Overview of Monte Carlo Simulation, Probability Review and Introduction to Matlab

1 Overview of Monte Carlo Simulation

1.1 Why use simulation?

- To understand complex stochastic systems

- To control complex stochastic systems

Such systems are often too complex to be understood and / or controlled using analytic or numerical methods

- Analytical Methods

    - can examine many decision points at once
    - but limited to simple models

- Numerical Methods

    - can handle more complex models but still limited
    - often have to repeat computation for each decision point

- Simulation

    - can handle very complex and realistic systems
    - but has to be repeated for each decision point

Examples

- Telecommunications networks

- Financial markets

- Weather forecasting

- Revenue Management

- Insurance

- Engineering

- Statistics

Before using simulation, we first need to build a mathematical *model* of the system

1.2 Modelling the System

- A good model should
  - facilitate a good understanding of the system, either analytically, numerically, or through simulation
  - capture the salient details of the system, omitting factors that are not relevant
- The steps in modelling
  - identify the system, including the features and state variables that need to be included and those that should be excluded
  - make necessary probabilistic assumptions about state variables and features
  - test these assumptions
  - identify the inputs and initial values of the state variables
  - identify the performance measure that we wish to obtain from the system
  - identify the mathematical relationship between the terminal state variables and the performance measure
  - Solve the model either analytically, numerically or using simulation

Example 1 (An Inventory Problem)

A retailer sells a perishable commodity and each day he places an order for Q units. Each unit that is sold gives a profit of 60 cents and units not sold at the end of the day are discarded at a loss of 40 cents per unit. The demand, D, on any given day is uniformly distributed on $[80, 140]$. How many units should the retailer order to maximize expected profit?

Solution 1: Use Simulation

- Let P denote profit. Then
$$P = \begin{cases} 0.6Q & \text{if } D \geq Q \\ 0.6D - 0.4(Q - D) & \text{if } Q > D \end{cases}$$

- Set $Q = 80$
- Generate n replications of D
- For each replication, compute the profit or loss, P_i
- Estimate $\mathsf{E}[P]$ with $\sum P_i / n$
- Repeat for different values of Q
- Select the value that gives the biggest estimated profit

Solution 2: Solve Analytically!

- Can write expected profit as

$$\mathsf{E}[P] = \int_Q^{140} \frac{0.6Q}{60} \, dx \; + \; \int_{80}^Q \frac{(0.6x - 0.4(Q - x))}{60} \, dx$$

- Use calculus to find the optimal $Q^* = 116$

Analytic solution gives the *exact* answer and requires far less work than simulation! But this problem is simple and for more complex problems, simulation is often the only option. ∎

1.3 The Issues in Simulation

In this course we will generally assume that the model has been developed and that we have decided to use simulation to analyze the model. A number of questions then arise:

- What distribution do the random variables have?

- How do we generate these random variables for the simulation?

- How do we analyze the output of the simulation?

- How many simulation *runs* do we need?

- How do we improve the *efficiency* of the simulation?

We will answer all of these questions in this course. Consider now the following example where we wish to price Asian options.

Example 2 (Pricing Asian Options)

Consider a stock whose price at time t is denoted by S_t. For a given positive integer, m, the arithmetic mean, A, and the geometric mean, G, are given by

$$A := \frac{\sum_{j=1}^{m} S_{\frac{jT}{m}}}{m}$$

$$G := \sqrt[m]{\prod_{j=1}^{m} S_{\frac{jT}{m}}}$$

Now consider an asset that pays

$$[A - K]^+ := \max(A - K, 0)$$

at time T. This is a *European call option* on the arithmetic mean. K is the *strike price* and is fixed in advance at $t = 0$. T is the *expiration date*.

Similarly, a call option on the geometric mean has payoff at time T given by

$$[G - K]^+ := \max(G - K, 0).$$

Finance theory implies that the fair price, C_A, of the arithmetic option is given by

$$C_A := \mathsf{E}^Q[e^{-rT}(A - K)^+] \tag{1}$$

where the expectation in (1) is taken with respect to the *risk-neutral* probability distribution, Q, and r is the risk-free interest rate.
Similarly, the price, C_G, of the geometric option is given by

$$C_G := \mathsf{E}^Q[e^{-rT}(G - K)^+].$$

We can often compute C_G in closed form but this is generally not true of C_A. So in order to find C_A we could use simulation. How do we do this?

Solution

1. Generate random values of $S_{\frac{jT}{m}}$ for $j = 1, \ldots, m$

2. Compute $a_1 := \left[\frac{1}{m} \sum_{j=1}^{m} S_{\frac{jT}{m}} - K\right]^+$

3. Repeat n times

4. Estimate given by

$$\widehat{C}_A = e^{-rT} \frac{(a_1 + \ldots + a_n)}{n}$$

Questions

- How do we generate sample values of $S_{\frac{jT}{m}}$?

- How large should n be?

- How do we assess the quality of $\widehat{C}_A$?

$$\mathsf{E}[\widehat{C}_A] = ?$$

$$\mathsf{Var}(\widehat{C}_A) = ?$$

- How do we make the simulation more efficient?

Note that we have assumed that we know the model in this example:

- The state variables have been identified

- The goal is to price the option

- The mathematical relationship between the inputs and outputs have been identified

1.4 Variance Reduction Techniques

A particularly important topic is *simulation efficiency*. In particular we will wish to develop techniques that will improve the efficiency of our simulations and this leads us to *variance reduction* techniques:

1. Common random numbers

2. Antithetic variates

3. Control variates

4. Conditional expectations

5. Stratified sampling

6. Importance sampling

We now give a very brief introduction to and motivation for variance reduction methods.

Antithetic Variates

- S_t often assumed to be lognormally distributed i.e. $\log(S_t) \sim \mathsf{N}(\mu, \sigma^2)$

- Do we need to generate 2 independent $\mathsf{N}(\mu, \sigma^2)$ variables in order to generate 2 sample values of S_t?

- Maybe not: if $X \sim \mathsf{N}(0,1)$ then $-X \sim \mathsf{N}(0,1)$

- How can this help?

  - based upon

$$\mathsf{Var}(W + Y) = \mathsf{Var}(W) + \mathsf{Var}(Y) + 2\mathsf{Cov}(W, Y)$$

Control Variates

- Based on the idea of 'using what you know'

- What do we know in the Asian option example?

    - the price of the call option on the *geometric* mean!

- How can this be used? Maybe

$$\widehat{\widehat{C}}_A = \widehat{C}_A + b(C_G - \widehat{C}_G)$$

- Why might this work?

- How do we choose b?

Conditional Monte Carlo and Stratified Monte Carlo

- A very useful formula

$$\mathsf{Var}(X) = \mathsf{Var}(\mathsf{E}[X|Y]) + \mathsf{E}[\mathsf{Var}(X|Y)]$$

- Useful for certain variance reduction techniques

    - conditional Monte Carlo simulation
    - stratified Monte Carlo simulation

- Why? We will see later in the course.

Importance Sampling

- Sometimes it is good to sample from a different distribution!

- Let X be a random variable with density $f(x)$

- Suppose we want to estimate $\theta = \mathsf{E}[h(X)]$

- Let $g(x)$ be another density function

- Then

$$
\begin{aligned}
\theta &= \mathsf{E}[h(X)] \\[2mm]
&= \int h(x)f(x) \ dx \\[2mm]
&= \int \frac{h(x)f(x)g(x)}{g(x)} \ dx \\[2mm]
&= \mathsf{E}_g\left[\frac{h(x)f(x)}{g(x)}\right]
\end{aligned}
$$

- Often very useful for simulating rare events

2 Probability Review

We now review some probability theory. Monte Carlo simulation relies extensively on probability theory so it is very important that you understand this material well.

2.1 <u>Discrete Probability Distributions</u>

Example 3 (Binomial Distribution)

- Say X has a Binomial distribution, or $X \sim \mathsf{B}(n, p)$, if

$$\mathsf{P}(X = r) = \left(\begin{array}{c} n \\ r \end{array} \right) p^r (1-p)^{n-r}$$

- For example, X could be the number of heads in n independent coin tosses, where $p = \mathsf{P}(\text{head})$

- $\mathsf{E}[X] = np$

- $\mathsf{Var}[X] = np(1-p)$

- Assuming we know how to generate random coin tosses, how would we generate a $\mathsf{B}(n, p)$ random variable?

▌

Example 4 (Geometric Distribution)

- Say X has a Geometric distribution if

$$\mathsf{P}(X = r) = (1-p)^{r-1} p$$

- For example, X could be the number of independent coin tosses required until 1st head where $p = \mathsf{P}(\text{head})$

- $\mathsf{E}[X] = 1/p$

- $\mathsf{Var}[X] = (1-p)/p^2$

- Assuming we know how to generate random coin tosses, how would we generate a geometric random variable?

▌

Example 5 (Poisson Distribution)

- Say X has a Poisson(λ) distribution if

$$\mathsf{P}(X = r) = \frac{\lambda^r e^{-\lambda}}{r!}$$

- $\mathsf{E}[X] = \lambda$

- $\mathsf{Var}[X] = \lambda$

- How do we generate a Poisson(λ) random variable?

   − could we use coin tosses?

▌

2.2 Continuous Probability Distributions

We now review some *continuous* distributions, beginning with the most important distribution in simulation, the uniform distribution. But first, we review some notation and important facts regarding continuous random variables. This also applies to discrete random variables but with integrals replaced by summations.

Definition 1 *The* **cumulative distribution function** *(CDF), $F(x)$, of a rv, X, is defined by*

$$F(x) := P(X \leq x).$$

Definition 2 *A continuous rv, X, has* **probability density function** *(PDF), $f(x)$, if $f(x) \geq 0$ and $\forall x$*

$$P(X \in A) = \int_A f(y) \ dy.$$

The CDF and PDF are related by

$$F(x) = \int_{-\infty}^{x} f(y) \ dy$$

so that if you differentiate the CDF you obtain the pdf. It is often convenient to observe that

$$P\left(X \in (x - \frac{\epsilon}{2}, x + \frac{\epsilon}{2}) \right) \approx \epsilon f(x)$$

Definition 3 *For a continuous random vector (X, Y) the* **joint pdf**, $f(x, y)$, *is non-negative and satisfies*

$$P((X, Y) \in A) = \int \int_A f(x, y) \ dx \ dy.$$

Definition 4 *The* **marginal PDF** *of Y is given by*

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) \ dx$$

Definition 5 *The* **conditional PDF** *of X given Y is given by*

$$f_{x|y}(x|y) = \frac{f(x, y)}{f_Y(y)}$$

Recall that if X and Y are independent then $f(x, y) = f_X(x) f_Y(y)$.

Example 6 (**Uniform Distribution**)

- Say X has the uniform distribution on (a, b) if

$$f(x) = \begin{cases} 1/(b-a) & x \in (a, b) \\ \\ 0 & \text{otherwise} \end{cases}$$

  where $f(x)$ is the pdf of X
- Write $X \sim U(a, b)$
- $E[X] = (a + b)/2$
- $Var[X] = (b - a)^2/12$

- Suppose for now we can **simulate** (i.e. generate) independent and identically (**IID**) $U(0,1)$ random variables (**rv's**). How do we generate a random coin toss with $P(\text{Head}) = p$?

Example 7 (Normal Distribution)

- Say X has a Normal distribution, or $X \sim N(\mu, \sigma^2)$, if

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

- $E[X] = \mu$
- $Var[X] = \sigma^2$
- How do we generate a normal random variable?

Example 8 (Exponential Distribution)

- Say X has an Exponential(λ) distribution if

$$f(x) = \lambda e^{-\lambda x}, \qquad x > 0$$

- $E[X] = 1/\lambda$
- $Var[X] = 1/\lambda^2$
- How do we generate an exponential random variable?

Example 9 (Lognormal Distribution)

- Say X has a Lognormal distribution, or $X \sim LN(\mu, \sigma^2)$, if

$$\log(X) \sim N(\mu, \sigma^2)$$

- $E[X] = \exp(\mu + \sigma^2/2)$
  - not μ!
- $Var[X] = \exp(2\mu + \sigma^2)\left(\exp(\sigma^2) - 1\right)$
  - not σ^2!
- How do we generate a lognormal random variable?
- Very important in financial applications

Example 10 (Gamma Distribution)

- Say X has a Gamma(n, λ) distribution if

$$f(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{n-1}}{(n-1)!}, \qquad x > 0$$

- $E[X] = n/\lambda$
- $Var[X] = n/\lambda^2$
- How do we generate a gamma random variable?

2.3 Covariance

- The covariance of X and Y is defined by

$$\mathsf{Cov}(X, Y) := \mathsf{E}\left[(X - \mu_x)(Y - \mu_y)\right]$$

- Then $\mathsf{Var}(aX + bY) = a^2\mathsf{Var}(X) + b^2\mathsf{var}(Y) + 2ab\mathsf{Cov}(X, Y)$

- If X and Y independent, then $\mathsf{Cov}(X, Y) = 0$

- If $X_1, \ldots, X_n$ are IID, then

$$\mathsf{Var}\left(\frac{X_1 + \ldots + X_n}{n}\right) = \frac{\mathsf{Var}(X_1)}{n}$$

- Note $\mathsf{Cov}(X, Y) = 0 \nRightarrow X, Y$ independent

  - unless, for example, the joint distribution of (X, Y) is bivariate normal

2.4 Conditional Expectations and Conditional Variances

For random variables, X and Y, we have the conditional expectation identity

$$\mathsf{E}[X] = \mathsf{E}\left[\mathsf{E}[X|Y]\right]$$

and the conditional variance identity

$$\mathsf{Var}(X) = \mathsf{Var}(\mathsf{E}[X|Y]) + \mathsf{E}[\mathsf{Var}(X|Y)].$$

These identities, while useful in general, are of particular useful in simulation.

Example 11 (A Random Sum of Random Variables)

Let $W = X_1 + X_2 + \ldots + X_n$ where the X_i's are IID and n is also a rv, independent of the X_i's.

1. What is $\mathsf{E}[W]$?
   - Use $\mathsf{E}[X] = \mathsf{E}\left[\mathsf{E}[X|Y]\right]$
   - What is X and what is Y ?

2. What is $\mathsf{Var}(W)$?
   - Use $\mathsf{Var}(X) = \mathsf{Var}(\mathsf{E}[X|Y]) + \mathsf{E}[\mathsf{Var}(X|Y)]$
   - Again, what is X and what is Y ?

∎

Example 12 (Chickens and Eggs)

A hen lays N eggs where $N \sim \mathsf{Poisson}(\lambda)$. Each egg hatches with probability p, independently of the other eggs and N. Let K be the number of chickens.

1. What is $\mathsf{E}[K|N]$?

2. What is $\mathsf{E}[K]$?

3. What is $\mathsf{E}[N|K]$?

∎

2.5 Laws of Large Numbers

- Suppose $X_1, \ldots, X_n$ IID with mean μ. Then for any $\epsilon > 0$

$$\mathsf{P}\left(\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| > \epsilon\right) \to 0 \qquad \text{as } n \to \infty$$

  - this is *the Weak Law of Large Numbers*
  - see Ross for proof under assumption of finite variances

- The *Strong Law of Large Numbers* says

$$\lim_{n \to \infty} \frac{X_1 + \ldots + X_n}{n} = \mu$$

  with probability 1. This law will justify our simulation methodology.

2.6 Central Limit Theorem

- Suppose $X_1, \ldots, X_n$ IID with mean μ and finite variances σ^2

- Define

$$S_n := \frac{X_1 + \ldots + X_n}{n}$$

- Then

$$\frac{S_n - \mu}{\sigma/\sqrt{n}} \to \mathsf{N}(0, 1)$$

- Very useful for analyzing simulation output

2.7 The Poisson Process

- Consider a stochastic process where events (or arrivals) occur at random points in time

- Let $N(t)$ denote the number of events (or arrivals) that have occurred by time t

- Assume $N(t)$ has the following properties

  1. $N(0) = 0$
  2. number of arrivals in non-overlapping time intervals are independent
  3. distribution of number of arrivals in a time interval only depends on length of interval
  4. for $h > 0$ very small,

$$
\begin{aligned}
\mathsf{P}\left(N(h) = 1\right) &\approx \lambda h \\
\mathsf{P}\left(N(h) \geq 2\right) &\approx 0 \\
\mathsf{P}\left(N(h) = 0\right) &\approx 1 - \lambda h
\end{aligned}
$$

- Then say $N(t)$ is a Poisson process with *intensity* λ

- Very useful for modelling 'arrival' processes. Good for modelling

  - emission of particles from a radioactive source
  - market crashes or other 'big' events
  - arrival of customers to a queue

- It can be shown that

$$P\left(N(t) = r\right) = \frac{(\lambda t)^r e^{-\lambda t}}{r!}$$

  - i.e. $N(t)$ is Poisson(λt)

- Let X_i be interval between the $(i-1)^{th}$ and i^{th} arrivals

  - also called the ith interarrival time

- Assumptions imply X_i's are IID

- What is the distribution of X_i?

- $P(X_1 > t) = P\left(N(t) = 0\right) = \ldots?$

Later in the course we will review Brownian motion and any other topics from probability theory that we might need.

Example 13 (A Teaser)

The king comes from a family of 2 children. What is the probability that the other child is his brother?

Notation:

BG means 1st child a boy, second child a girl
BB means both children are boys etc.

What is it we want to compute?

$\blacksquare$

---

3 Introduction to Matlab

Matlab should be available in the computing labs at Columbia. If you wish to purchase it for yourself, then you should purchase the student edition. Make sure you also obtain the *Statistics* toolbox. The *Optimization* and *Financial* toolboxes are often useful as well but we will probably not need them in this course. The easiest way to learn Matlab is by working through examples and you should download one of the many Matlab tutorials that can be found online. They are relatively short and working through them is probably the best way to start.

- Some online tutorials can be found at

  http://www.eece.maine.edu/mm/matweb.html

- A particularly good one by Kermit Sigmon can be found at

  http://ise.stanford.edu/Matlab/matlab-primer.pdf

- We will use Matlab extensively in this course

  - in lectures
  - and in assignments

- But you may use other software if you wish