UNIVERSITY OF SOUTHERN CALIFORNIA

DEPARTMENT OF PREVENTIVE MEDICINE

SYLLABUS: PM 599: BAYESIAN STATISTICS AND COMPUTATIONAL METHODS/ SUMMER 2008

Instructor: Paul Marjoram, Ph.D.
Soto Bldg 202V, HSC
Phone: (323) 442-0111
Email: pmarjora@usc.edu

Teaching Assistants: TBD

Office Hours : TBD

Class Meets: Thursday, 12noon-3:00pm.

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 entirely by numerous small projects, and one larger project towards the end of class. 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.

Finally, we note that since this is the first time this course has been taught, its exact contents are somewhat open-ended in nature. If, as we go, the class becomes interested in a particular area, we will focus more time than planned on going deeper into that area, at the expense of other planned material. As such, I encourage you to become fully involved with the class and the directions it might take. The more involved you become, the more enjoyable it will be.

Required Textbook: "Scientific Programming and Simulation Using R" – Owen Jones, Robert Maillardet and Andrew Robinson, CRC Press. This book will not (I think) be available from the USC bookstore, but can ordered from Amazon ($70 to buy; $50 to rent).
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

Grading: By projects.

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.

PROPOSED 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)"].

Weeks 1-2: Introduction to R. Methods for finding function roots and fixed points. (Chapter 10 of course text).
Week 3: Numerical Integration (Chapter 11).
Week 4-5: Optimization and Regression (Chapter 12).
Week 6: Probability and Stochastic Simulation (Chapter 18).
Week 7: Monte Carlo Estimation: Likelihood Estimation and Numerical Integration (Chapter 19).
Week 8: Case Study: Epidemics (Chapter 21).
Week 9: Introduction to Final Project (Chapter 22).
Week 10: Markov Chain Monte Carlo [MCMC] Methods (external material).
Week 11: Accept/Reject Algorithms (external material).
Week 12: Importance Sampling (Chapter 22).
Week 13: Approximate Bayesian Computation (external material).
Week 14: Expectation-Maximization Algorithms and Regression (external material)
Week 15: Recap.

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://www.usc.edu/dept/publications/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/.