CSCI 1450: Probability and Computing

Brown University, Spring 2016

Probability and statistics have become indispensable tools in computer science. Probabilistic methods and statistical reasoning play major roles in machine learning, cryptography, network security, communication protocols, web search engines, robotics, program verification, and more. This course introduces the basic concepts of probability and statistics, focusing on topics that are most useful in computer science applications. Topics include: modeling and solution in sample space, random variables, simple random processes and their probability distributions, Markov processes, limit theorems, and basic elements of Bayesian and frequentist statistical inference.

Prerequisites: Two semesters of college-level calculus, at the level of MATH 0090 and MATH 0100. Basic programming experience required for homework assignments.

Administrative Information

Textbooks: *Introduction to Probability*, second edition. Dimitri P. Bertsekas & John N. Tsitsiklis, Athena Scientific, 2008. *Probability* (supplemental readings). Jim Pitman, Springer, 1993.

Lectures: Tuesdays and Thursdays from 2:30-3:50pm, 85 Waterman Street, room 130.

Recitations: Mondays from 5:30-6:30pm, CIT room 368. Led by the graduate teaching assistant.

Instructor: Prof. Erik Sudderth (sudderth@cs.brown.edu; 401-863-7660; CIT room 555)

Graduate Teaching Assistant: David Meierfrankenfeld (nfelddav@cs.brown.edu; CIT room 453)

Undergraduate Teaching Assistants: Nikhil Patel (Head UTA), Tiffany Citra, Vincent Kubala, Ivaylo Petrov, Christopher Robotham, & Isaac Semaya

Office Hours: See the course website for a detailed schedule.

Exams and Course Grades

Overall course grades will be assigned as follows: 50% homeworks, 20% midterm exam, 30% final exam. The midterm exam will be given during the normal lecture time on Thursday, March 17. The final exam will be given on Friday, May 20 from 2:00pm-5:00pm. *Exams **must** be taken at these times. Exceptions are granted only for medical or family emergencies.*

Homework Assignments

There will be ten homework assignments, each due one week after it is handed out. Homework problems will emphasize probabilistic derivations, calculations, and reasoning. Most homeworks will also have one problem requiring Matlab implementation of simple methods from probability or statistics. The scores of all ten assignments will be averaged equally to determine an overall homework score (we will not "drop" any homeworks). Homeworks will be submitted electronically.

Collaboration Policy Students may discuss and work on homework problems in groups. However, each student must write up their solutions independently, and do any required programming independently. You may *not* directly copy solutions from other students, or from materials distributed in previous versions of this or other courses. You may *not* make your solutions available to other students: files in your home directory may not be world-readable, and you may not post your solutions to public websites.

Late Submission Policy Homework assignments are due by 11:59pm on Thursday evenings. Your answers may be submitted up to 4 days late (by Monday evening); after this point, solutions will be distributed and handins will no longer be accepted. You may submit up to two late assignments without penalty. For each subsequent late assignment, 20 points (out of a maximum of 100) will be deducted from the overall score. *Exceptions to this policy are only given in very unusual circumstances, and any extensions must be requested in advance by e-mail to the instructor.*

Syllabus: Summary of Course Topics

Probability Models sample spaces, axioms of probability, sets and counting, conditioning, Bayes' rule, independence

Discrete Random Variables probability mass functions, expectation and moments, conditioning and independence, functions of random variables, conditional expectation

Continuous Random Variables probability density functions and cumulative distributions, expectation and moments, conditioning and independence, functions and derived distributions

Normal Distributions covariance and correlation, bivariate distributions, linear regression

Limit Theorems Markov and Chebyshev inequalities, weak law of large numbers, convergence in probability, central limit theorem

Discrete-time Markov Chains classification of states, steady-state behavior and equilibrium distributions, absorption probabilities

Monte Carlo Methods pseudo-random number generation, Monte Carlo integration

Bayesian Statistical Inference posterior distributions, hypothesis testing, maximum a posteriori (MAP) estimation, least mean squares estimation

Frequentist Statistical Inference maximum likelihood parameter estimation, Neyman-Pearson hypothesis testing, significance tests