Syllabus for Math 319/419:
Applied Probability and Stochastic Processes for Biology
(short title: Biological Stochastic Processes)
Crosslistings: BIOL/EECS/MATH/SYBB 319 and BIOL/EBME/MATH/PHOL/SYBB 419

Instructor: Peter Thomas
Assoc. Prof. of Mathematics, Applied Mathematics, and Statistics
Yost Hall 212

Spring Semester 2018

Catalog Description

Applications of probability and stochastic processes to biological systems. Mathematical topics will include: introduction to discrete and continuous probability spaces (including numerical generation of pseudo random samples from specified probability distributions), Markov processes in discrete and continuous time with discrete and continuous sample spaces, point processes including homogeneous and inhomogeneous Poisson processes and Markov chains on graphs, and diffusion processes including Brownian motion and the Ornstein-Uhlenbeck process. Biological topics will be determined by the interests of the students and the instructor. Likely topics include: stochastic ion channels, molecular motors and stochastic ratchets, actin and tubulin polymerization, random walk models for neural spike trains, bacterial chemotaxis, signaling and genetic regulatory networks, and stochastic predator-prey dynamics. The emphasis will be on practical simulation and analysis of stochastic phenomena in biological systems. Numerical methods will be developed using a combination of `MATLAB`, the `R` statistical package, `MCell`, and/or `URDME`, at the discretion of the instructor. Student projects will comprise a major part of the course. Offered as BIOL 319, EECS 319, MATH 319, SYBB 319, BIOL 419, EBME 419, MATH 419, PHOL 419, and SYBB 419 .

Logistics

- **Course Meetings:** Mondays & Wednesdays 2:15-3:30 p.m.
- **Location:** CWRU active learning classroom Thwing 101.
- **Office Hours:** Yost Hall room 212 MWF 9:45-10:30 a.m., and A.W. Smith Hall room 329 MWF 11:25-11:40 a.m., and by appointment (pjthomas—@—case—dot—edu), or 216-386-3623.

Background & Topics

Mathematical models of biological systems frequently involve systems of ordinary or partial differential equations. While these deterministic models can give important insights into biological behavior, they overlook the effects of molecular fluctuations on biological dynamics. This course will explore applications of probability theory and stochastic processes in biological systems. It is a natural extension of the biological dynamics courses (BIOL 300 or BIOL 306) or a first course in differential equations (MATH 224 or 228) or a first course in mathematical modeling (MATH 441) and any of these can serve as a prerequisite. Students should be comfortable with multivariable calculus (MATH 223 or 227) and linear algebra (MATH 201 or MATH 307). A first course in probability (MATH 308) while helpful background, is not strictly required. While the mathematical content will be appropriate for a 300 level undergraduate or 400 level introductory graduate course, the emphasis will be on applications, and on practical matters such as numerically simulating stochastic phenomena in biological systems using numerical platforms such as the `R` statistical package, `MATLAB`, `URDME`, and `MCell`.

Mathematical topics to be covered include applications of:

- Discrete and continuous probability spaces. Expectation, independence, conditional probability. Bayes' theorem.
- Numerical techniques for generating samples from different probability distributions.
- Random walks: Markov processes in discrete and continuous time with discrete and continuous space variables.
- Diffusion processes: Markov processes in continuous time and space obtained as the limit of a random walk; Wiener and Ornstein-Uhlenbeck processes.
- Point processes: Poisson, inhomogeneous Poisson, Markov chains on graphs.
- Numerical methods for generating each type of process, include Gillespie's algorithm for exact stochastic simulation of coupled chemical reactions, and extensions (Gibson/Bruck, Random Time Change representation, time-varying propensities).
- Statistical analysis of time series; Power spectra of random processes.
- Kurtz's theorem: a continuous time Markov model for a set of chemical reactions converges to the right ODE in the limit of a large, well-mixed volume for finite times.
- "Keizer's paradox" and the correspondence of master equation and mass action chemical kinetics models. Concentration robustness and deficiency of chemical reaction models.

Biological applications will be determined by the mutual interests of the students and the instructor. Suggestions are welcomed. A list of tentative topics includes:

- Stochastic membrane ion channel kinetics.
- Simulation of biochemical and genetic regulatory networks.
- Stochastic predatory-prey models.
- Molecular motors and stochastic ratchets.
- Dynamics of actin and tubulin polymerization, actin treadmilling, cell motility, dynamic instability and tubulin "catastrophes".

- Stochastic treatment of Michaelis Menten kinetics.
- Random walk models for neural spike trains.
- Bacterial random walks and chemotaxis.

Course Requirements

The course work will include regular homework and/or in-class assignments (worth 50% of the grade), class preparation and participation (worth 10% of the grade) and a course project (worth 40% of the grade). There will not be a final exam. The projects will be evaluated based on a written report of 15-20 pages (including figures and references; the text should be 10-15 pages). With permission of the instructor, the project may be done individually or in a group of two. Students are required to turn in a detailed project proposal before spring break. Depending on the time available and course enrollment, the project may include a class presentation as well as a written report. Late homework assignments will be given a random score as follows. If the same assignment would have gotten a score of s, had it been turned in on time, the score will instead be $u^{k/5}s$, where u is uniformly distributed on the unit interval $[0, 1)$, and k is the number of weekdays the assignment is late (rounded up to the next largest integer). It is not in your best interest to fall behind on the homework assignments!

Books & Materials

There are two required books (three for students in Math 419) and several software packages that are strongly recommended.

Required Books & Materials

- *Stochastic Modelling for Systems Biology, 2/E* (2012), by Darren J. Wilkinson. Available from Chapman & Hall / CRC Press.
  This book gives an excellent undergraduate-level introduction to probability and stochastic processes in the context of biological modeling. It includes material on representation of chemical reaction systems *via* Petri nets, detailed presentation of Gillespie's exact simulation algorithm and its variants, and Bayesian inference for stochastic biochemical network systems. The book and the author's website `https://www.staff.ncl.ac.uk/d.j.wilkinson/smfsb/2e/index.html` provide example simulations built in the `R` programming language. The book also provides an introduction to `R` and the Systems Biology Markup Language (SBML, `http://sbml.org/`).
- *Stochastic Processes in Cell Biology* (2014), by Paul C. Bressloff. Available from Springer.
  This book gives an advanced undergraduate / beginning graduate level survey of mathematical modeling of stochastic processes in cell biology. It treats diffusion models, stochastic ion channels, molecular motors, signal transduction pathways, stochastic gene expression regulatory networks, and several other topics. It includes a chapter on WKB methods and mathematical background on probability theory (martingales, stopping times, branching processes) at a level appropriate for graduate students in mathematics.
- (Required for Math 419 students only). *Stochastic Analysis of Biochemical Systems* (2015), by David F. Anderson and Thomas G. Kurtz. Available from Springer. This terse (75 page) book provides a

rigorous treatment of stochastic biochemical network models based on a continuous time, discrete state framework.

- We will take advantage of several different numerical simulation platforms during the course. Students should allow time to download and install the following:
  - `Matlab` (a widely used general purpose numerical computing platform, available to CWRU students through site-license from the CWRU software center, `https://softwarecenter.case.edu`).
  - `R` (another widely used general purpose platform, prevalent in the statistics community; available from `https://www.r-project.org`).
  - `StochSS` (stochastic simulation service, a specialized set of software tools for building and simulating chemical reaction network models; allows interconversion between ODE models, well-mixed discrete population models, and spatially-distributed population models in simple geometries; available from `http://www.stochss.org` both as a web-based service and as a set of tools one can download, install and run locally).
  - `MCell/DReAMM` (state-of-the-art Monte Carlo platform for spatially distributed chemical reaction simulation in complex 3D geometries, using `Blender`; available from `http://mcell.org`).
  - `XPP/AUTO` (NOT required for this course; but easy to install and useful as a general purpose tool for dynamical systems, phase plane analysis, bifurcation analysis, averaging, and other things; available from `www.math.pitt.edu/~bard/xpp/xpp.html`)

Recommended Books & Materials

Mathematical biology is a broad interdisciplinary field and there are many places to find useful information. Here are some references that I recommend.

- Pinsky & Karlin, *An Introduction to Stochastic Modeling (4/E)*, Academic Press, 2011. A well-written general introduction to probability and stochastic processes at the advanced undergraduate level.
- Howard Berg, *Random Walks in Biology*, Princeton University Press, 1993. A beautifully written book introducing ideas related to random walks in a semiquantitative setting based on the author's pioneering work on chemotaxis in motile prokaryotes.
- Crispin Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences (4/E)*, Springer, 2009. An indispensable reference for all manner of stochastic modeling, particularly stochastic differential equations, the Itô calculus, and Fokker-Planck equations. Accessible to advanced undergraduates and beginning graduate students.
- Linda Allen, *An Introduction to Stochastic Processes, with Applications to Biology*, Prentice Hall, 2003. A useful reference touching on many topics related to the course.
- Ross, *First Course in Probability*, seventh edition, Pearson Prentice Hall, 2006. A standard textbook for a first course in probability.
- Anderson, Seppäläinen and Valkó, *Introduction to Probability*, Cambridge, 2018. Another introductory textbook for a first course in probability.
- Grimmett & Stirzaker, *Probability and Random Processes*, third edition, Oxford University Press, 2005. A standard upper level undergraduate or beginning graduate text in probability theory.
- Lawler, *Introduction to Stochastic Processes.* A readable upper-level undergraduate or beginning graduate text on topics including Markov chains and Brownian motions.

- Denny & Gaines, *Chance in Biology: Using Probability to Explore Nature.* Princeton University Press, 2000. A very readable introduction to selected applications of probability theory in biology. Inexpensive and frequently illuminating. Accessible to undergraduates from all disciplines.

- Edelstein-Keshet, *Mathematical Models in Biology*, second edition, SIAM Classics in Applied Mathematics, 2005. A classic introduction to mathematical biology. Does not treat stochastic phenomena, but covers all the concepts from multivariable calculus and differential equations that will be needed throughout the course.

- Ellner & Guckenheimer, *Dynamic Models in Biology*, Princeton University Press, 2006. Another book mainly concerning deterministic models, but chapter three has a very approachable treatments of stochastic ion channels.

- Fall, Marland, Wagner & Tyson (Eds.), *Computational Cell Biology*, Springer series on Interdisciplinary Applied Mathematics / Mathematical Biology, 2002. Contains several chapters on stochastic phenomena, including ion channels and molecular motors.