STAT 721: Stochastic Processes

Instructor: Ray Bai, PhD

Department of Statistics, University of South Carolina

January 8, 2024

STAT 721 Spring 2024 Class Logistics

Class Meetings: MWF 1:10-2:00 pm
Place: LeConte 107
Course Webpage: `http://blackboard.sc.edu/`

Grading:

- 70% homework
- 30% project

A: 90-100, B+: 80-89, B: 70-79, C+: 60-69, C: 0-60

Answers to theoretical exercises may be handwritten and turned in to the instructor in-person. Answers to coding exercises should be typed, and the answers and code should be submitted through Blackboard.

The project report must be typed. See project guidelines.

Computing: Please use one of the following languages for computing: R, Python, MATLAB, or C/C++.

STAT 721 Learning Outcomes

Apart from learning about stochastic processes, the other main learning outcomes from the course are as follows:

1. Instill good programming practices, such as:
   - proper exception/error handling;
   - optimizing code for runtime **and** memory.

   The best way to learn/improve coding skills is simply to do it, not watch the instructor read lines of code. The homework exercises are designed to teach coding. Students should ask the instructor for help debugging their code.

2. Learn to set up numerical experiments/simulation studies, analyze real datasets, and present the results through figures and tables.

3. Practice oral and written communication through giving presentations and writing a project report in the style of a journal or conference article.

Stochastic Processes

A **stochastic process** is a collection of random variables $\{Y(x) : x \in \mathcal{X}\}$, where x is a (possibly multi-dimensional) parameter that runs over an index set $\mathcal{X}$. For example:

- $\mathcal{X}$ could be a discrete index set, i.e. $\mathcal{X} \subseteq \mathbb{N}$ or $\mathcal{X} \subseteq \mathbb{Z}$.
- $\mathcal{X}$ could be a continuous time interval $[0, T] \subset \mathbb{R}^+$.
- $\mathcal{X}$ could be a continuous domain $\mathcal{A} \subset \mathbb{R}^d, d \geq 1$. For instance, $\mathcal{A}$ could be a *spatial* domain (for point processes) or simply the domain of the covariates (for Gaussian processes).

We can alternatively write a stochastic process as $\{Y_x : x \in \mathcal{X}\}$. If the index set is time, we frequently use t, i.e. $Y(t)$ or Y_t.

Our aim is to study stochastic processes as modeling tools.

Applications of Stochastic Processes

- Modeling temporal and/or spatial patterns of natural disasters, queuing systems, spread of infectious disease, etc. (**Poisson processes**)
- Modeling the prices and returns of financial instruments and assets (**random walks and geometric Brownian motion**)
- Nonparametric regression (**Gaussian processes**)
- Simulating from intractable probability distributions and Bayesian computation (**Markov chain Monte Carlo**)
- Clustering and discovering *new* grouping structures such as new semantic structures in natural language processing, new disease subtypes, etc. (**Dirichlet processes**)
- Training intelligent agents such as self-driving vehicles, algorithmic trading programs, precision medicine software, etc. to learn optimal actions (**reinforcement learning**)

STAT 721 Topics and Applications

Stochastic Processes	**Applications**
point processes	**temporal and spatial statistics**
random walks, Brownian motions	**mathematical finance**
Gaussian processes	**nonparametric regression**
Discrete-time Markov chains	**Markov chain Monte Carlo**
Dirichlet processes	**clustering**
Markov decision processes	**reinforcement learning**