Monte Carlo Methods (ORIE 7791)
Spring 2009, 3 credits.
Prof. D. Woodard

Summary
Monte Carlo methods for integration and counting problems arising in Bayesian statistics, statistical physics, and theoretical computer science. Markov chain techniques, including Metropolis-Hastings, Gibbs sampling, data augmentation, and parallel tempering. Convergence diagnosis and validation of results. Rates of convergence, mixing times, and central limit theorems. Proven improvements in efficiency due to the use of Markov chain Monte Carlo, e.g. for approximating the permanent of a matrix and sampling from a multivariate normal distribution. Computational complexity of the integration problem.

Prerequisites
PhD students only. Coursework (senior level or above) in statistics, linear algebra and stochastic processes.

Course web sites
Materials for the class can be found on Blackboard once you have registered:
http://blackboard.cornell.edu
The course website contains the syllabus:
http://people.orie.cornell.edu/woodard/monteCarloMethods.html
Visit Blackboard before the next class to make sure that you can access the course information, and to sign up for the course email list.
Please also visit the ORIE intranet site to request an account:
http://intranet.orie.cornell.edu

Instructor
Prof. D. Woodard
228 Rhodes Hall
Web page (contains contact information):
http://people.orie.cornell.edu/woodard
Office hours: Mondays and Fridays 3-3:50 and by appointment (no drop-ins)
Only very short questions that have not been previously answered in class will be answered via email. For this reason please make sure that one of the office hours is at a time when you can attend if necessary.

Lectures
Lectures are Tuesdays and Thursdays 1:05-2:20 PM in Phillips 307

Exams

Prelim 1 is in class on **Thursday, March 12**. Prelim 2 is in class on **Thursday, April 30**. **No makeup exams will be given**, so class attendance is mandatory on these days. Class attendance is also mandatory **Thursday, April 16** and **Tuesday, April 21**; these are the project presentation days.

Homework

A hard copy of your homework solutions must be put in the course drop box (on the second floor of Rhodes) before the scheduled due time; no homework assignments are accepted via email, submission in person, or any other manner. You may discuss the content of the homework with other students in the class, but your submitted product must be your own work. Your lowest homework grade will be dropped; this is in order to accommodate sickness, family emergency, etc. without a formal process. If you miss a single homework due to sickness etc. then it must count as your dropped assignment.

Textbooks

Required:

- Robert and Casella (2004), Monte Carlo Statistical Methods

Optional:

- Liu (2004), Monte Carlo Strategies in Scientific Computing

One copy of the Robert and Casella text is on reserve in the Engineering library.

Software

S-PLUS 8.1 with the FlexBayes library (included with S-PLUS 8.1) and its dependencies. A free student version of S-PLUS (I think it is version 8.1) is available at http://elms03.e-academy.com/splus/
Instructions for installing the FlexBayes dependencies are available by opening S-PLUS and calling library(FlexBayes) and then help(FlexBayes).

Good references for the S-PLUS programming languages and its sister language R include:

- "An Introduction to R", found at http://www.r-project.org/
- The "User's Guide" or "Getting Started Guide" for S-PLUS
- The book "Data Mining with R"

Grading

Grade allocation is as follows: 10% homework, 25% project and 65% exams. In case of a grading error you may resubmit the assignment or exam within one week of when it was given back to you, with a written explanation of the grading error. The entire assignment or exam will be regraded, so your final grade may be lower or higher than your original grade. Any errors in your recorded grades on Blackboard must be reported by one week after the last day of class in order to be corrected for your final grade.

Project with presentation

Choice of the following:

- Implementation and application of a method from class
- Present another Markov chain method (e.g. reversible jump)
- Present a paper, such as a mixing time analysis of a Markov chain algorithm

Projects will be presented during the last week of classes.

Academic Integrity

Any violations of the Cornell Code of Academic Integrity or the principle of academic integrity will be punished at minimum by failure of the course. On the course web page there is a link to Cornell's Code of Academic Integrity. Be sure that you are familiar with this code.

Schedule

Week	Topic	Reading
Week 1	Motivation • Bayesian statistics • Statistical physics • Theoretical computer science	
Week 2	Deterministic methods for integration Monte Carlo integration	
Week 3	Random variable generation	
Week 4	Markov chains	
Week 5	Metropolis-Hastings	
Week 6	Convergence diagnosis	
Week 7	Gibbs sampler	
Weeks 8 & 9	Data augmentation	
Week 10	Central Limit Theorems Rate of convergence	
Week 11	Strategies for speeding convergence	
Week 12	Proven improvements in efficiency due to use of Markov chains	
Week 13	Computational complexity of the integration problem	
Week 14	Validation of Markov chain results for Bayesian statistics	