C303 Stochastic Simulation

Credits. 3

Prerequisites. C106, C107, C110

Dependent Courses. None

Attributions. Core; C, T, L

Rationale, Outlook, Purpose, Objectives, and Goals. To familiarize the student with simulation methods (together with their modeling contexts) involving the use of randomness, including methods for sampling from probability distributions, Monte Carlo integration, Markov chain Monte Carlo methods, etc.

Syllabus.

1. Randomness. Randomness in natural processes: decaying nuclei, chaotic oscillators, leaky faucets, cosmic ray showers, etc. Randomness as complexity, non-compressibility of information, unpredictability, ignorance, statistical independence. Randomness as a modeling assumption. Randomness and entropy.
2. Pseudo-Random Number Generators. Generating deterministic sequences of numbers that appear random. Uniform pseudo-random number generators their properties. Breaking correlations via shuffling. Mersenne Twister and other state-of-the-art generators: an overview. Simple transformations from the Uniform. Other distributions as transformations from the Uniform: exponential, Cauchy, Beta, etc. $N(0, 1)$ using Box-Müller and other methods. Arbitrary distributions and acceptance-rejection sampling. Testing for randomness: how random is random enough? `DIEHARD` and other test batteries.
3. (Optional) Correlated Random Numbers. Normal random numbers with pre-specified correlations. Nataf transformation.
4. Monte Carlo Integration. Estimating π using a dartboard. Estimating one-dimensional integrals: basic MC integration. Importance sampling for better estimators and tighter errorbars. Deterministic vs. Stochastic: Behaviour of the error as function of the number of dimensions.
5. Sampling and Integration in More Than One Dimension. Markov chains, their properties, and limit theorems. Metropolis, Metropolis-Hastings and Gibbs sampling. Master equation, detailed balance, and why Metropolis-Hastings works. Relationship between Metropolis-Hastings, Metropolis, and Gibbs. Relationship between Metropolis and rejection sampling. A survey of illustrative problems involving high-dimensional distributions, integration/expectation, and simulations. Practical considerations: the adjustable step length parameter, behaviour of Markov chain Monte Carlo methods when the distribution being sampled is multimodal, burn-in or equilibration behaviour, detecting equilibration/convergence of the Markov chain, convergence diagnostics, correlations and error bars on estimates, etc.
6. (Optional) Specialized (M&)S Methods Involving Randomness. Reaction kinetics, epidemiology, and population dynamics: The Gillespie method. Agent-based stochastic models in epidemiology and other fields. Tutorial on stochastic differential equations. Discrete versus continuous, stochastic versus deterministic: What is more appropriate/useful for given problem?

Suggested Texts/References.

1. Charles M. Grinstead and J. Laurie Snell, *Introduction to Probability.* American Mathematical Society, 1997 or later. https://math.dartmouth.edu/~prob/prob/prob.pdf
2. Hoel, Port, and Stone, *Introduction to Stochastic Processes.* Houghton Mifflin, 1972.
3. George Casella and Edward I. George, *Explaining the Gibbs Sampler*, The American Statistician **46**(3) 167–174 (1992).
4. Brooks, Gelman, Jones, and Meng (eds.), *Handbook of Markov Chain Monte Carlo.* Chapman and Hall/CRC, 2011.
5. Liang, Liu, and Carrol, *Advanced Markov Chain Monte Carlo Methods: Learning from Past Samples.* Wiley, 2010.
6. Gilks, Richardson, and Spiegelhalter, *Markov Chain Monte Carlo Methods in Practice.* Chapman and Hall, 1996.

Notes on Pedagogy. This syllabus is based on an outline for a longer course that was refined over several course deliveries by the contributor (see below). Depending on the background and capabilities of the students, this outline may need to be somewhat diluted or intensified – without compromising upon the essential content and the goals for the course. This course needs sufficient level of hands-on activities, and students require adequate computing skills. What is not mentioned explicitly in the syllabus is the modeling contexts in which stochastic methods can be useful. Exposing the student to such modeling contexts is a must. Specific modeling contexts can be chosen by the instructor according to her/his field of specialization.

Contributor/s. Mihir Arjunwadkar (http://cms.unipune.ac.in/~mihir)