STAT 721: Stochastic Processes

Spring 2024

Instructor:	Ray Bai	**Time:**	1:10-2:00 pm MWF
Email:	RBAI@mailbox.sc.edu	**Place:**	LeConte 107

Course Page: `https://blackboard.sc.edu/` (Check regularly for updates)

Office Hours: By appointment. I am also very accessible by e-mail.

Course Description: Stochastic processes are probabilistic (non-deterministic) systems indexed by time or space. They are very useful for modeling a variety of phenomena that vary in a random manner, such as the volatility of financial assets or the spatial distribution of crimes in a city. Stochastic processes are also very useful for many machine learning tasks such as regression, clustering, and reinforcement learning.

This course introduces methodology, applications, and theory for stochastic processes. It focuses on the following five topics: 1) point processes, 2) mathematical finance, 3) Bayesian nonparametrics, 4) Markov chain Monte Carlo (MCMC), and 5) reinforcement learning. The tentative schedule of topics is as follows:

- **Weeks 1-2**: point processes (Poisson processes, spatial point processes, nonhomogeneous Poisson processes, simulation algorithms)
- **Weeks 3-5**: mathematical finance (pricing of European options, random walks, Brownian motion, stochastic differential equations, binomial model, Black-Scholes model)
- **Weeks 6-8**: Gaussian processes (Bayesian inference, stationary and nonstationary kernels, GP regression, empirical Bayes, scalable GPs for big data)
- **Week 9-10**: Markov chain Monte Carlo (Markov chains, stationarity and convergence, Metropolis-Hastings, Gibbs sampling, Hamiltonian Monte Carlo, MCMC for Bayesian inference)
- **Week 11**: Dirichlet processes (Kolmogorov consistency definition, stick-breaking process, Chinese restaurant process, Dirichlet process mixture models)
- **Weeks 12-14**: reinforcement learning (Markov decision process, dynamic programming, exploration vs. exploitation, Q-learning, policy gradients, deep reinforcement learning)
- **Week 15**: class project presentations

Learning Outcomes:

1. Be able to apply stochastic processes to a variety of tasks, including spatial statistics, financial modeling, Bayesian inference, regression, clustering, and training intelligent agents.
2. Develop programming and algorithm design skills by implementing the models covered in the course.
3. Understand the theory of stochastic processes for point processes, option valuation, Bayesian inference, MCMC, and reinforcement learning and appropriately apply this theory to solving problems.
4. Practice effective communication skills through writing scientific reports and making presentations.

Prerequisites:

- STAT 712 **or** STAT 511/512 (or equivalent)
- MATH 344 or MATH 544 (or equivalent)

Students must also be able to write functions in one programming language such as R, Python, MATLAB, or C++. These are strict requirements. Students who cannot code basic functions or who have not taken probability/statistics/linear algebra at the levels specified above may not register for the course. Students who have only taken STAT 712 may register for the class. However, if you have *not* taken STAT 712, then you must have taken *both* STAT 511 and STAT 512.

Main References: We will use typed handouts prepared by the instructor.

Computing: This course involves programming. Please use one of the following languages: Python, R, MATLAB, C, or C++.

Homework: There will be five homework assignments which consist of both theoretical and computational exercises. Your answers to theoretical exercises may be handwritten and can be turned in to the instructor in person. However, the answers to the programming exercises must be typed. Your answers to programming exercises are to be submitted through Blackboard, along with the computer code.

Students are also allowed to collaborate with other students in the class, but the collaboration must be acknowledged on the homework that you hand in.

Project: Students will research a topic of their choosing, prepare a 15-minute presentation, and write a short report in the style of a journal article: abstract, introduction, method, data analysis, and a bibliography. The last week of the semester will be devoted to project presentations. Some potential examples of projects include:

- numerical methods for pricing American options
- Hawkes process for modeling social media contagion
- Vecchia approximation methods for Gaussian processes
- Bayesian additive regression trees
- Monte Carlo Tree Search or other reinforcement learning algorithms not covered in lecture
- Gillespie algorithm for simulating stochastic equations of molecular reactions

Projects must be approved in advance by the instructor, and no two students may do the same topic for their project. If you have an idea of what you want to do for your project, please "claim" it early. Detailed instructions for the presentation and the report will be given at a later date.

Grading: Your grade will be determined by homework (70%) and the project (30%). The tentative grading scale is as follows: 90-100 for an A, 80-89 for a B+, 70-79 for a B, 60-69 for a C+, 0-59 for a C.

Honor Code: See the Carolinian Creed in the *Carolina Community: Student Handbook and Policy Guide*. The *minimum* punishment for violations of the USC Honor Code is a grade of zero for the work in question. In accordance with university policy, there may be other punishments, including an automatic F in the class and/or expulsion from the university.

Accommodation: If you need special accommodations for any aspects of the course, please contact me before or during the first week of the semester. Note that reasonable accommodations are available for students with a documented disability. If you have a disability and may need accommodations to fully participate in this class, contact the Office of Student Disability Services by phone (803-777-6142) or e-mail sasds@mailbox.sc.edu. All accommodations must be approved through the Office of Student Disability Services.