Statistics 221 – Statistical Computing Methods

Instructor: Mark Irwin   Office Hours: Tuesday 1:30 – 2:30,
Office: Science Center 235   Thursday 10:30 – 11:30,
Phone: 617-495-5617   or by appointment
e-mail: irwin@stat.harvard.edu   Web site: http://www.courses.fas.harvard.edu/~stat221/

Objectives

A study of computing methods commonly used in statistics. Topics include generation of random numbers, Monte Carlo methods, optimization methods, numerical integration, resampling methods such as the Bootstrap and the Jackknife, and advanced Bayesian computational tools such as the Gibbs sampler, Metropolis Hastings, the method of auxiliary variables, marginal and conditional data augmentation, slice sampling, exact sampling, and reversible jump MCMC. Computer programming exercises apply the methods discussed in class.

Prerequisites

Linear algebra, Statistics 111, and knowledge of a computer programming language. Statistics 220 is recommended.

Lectures

Wednesday and Friday, 3:30 – 5:00, Science Center 102B

Grading

Course grades will be determined based on homework assignments (4 or 5 during the term) and the term project, which will involve discussing a recent and/or important paper involving statistical computing issues in lecture. A list of suggested papers for the term project will be presented later in the term.

Computing

There will not be a set package or language for this course. However it is suggested that programming projects be completed in Matlab or S-Plus/R. For many problems, Matlab will probably be preferable.

Textbooks and References

Required Text:

Lange K (1999). Numerical Analysis for Statisticians. Springer-Verlag.

Optional Texts:

Gilks WR, Richardson S, and Spiegelhalter DJ (1996). Markov Chain Monte Carlo in Practice. Chapman & Hall / CRC.

Gelman A, Carlin JB, Stern HS, and Rubin DB (2004). Bayesian Data Analysis, (2nd edition). Chapman & Hall / CRC.

Other References:

Efron B and Tibshirani R (1993). An Introduction to the Bootstrap. Chapman and Hall.

Liu J (2001). Monte Carlo Strategies in Scientific Computing. Springer-Verlag.

Ripley BD (1987). Stochastic Simulation. Wiley.

Tanner M (1996). Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions, 3rd Edition. Springer-Verlag.

Thisted RA (1988). Elements of Statistical Computing: Numerical Computation. CRC Press.

Splus / R:

Venables WN and Ripley BD (2002). Modern Applied Statistics with S (4th edition), Springer-Verlag.

Venables WN and Ripley BD (2002). S Programing, Springer-Verlag.

Krause A and Olson M (2002). Basics of S-PLUS, Springer-Verlag.

Matlab

Hanselman D and Littlefield BR (2000). Mastering Matlab 6, Prentice Hall.

Marchand P (1999). Graphics and GUIs with MATLAB, (2nd edition), CRC Press.

Other references and papers will be added to the web site during the course.