATEMENT ON ACADEMIC INTEGRITY:
USC seeks to maintain an optimal learning environment. General principles of academic honesty include the concept of respect for the intellectual property of others, the expectation that individual work will be submitted unless otherwise allowed by an instructor, and the obligations both to protect one's own academic work from misuse by others as well as to avoid using another's work as one's own. All students are expected to understand

and abide by these principles. *Scampus,* the Student Guidebook, contains the Student Conduct Code in Section 11.00, while the recommended sanctions are located in Appendix A: http://web-app.usc.edu/scampus/gov/. Students will be referred to the Office of Student Judicial Affairs and Community Standards for further review, should there be any suspicion of academic dishonesty. The Review process can be found at: http://www.usc.edu/student-affairs/SJACS/.

Academic Conduct:

Plagiarism – presenting someone else's ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in *SCampus* in Part B, Section 11, "Behavior Violating University Standards" policy.usc.edu/scampus-part-b. Other forms of academic dishonesty are equally unacceptable. See additional information in *SCampus* and university policies on scientific misconduct, http://policy.usc.edu/scientific-misconduct.

Support Systems:

Student Counseling Services (SCS) – (213) 740-7711 – 24/7 on call
Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.
engemannshc.usc.edu/counseling

National Suicide Prevention Lifeline – 1 (800) 273-8255
Provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week. www.suicidepreventionlifeline.org

Relationship and Sexual Violence Prevention Services (RSVP) – (213) 740-4900 – 24/7 on call Free and confidential therapy services, workshops, and training for situations related to genderbased harm.
engemannshc.usc.edu/rsvp

Sexual Assault Resource Center
For more information about how to get help or help a survivor, rights, reporting options, and additional resources, visit the website: sarc.usc.edu

Office of Equity and Diversity (OED)/Title IX Compliance – (213) 740-5086
Works with faculty, staff, visitors, applicants, and students around issues of protected class.
equity.usc.edu

Bias Assessment Response and Support
Incidents of bias, hate crimes and microaggressions need to be reported allowing for appropriate investigation and response. studentaffairs.usc.edu/bias-assessment-response-support

The Office of Disability Services and Programs
Provides certification for students with disabilities and helps arrange relevant accommodations.
dsp.usc.edu

Student Support and Advocacy – (213) 821-4710
Assists students and families in resolving complex issues adversely affecting their success as a student EX: personal, financial, and academic. studentaffairs.usc.edu/ssa

Diversity at USC
Information on events, programs and training, the Diversity Task Force (including representatives for each school), chronology, participation, and various resources for students. diversity.usc.edu

USC Emergency Information
Provides safety and other updates, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible. emergency.usc.edu

USC Department of Public Safety – UPC: (213) 740-4321 – HSC: (323) 442-1000 – 24-hour emergency or to report a crime.
Provides overall safety to USC community. dps.usc.edu