Probability and Stochastic Processes
Lecture 19: Simulation of Random Variables

Prof. Washim Uddin Mondal
Department of Electrical Engineering
Indian Institute of Technology Kanpur

Monte Carlo Simulation

Objective

Given access to an oracle that can generate independent samples from an unknown CDF $F(\cdot)$, how to estimate the CDF $F(\cdot)$?

Solution

- Generate N sample values $x_1, \cdots, x_N$ from the oracle.
- Assign $F(x) = (\text{number of samples} \leq \text{x})/N$.
- Compute $F(x)$ over a suitable range of x.
- Choose an appropriate resolution Δx.

Simulation of a Uniform Random Variable

Fundamentals

- Every modern programming language provides access to an oracle that generates IID samples from a uniform distribution over $(0, 1)$.
- In MATLAB, the oracle is $\mathrm{rand}()$ function.
- The working principle of such an oracle is beyond our scope.

Simulation of $\mathrm{Unif}(a, b)$

- If $U \sim \mathrm{Unif}(0, 1)$, then $X = a + (b - a)U \sim \mathrm{Unif}(a, b)$.

Simulation of Other Random Variables

Proposition 19.1

If $U \sim \text{Unif}(0,1)$, and F is a valid CDF, then $X \sim F$, where

$$X = \inf\{x \in \mathbb{R} \mid U \leq F(x)\}$$

Proof of Proposition 19.1

- It is evident from the definition of X that $\{U \leq F(x)\} \subset \{X \leq x\}$.
- Since F is right-continuous, $U \leq F(X)$ pointwise. Also, $\{X \leq x\} \subset \{F(X) \leq F(x)\}$ because F is non-decreasing. Combining, we get $\{X \leq x\} \subset \{U \leq F(x)\}$.
- Hence, $\{X \leq x\} = \{U \leq F(x)\}$, and $\mathbf{P}(X \leq x) = \mathbf{P}(U \leq F(x)) = F(x)$.

Simulation of Discrete Random Variables

Proposition 19.2

If X is a discrete random variable with F_X as its CDF, then X can be expressed as follows

$$X = \inf\{x \in \mathrm{range}(X) \mid U \leq F_X(x)\}$$

where $U \sim \mathrm{Unif}(0,1)$.

Bernoulli Random Variables

If $U \sim \mathrm{Unif}(0,1)$, and $p \in (0,1)$, then $X \sim \mathrm{Bernoulli}(p)$, where

$$X = \begin{cases} 0 & \text{if } U \leq 1-p \\ 1 & \text{if } U > 1-p \end{cases}$$

Simulation of Discrete Random Variables

Geometric Random Variables

If $X \sim \text{Geom}(p)$, then one can write the following, where $U \sim \text{Unif}(0,1)$.

$$X = \min\{n \in \mathbb{N} \mid U \leq 1 - p^n\} = \min\left\{n \in \mathbb{N} \;\middle|\; n \geq \frac{\log(1-U)}{\log(p)}\right\}$$

Poisson Random Variables

If $X \sim \text{Poisson}(\lambda)$, then we have the following, where $U \sim \text{Unif}(0,1)$.

$$X = \min\left\{n \in \{0, 1, \cdots\} \;\middle|\; U \leq e^{-\lambda} \sum_{k=0}^{n} \frac{\lambda^k}{k!}\right\}$$

Simulation of Discrete Random Variables

Binomial Random Variables

- Simulation via a uniform random variable might be inefficient.
- If $X_i \sim \text{Bernoulli}(p)$, $\forall i \in \{1, \cdots, n\}$ are independent, then

$$X_1 + \cdots + X_n \sim \text{Binomial}(n, p)$$

Negative Binomial Random Variables

- Simulation via a uniform random variable might be inefficient.
- If $X_i \sim \text{Geom}(p)$, $\forall i \in \{1, \cdots, r\}$ are independent, then

$$X_1 + \cdots + X_r \sim \text{NB}(r, p)$$

Simulation of Continuous Random Variables

Proposition 19.3

If X is a continuous random variable with an invertible CDF F_X, then we can express $X = F_X^{-1}(U)$, where $U \sim \mathrm{Unif}(0,1)$.

Exponential Random Variable

- If $X \sim \exp(\lambda)$, then $F_X(x) = \left(1 - e^{-\lambda x}\right) \mathbf{1}(x \geq 0)$.
- If $U \sim \mathrm{Unif}(0,1)$, then we can write

$$X = -\frac{1}{\lambda} \log(1 - U)$$