Comps Description

Topics in Stochastic Processes and Probability
Faculty: Bob Dobrow
Prerequisite: Math 265. Math 365 is recommended.
Fall-Winter terms

Stochastic, or random, models are increasingly applied in many scientific disciplines, such as biology, computer science, statistics, and elsewhere. This comps will give students the opportunity to study such models in detail, and apply what they have learned in Math 265 and 365.

The expected number of students in this extended comps is about 9, which would break down into three fairly disjoint groups. The first part of the project will be a Math 395 seminar on stochastic process, which will extend what is covered in Math 365 this term. A focus of the seminar will be on student-directed reading from outside literature. Student presentations of their readings will play a role. I will lecture (somewhat minimally) based on students' interests. Our goal is that by the middle of Fall term, group projects will be proposed and discussed, and students will have picked their projects and their working groups. For the second half of Fall term and throughout the Winter each group will develop their own project independently of other groups. I will meet regularly with each group.

Motivated students who have taken 265 but not 365 will still be able to participate in this comps.

There is a wealth of material available for comps projects in stochastic processes. And there is enough flexibility in this comps so that if students have their own ideas or interests, we will be able to develop a suitable project proposal.

Following are three possible projects.

Stochastic models for the spread of infectious diseases. Discrete and continuous-time Markov chains, as well as branching processes and stochastic differential equations have been used to study the spread of diseases. The SIS and SIR epidemic models have been studied extensively. Here individuals are classified as Susceptible, Infectious, or immune (Removed). Systems of differential equations are used for deterministic models.

There is growing interest in stochastic models, where $S=S(t)$, $I(t)$, and $R(t)$ are random variables, possibly depending on time. In this project students will learn about stochastic epidemic models, study how they have been applied to different diseases and epidemics (e.g., measles, Ebola, avian flu), and perhaps even amass data to use for either applying, simulating and/or testing these models.

There is enough interesting material in this area, so that we may have more than one group on stochastic models in biology.

Cryptography: Random Walks and Markov Chain Monte Carlo. A fascinating article by mathematician Persi Diaconis called *The Markov Chain Monte Carlo Revolution* describes a novel use of random walk to decode text . To decode a simple substitution cipher one is looking for an unknown coding function

f: { code space } → { usual alphabet}.

A standard approach to decryption is to use frequency statistics of written English to guess at probable choices of f and try these out until the decrypted message makes sense. Diaconis describes a randomized Markov chain algorithm which performs a random walk on the space of all possible coding functions, to obtain a candidate f with high probability.

The suggested comps project would focus on extending and improving this algorithm. Diaconis' discussion is based on first-order transitions of consecutive alphabet letters (e.g., in English text, how often does *a* follow *b*? follow *c*? etc.) Can the algorithm be efficiently extended to higher order transitions? Can parameters be introduced and modified to improve the algorithm? For what types of text/codes does the algorithm work well, work poorly? There are many questions and issues to explore.

Random fields and image analysis. Markov random fields are spatial generalizations of Markov chains. A probability model is placed on the space of two-dimensional lattice *configurations,* with probabilities assigned to individual *sites* of the configuration. In the Markov setting, the probability at each site, conditional on the entire configuration, only depends on a small neighborhood of the given site. (In the regular Markov chain model, the probability of a future state, given the present and full past history only depends on the present state.)

One famous example is the Ising model of magnetism in physics.. These models have been used to study magnetism, but also for image analysis, compression, and reconstruction. In this setting, a configuration is a gray-level image, and sites are individual pixels.

For this project, students would learn about Markov random fields and their application to images, and perhaps use these techniques to reconstruct some fuzzy images.