3. Monte Carlo Simulations
References

1. Chapters 4 and 8, “Numerical Methods in Finance”

2. Chapters 17.6-17.7, “Options, Futures and Other Derivatives”

3. Monte Carlo Simulations

3.1 Overview
Monte Carlo Simulation

- Monte Carlo simulation, a quite different approach from binomial tree, is based on statistical sampling and analyzing the outputs gives the estimate of a quantity of interest.
Monte Carlo Simulation

• Typically, estimate an expected value with respect to an underlying probability distribution
  – *eg.* an option price may be evaluated by computing the expected payoff w.r.t. risk-neutral probability measure
• Evaluate a portfolio policy by simulating a large number of scenarios
• Interested in not only mean values, but also what happens on the tail of probability distribution, such as Value-at-risk
Monte Carlo Simulation in Option Pricing

• In option pricing, Monte Carlo simulations uses the risk-neutral valuation result
• More specifically, sample the paths to obtain the expected payoff in a risk-neutral world and then discount this payoff at the risk-neutral rate
Learning Objectives

- Understanding the stochastic processes of the underlying asset price.
- How to generate random samplings with a given probability distribution
- Choose an optimal number of simulation experiments
- Reliability of the estimate, e.g. a confidence interval
- Estimate error
- Improve quality of the estimates: variance reduction techniques, quasi-Monte Carlo simulations.
3. Monte Carlo Simulation

3.2 Modeling Asset Price Movement
Randomness in stock market

Daily IBM prices from January 2001 to January 2005
Outlines

1. Markov process
2. Wiener process
3. Generalized Wiener process
4. Ito’s process, Ito’s lemma
5. Asset price models
References

1. Chapter 12, “Options, Futures, and Other Derivatives”
A particular type of stochastic process whose future probability depends only on the present value.

A stochastic process $x(t)$ is called a Markov process if for every $n$ and $t_1 < t_2 < \cdots < t_n$, we have

$$P(x(t_n) \leq x_n \mid x(t) \text{ for all } t \leq t_{n-1}) = P(x(t_n) \leq x_n \mid x(t_{n-1}))$$

The Markov property implies that the probability distribution of the price at any particular future time is not dependent on the particular path followed by the price in the past.
Modeling of asset prices

• Really about modeling the arrival of new info that affects the price.
• Under the assumptions,
  ▪ The past history is fully reflected in the present asset price, which does not hold any further information
  ▪ Markets respond immediately to any new information about an asset
• The anticipated changes in the asset price are a Markov process
Weak-Form Market Efficiency

• This asserts that it is impossible to produce consistently superior returns with a trading rule based on the past history of stock prices. In other words technical analysis does not work.

• A Markov process for stock prices is clearly consistent with weak-form market efficiency
Example of a Discrete Time Continuous Variable Model

- A stock price is currently at $40.
- Over one year, the change in stock price has a distribution $\mathcal{N}(0,10)$ where $\mathcal{N}(\mu,\sigma)$ is a normal distribution with mean $\mu$ and standard deviation $\sigma$. 
Questions

• What is the probability distribution of the change of the stock price over 2 years?
  – ½ years?
  – ¼ years?
  – $\Delta t$ years?

• Taking limits we have defined a continuous variable, continuous time process
Variances & Standard Deviations

• In Markov processes, changes of the variable in successive periods of time are independent. This means that:
  – Means are additive
  – Variances are additive
  – Standard deviations are not additive
Variances & Standard Deviations
(continued)

- In our example it is correct to say that the variance is 100 per year.
- It is strictly speaking not correct to say that the standard deviation is 10 per year.
A Wiener Process (See pages 265-67, Hull)

- Consider a variable $z$ follows a particular Markov process with a mean change of 0 and a variance rate of 1.0 per year.
- The change in a small interval of time $\Delta t$ is $\Delta z$
- The variable follows a Wiener process if

1. $\Delta z = \varepsilon \sqrt{\Delta t}$ where $\varepsilon$ is $N(0,1)$

2. The values of $\Delta z$ for any 2 different (non-overlapping) periods of time are independent
Properties of a Wiener Process

\[ \Delta z \] over a small time interval \( \Delta t \):
- Mean of \( \Delta z \) is 0
- Variance of \( \Delta z \) is \( \Delta t \)
- Standard deviation of \( \Delta z \) is \( \sqrt{\Delta t} \)

\[ \Delta z \] over a long period of time, \( T \):
- Mean of \( [z(T) - z(0)] \) is 0
- Variance of \( [z(T) - z(0)] \) is \( T \)
- Standard deviation of \( [z(T) - z(0)] \) is \( \sqrt{T} \)
Taking Limits . . .

• What does an expression involving $dz$ and $dt$ mean?
• It should be interpreted as meaning that the corresponding expression involving $\Delta z$ and $\Delta t$ is true in the limit as $\Delta t$ tends to zero.
• In this respect, stochastic calculus is analogous to ordinary calculus.
Generalized Wiener Processes
(See page 267-69, Hull)

• A Wiener process has a drift rate (i.e. average change per unit time) of 0 and a variance rate of 1

• In a generalized Wiener process the drift rate and the variance rate can be set equal to any chosen constants

• The variable \( x \) follows a generalized Wiener process with a drift rate of \( a \) and a variance rate of \( b^2 \) if

\[
dx = a \, dt + b \, dz
\]
Generalized Wiener Processes
(continued)

\[ \Delta x = a \Delta t + b \varepsilon \sqrt{\Delta t} \]

- Mean change in $x$ in time $\Delta t$ is $a \Delta t$
- Variance of change in $x$ in time $\Delta t$ is $b^2 \Delta t$
- Standard deviation of change in $x$ in time $\Delta t$ is $b \sqrt{\Delta t}$
Generalized Wiener Processes
(continued)

\[ \Delta x = a \, T + b \, \varepsilon \sqrt{T} \]

- Mean change in \( x \) in time \( T \) is \( aT \)
- Variance of change in \( x \) in time \( T \) is \( b^2 T \)
- Standard deviation of change in \( x \) in time \( T \) is \( b \sqrt{T} \)
Example

- A stock price starts at 40 and at the end of one year, it has a probability distribution of N(40,10)
- If we assume the stochastic process is Markov with no drift then the process is
  \[ dS = 10dz \]
- If the stock price were expected to grow by $8 on average during the year, so that the year-end distribution is N (48,10), the process would be
  \[ dS = 8dt + 10dz \]
- What the probability distribution of the stock price at the end of six months?
• In an Itô process the drift rate and the variance rate are functions of time

\[ dx = a(x,t) \, dt + b(x,t) \, dz \]

• The discrete time version, i.e., \( \Delta x \) over a small time interval \( (t, t+\Delta t) \)

\[ \Delta x = a(x,t) \Delta t + b(x,t) \varepsilon \sqrt{\Delta t} \]

is only true in the limit as \( \Delta t \) tends to zero
Why a Generalized Wiener Process is not Appropriate for Stocks

• For a stock price we can conjecture that its expected percentage change in a short period of time remains constant, not its expected absolute change in a short period of time

• We can also conjecture that our uncertainty as to the size of future stock price movements is proportional to the level of the stock price
A Simple Model for Stock Prices
(See pages 269-71, Hull)

\[ dS = \mu S \, dt + \sigma S \, dz \]

where \( \mu \) is the expected return \( \sigma \) is the volatility.
The discrete time equivalent is

\[ \Delta S = \mu S \Delta t + \sigma S \varepsilon \sqrt{\Delta t} \]
Example: Simulate asset prices

Consider a stock that pays no-dividends, has an expected return 15% per annum with continuous compounding and a volatility of 30% per annum. Observe today’s price $1 per share and with \( \Delta t = 1/250 \) yr. Simulate a path for the stock price.

Given \( S_{i-1} \), we compute

\[
\Delta S = \mu S_{i-1} \Delta t + \sigma S_{i-1} \epsilon \sqrt{\Delta t}
\]

Then

\[
S_i = S_{i-1} + \Delta S
\]
Example: Simulate stock prices

<table>
<thead>
<tr>
<th>Stock Price S at start of period</th>
<th>Random sample for $\varepsilon$</th>
<th>$\Delta S$ during a time period $dt$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0000</td>
<td>-1.3457</td>
<td>-0.0249</td>
</tr>
<tr>
<td>0.9751</td>
<td>0.6132</td>
<td>0.0119</td>
</tr>
<tr>
<td>0.9870</td>
<td>1.1928</td>
<td>0.0229</td>
</tr>
<tr>
<td>1.0099</td>
<td>0.0114</td>
<td>0.0008</td>
</tr>
<tr>
<td>1.0108</td>
<td>0.2667</td>
<td>0.0057</td>
</tr>
<tr>
<td>1.0165</td>
<td>-0.1590</td>
<td>-0.0025</td>
</tr>
<tr>
<td>1.0140</td>
<td>-0.5652</td>
<td>-0.0103</td>
</tr>
<tr>
<td>1.0037</td>
<td>0.7189</td>
<td>0.0143</td>
</tr>
<tr>
<td>1.0180</td>
<td>-1.3431</td>
<td>-0.0253</td>
</tr>
<tr>
<td>0.9927</td>
<td>-1.3913</td>
<td>-0.0256</td>
</tr>
</tbody>
</table>
Example: Simulate stock prices
Itô’s Lemma (See pages 273-274, Hull)

- If we know the stochastic process followed by $x$, Itô’s lemma tells us the stochastic process followed by some function $G(x, t)$
- Since a derivative security is a function of the price of the underlying and time, Itô’s lemma plays an important part in the analysis of derivative securities
Taylor Series Expansion

A Taylor’s series expansion of $G(x, t)$ gives

$$
\Delta G = \frac{\partial G}{\partial x} \Delta x + \frac{\partial G}{\partial t} \Delta t + \frac{1}{2} \frac{\partial^2 G}{\partial x^2} \Delta x^2
$$

$$
+ \frac{\partial^2 G}{\partial x \partial t} \Delta x \Delta t + \frac{1}{2} \frac{\partial^2 G}{\partial t^2} \Delta t^2 + \ldots
$$
Ignoring Terms of Higher Order Than $\Delta t$ and $\Delta x$

In ordinary calculus we have

$$\Delta G = \frac{\partial G}{\partial x} \Delta x + \frac{\partial G}{\partial t} \Delta t$$

In stochastic calculus this becomes

$$\Delta G = \frac{\partial G}{\partial x} \Delta x + \frac{\partial G}{\partial t} \Delta t + \frac{1}{2} \frac{\partial^2 G}{\partial x^2} \Delta x^2$$

because $\Delta x$ has a component which is of order $\sqrt{\Delta t}$
Substituting for $Δx$

Suppose

$$dx = a(x, t)dt + b(x, t)dz$$

so that

$$Δx = a \Delta t + b \varepsilon \sqrt{Δt}$$

Then ignoring terms of higher order than $Δt$

$$ΔG = \frac{∂G}{∂x} Δx + \frac{∂G}{∂t} Δt + \frac{1}{2} \frac{∂^2 G}{∂x^2} b^2 \varepsilon^2 Δt$$
The $\varepsilon^2\Delta t$ Term

Since $\varepsilon \approx N(0,1)$, $E(\varepsilon) = 0$

$$E(\varepsilon^2) - [E(\varepsilon)]^2 = 1$$

$E(\varepsilon^2) = 1$

It follows that $E(\varepsilon^2\Delta t) = \Delta t$

The variance of $\Delta t$ is proportional to $\Delta t^2$ and can be ignored. Hence

$$\Delta G = \frac{\partial G}{\partial x} \Delta x + \frac{\partial G}{\partial t} \Delta t + \frac{1}{2} \frac{\partial^2 G}{\partial x^2} b^2 \Delta t$$
Taking Limits

Taking limits

\[ dG = \frac{\partial G}{\partial x} dx + \frac{\partial G}{\partial t} dt + \frac{1}{2} \frac{\partial^2 G}{\partial x^2} b^2 dt \]

Substituting

\[ dx = a \ dt + b \ dz \]

We obtain

\[ dG = \left( \frac{\partial G}{\partial x} a + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial x^2} b^2 \right) dt + \frac{\partial G}{\partial x} b \ dz \]

This is Ito's Lemma
The stock price process is

\[ dS = \mu S \, dt + \sigma S \, dz \]

For a function \( G \) of \( S \) and \( t \),

\[
dG = \left( \frac{\partial G}{\partial S} \mu S + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial G}{\partial S} \sigma S \, dz
\]
Example

If \( G = \ln S \), then

\[
dG = ?
\]

\[
dG = \left( \mu - \frac{\sigma^2}{2} \right) dt + \sigma \ dz
\]

Generalized Wiener process
A generalized Wiener process for stock price

The discrete version of this is

\[ \ln S(t + \Delta t) - \ln S(t) = \left( \mu - \sigma^2 / 2 \right) \Delta t + \sigma \varepsilon \sqrt{\Delta t} \]

or

\[ S(t + \Delta t) = S(t) e^{\left( \mu - \sigma^2 / 2 \right) \Delta t + \sigma \varepsilon \sqrt{\Delta t}} \]

\( S(t) \) follows a lognormal distribution. It is often referred to as geometric Brownian motion.
The corresponding density function for $S(t)$ is

$$f(x) = \frac{\exp\left(-\left(\log\left(\frac{x}{S_0}\right) - \left(\mu - \frac{\sigma^2}{2}\right)t\right)^2 \right)}{2\pi t \sigma x \sqrt{2\pi} t}$$

for $x > 0$ with $f(x) = 0$ for $x \leq 0$.

It can be proven that

$$E\left(S(t)\right) = S_0 e^{\mu t}$$

$$E\left(S(t)^2\right) = S_0^2 e^{(2\mu + \sigma^2)t}$$

$$\text{var}\left(S(t)\right) = S_0^2 e^{2\mu t} \left(e^{\sigma^2 t} - 1\right)$$
Example: Simulate stock prices
Models for stock price in risk neutral world (r: risk-free interest rate)

In a risk neutral world, two process for the underlying asset is:

\[
dS = \sigma S \, dz + \hat{\mu} S \, dt
\]

or

\[
S(t + \Delta t) - S(t) = \hat{\mu} S(t) \, \Delta t + \sigma S(t) \, \varepsilon \sqrt{\Delta t}
\]

\[
d \ln S = \left( r - \frac{\sigma^2}{2} \right) \, dt + \sigma \, dz
\]

or

\[
\ln S(t + \Delta t) - \ln S(t) = \left( r - \frac{\sigma^2}{2} \right) \, \Delta t + \sigma \varepsilon \, \sqrt{\Delta t}
\]
Which one is more accurate?

- It is usually more accurate to work with InS because it follows a generalized Wiener process and (2) is true for any $\Delta t$ while (1) is true only in the limit as $\Delta t$ tends to zero.

For all $T$,

$$\ln S(T) - \ln S(0) = \left( r - \frac{\sigma^2}{2} \right) T + \sigma \varepsilon \sqrt{T}$$

i.e.,

$$S(T) = S(0)e^{\left( r - \frac{\sigma^2}{2} \right) T + \sigma \varepsilon \sqrt{T}}$$
3. Monte Carlo Simulation

3.3 Generate pseudo-random variates
Generate pseudorandom variates

The usual way to generate pseudorandom variates, i.e., samples from a given probability distribution:

1. Generate pseudorandom numbers, which are variates from the uniform distribution on the interval \((0, 1)\). Denote as \(U(0, 1)\)

2. Apply suitable transformation to obtain the desired distribution
Generate U(0, 1) variables

The standard method is linear congruential generators (LCGs):

1. Given an integer number \( Z_{i-1} \),
   \[
   Z_i = (aZ_{i-1} + c) \pmod{m}
   \]
   where \( a, c, \) and \( m \) are chosen parameters

2. Return number \( \frac{Z_i}{m} \) which generate a U(0, 1) variable

Comments:
- \( Z_0 \): the initial number, the seed of the random sequence.
  - eg. \texttt{rand('seed', 0)} generates the same sequence each time
- LCGs generates rational numbers instead of real ones
- Periodic. The sequency \( \left\{ U_i = \frac{i}{m} \right\}_{i=0}^{m-1} \) has a maximum period of \( m \)
- Improved MATLAB function \texttt{rand('state', 0)} allows much longer periods
Generate $U(0, 1)$ variables
Generate U(0, 1) variables
Generate general distribution $F(x)$ from $U(0, 1)$

Given the distribution function $F(x) = P\{X \leq x\}$, we can generate random variates according to $F$ by the following

1. Draw a random number $U \sim U(0,1)$
2. Return $X = F^{-1}(U)$

**Proof**: Use the monotonicity of $F$ and the fact that $U$ is uniformly distributed:

$$P\{X \leq x\} = P\{F^{-1}(U) \leq x\} = P\{U \leq F(x)\} = F(x)$$
• How to generate samples according to standard normal distribution $N(0,1)$ from $U(0, 1)$?

- No analytical form of the inverse of the distribution function
- Support is not finite
Generate Samples from Normal Distribution

One simple way to obtain a sample from $\mathcal{N}(0,1)$ is as follows

$$x = \sum_{i=1}^{12} U_i - 6$$

where $U_i$ are independent random numbers from $U(0,1)$, and $x$ is the required samples from $\mathcal{N}(0,1)$

Note:
- It is from Central Limit Theorem.
- The approximation is satisfactory for most purposes
Central Limit Theorem

Let $X_1, X_2, X_3, \ldots$ be a sequence of independent random variables which has the same probability distribution.

Consider the sum $S_n = X_1 + \ldots + X_n$. Then the expected value of $S_n$ is $n \mu$ and its standard deviation is $\sigma \cdot n^{1/2}$. Furthermore, informally speaking, the distribution of $S_n$ approaches the normal distribution $N(n \mu, \sigma n^{1/2})$ as $n$ approaches $\infty$.

Or the standardized $S_n$, i.e., $Z_n = \frac{S_n - n\mu}{\sigma \sqrt{n}}$.

Then the distribution of $Z_n$ converges towards the standard normal distribution $N(0,1)$ as $n$ approaches $\infty$. 
Uniform random variable X:

PDF:

\[ f(x) = \begin{cases} 
\frac{1}{b-a}, & \text{if } x \in (a,b) \\
0, & \text{otherwise}
\end{cases} \]

Mean:

\[ E[X] = \frac{b+a}{2} \]

Variance:

\[ \text{Var}(X) = \frac{(b-a)^2}{12} \]
Generate Samples from $N(0,1)$: Box-Muller algorithm

The Box-Muller algorithm can be implemented as follows:
1. Generate two independent uniform samples from $U_1, U_2 \sim U(0,1)$
2. Set $R^2 = -2 \log U_1$ and $\theta = 2\pi U_2$
3. Set $X = R \cos \theta$ and $Y = R \sin \theta$

where $X$ and $Y$ are independent samples from $N(0,1)$.

Note:
- Box and Muller (1958), Ann. Math. Statist., 29:610-611
- Common and more efficient than the previous method.
- But it is costly to evaluate trigonometric functions
Generate Samples from N(0,1): Improved Box-Muller algorithm

The polar rejection (improved Box-Muller) method:

1. Generate two independent uniform samples from \( U_1, U_2 \sim U(0,1) \)

2. Set \( V_1 = 2U_1 - 1, V_2 = 2U_2 - 1, \) and \( S = V_1^2 + V_2^2 \)

3. If \( S > 1 \), return to step 1; otherwise, return \( U_1, U_2 \sim U \)

\[ X = \sqrt{-\frac{2 \ln S}{S}} V_1 \quad \text{and} \quad Y = \sqrt{-\frac{2 \ln S}{S}} V_2 \]

where \( X \) and \( Y \) are required independent samples from \( N(0,1) \).

Note:
How to generate samples according to normal distribution $N(\mu, \sigma)$ from $N(0, 1)$?
To obtain correlated normal samples $\varepsilon_1$ and $\varepsilon_2$, we first
1. generate two independent standard normal variates $x_1, x_2 \sim N(0,1)$
2. and set
   
   $\varepsilon_1 = \mu_1 + x_1$

   $\varepsilon_2 = \mu_2 + \rho x_1 + x_2 \sqrt{1 - \rho^2}$

where $\rho$ is the coefficient of correlation.
To obtain $n$ correlated normal samples $\varepsilon_1, \ldots, \varepsilon_n$, we first

1. generate $n$ independent standard normal variates $x_1, \ldots, x_n \sim N(0, 1)$

2. and set

$$\varepsilon = \mu + U^T X$$

Here, $U^T U = \Sigma = \begin{bmatrix} \rho_{ij} \end{bmatrix}$ is the Cholesky decomposition of the covariance matrix $\Sigma$. 
Simulate the correlated asset prices

To simulate the payoff of a derivative depending on multi-variables, how can we generate samples of each variables?

\[ S_i(t + \Delta t) = S_i(t) e^{(\mu_i - \sigma_i^2/2) \Delta t + \sigma_i \epsilon_i \sqrt{\Delta t}} \]
3. Monte Carlo Simulation

3.4 Choose the number of simulations
How many trials is enough?

\( X_i \): Sequence of independent samples from the same distribution

\[
\bar{X}(M) \equiv \frac{1}{M} \sum_{i=1}^{M} X_i: \text{Sample mean, an unbiased estimator to the quantity of interest } \mu = E[ X_i ]
\]

\[
S^2(M) = \frac{1}{M-1} \sum_{i=1}^{M} \left[ X_i - \bar{X}(M) \right]^2: \text{Sample variance}
\]

\[
E \left[ \left( \bar{X}(M) - \mu \right)^2 \right] = \text{Var} \left[ \bar{X}(M) \right] = \frac{\sigma^2}{M}: \text{the expected value of square error as a way to quantify the quality of our estimator}
\]

Note that \( \sigma^2 \) can be estimated by the sample variance \( S^2(M) \).
Number of Simulations and Accuracy

The number of simulation trials carried out depends on the accuracy required.

M: number of independent trials carried out

μ: expected value of the derivative,

S(M): standard deviation of the discounted payoff given by the trials.
The standard error of the estimate of the option price based on sample size $M$ is:

$$\frac{\sigma}{\sqrt{M}} \approx \sqrt{\frac{S^2(M)}{M}}$$

Clearly as $M$ increases, our estimate is more accurate.
(1-\(\alpha\)) Confidence Interval of Monte Carlo Simulation

In general, the confident interval at level (1- \(\alpha\)) may be computed by

\[
\bar{X}(M) - z_{1-\alpha/2} \sqrt{\frac{S^2(M)}{M}} < \mu < \bar{X}(M) + z_{1-\alpha/2} \sqrt{\frac{S^2(M)}{M}}
\]

where \(z_{1-\alpha/2}\) is the critical number from the standard normal distribution.

**Note:** Uncertainty is inverse proportional to the square root of the number of trials
A 95% confidence interval for the price $V$ of the derivative is therefore given by:

$$
\bar{X}(M) - 1.96 \sqrt{\frac{S^2(M)}{M}} < \mu < \bar{X}(M) + 1.96 \sqrt{\frac{S^2(M)}{M}}
$$
Absolute error

Suppose we want to controlling the absolute error such that, with probability \((1-\alpha)\),

\[
\left| \bar{X}(M) - \mu \right| \leq \beta,
\]

where \(\beta\) is the maximum acceptable tolerance. Then the number of simulations \(M\) must satisfy

\[
z_{1-\alpha/2} \sqrt{S^2(M)/M} \leq \beta.
\]
Relative error

Suppose we want to controlling the relative error such that,
with probability \((1-\alpha)\),
\[
\frac{|\bar{X}(M) - \mu|}{\mu} \leq \gamma,
\]
where \(\gamma\) is the maximum acceptable tolerance.
Then the number of simulations \(M\) must satisfy
\[
\frac{z_{1-\alpha/2} \sqrt{S^2(M)/M}}{\bar{X}(M)} \leq \frac{\gamma}{1+\gamma}.
\]
3. Monte Carlo Simulation

3.5 Applications in Option Pricing
Monte Carlo Simulation

When used to value a derivative dependent on a market variable $S$, this involves the following steps:

1. Simulate 1 path for the stock price in a risk-neutral world
2. Calculate the payoff from the stock option
3. Repeat steps 1 and 2 many times to get many sample payoff
4. Calculate mean payoff
5. Discount mean payoff at risk-free rate to get an estimate of the value of the option
Example: European call

- Consider a European call on a stock with no-dividend payment. \( S(0) = 50 \), \( E = 52 \), \( T = 5 \) months, and annual risk-free interest rate \( r = 10\% \) and a volatility \( \sigma = 40\% \) per annum.

- Black-Schole Price: $5.1911
- MC Price with 1000 simulations: $5.4445
  Confidence Interval [4.8776, 6.0115]
- MC Price with 200,000 simulations: $5.1780
  Confidence Interval [5.1393, 5.2167]
Example: Butterfly Spread

• A butterfly spread consists of buying a European call option with exercise price $K_1$ and another with exercise price $K_3$ ($K_1 < K_3$) and by selling two calls with exercise price $K_2 = (K_1 + K_3)/2$, for the same asset and expiry date.

• Consider a butterfly spread with $S(0) = $50, $K_1 = $55, $K_2 = $60, $K_3 = 65$, $T = 5$ months, and annual risk-free interest rate $r = 10\%$ and a volatility $\sigma = 40\%$ per annum.

• Exact price: $0.6124$

• MC price with 100,000 simulations: $0.6095$

  Confidence Interval $[0.6017, 0.6173]$
Example: Arithmetic Average Asian Option

- Consider pricing an Asian average rate call option with discrete arithmetic averaging. The option payoff is

\[
\max \left\{ \frac{1}{N} \sum_{i=1}^{N} S(t_i) - K, 0 \right\}
\]

where \( t_i = i \Delta t \) and \( \Delta t = \frac{T}{N} \).

- Consider the option on a stock with no-dividend payment. \( S(0) = 50 \), \( K = 50 \), \( T = 5 \) months, and annual risk-free interest rate \( r = 10\% \) and a volatility \( \sigma = 40\% \) per annum.
Determining Greek Letters

For $\Delta$:
1. Make a small change to asset price
2. Carry out the simulation again using the same random number streams
3. Estimate $\Delta$ as the change in the option price divided by the change in the asset price

$$\frac{V_{MC}^* - V_{MC}}{\Delta S}$$

Proceed in a similar manner for other Greek letters
3. Monte Carlo Simulation

3.6 Further Comments
The parameters in the model

• Our analysis so far is useless unless we know the parameters $\mu$ and $\sigma$
• The price of an option or derivative is in general independent of $\mu$
• However, the asset price volatility $\sigma$ is critically important to option price. Typically, $20% < \sigma < 50%$
• *Volatility* of stock price can be defined as the standard deviation of the return provided by the stock in one year when the return is expressed using continuous compounding
Estimate volatility $\sigma$ from historical data

Assume:

$n + 1$: number of observation at fixed time interval  
$S_i$: Stock price at end of ith ($i = 0, 1, ..., n$) interval  
$\Delta t$: Length of observing time interval in years

$$u_i = \ln \left( \frac{S_i}{S_{i-1}} \right) : n \text{ number of approximation for the change of } \ln S \text{ over } \Delta t$$
Estimate volatility $\sigma$ from historical data

Compute the standard deviation of the $u_i$'s:

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (u_i - \bar{u})^2}$$

where the mean of the $u_i$'s is

$$\bar{u} = \frac{1}{n} \sum_{i=1}^{n} u_i$$

$$\therefore s \approx \sigma \sqrt{\Delta t} \quad \therefore \sigma \approx \hat{\sigma} = \frac{s}{\sqrt{\Delta t}}$$

The standard error of estimate is about $\frac{\hat{\sigma}}{\sqrt{2n}}$
Advantages and Limitations

• More efficient for options dependent on multiple underlying stochastic variables. As the number of variables increase, Monte Carlo simulations increases linearly while the others increases exponentially
• Easily deal with path dependent options and options with complex payoffs / complex stochastic processes
• Has an error estimate

• It cannot easily deal with American options
• Very time-consuming
Extensions to multiple variables

When a derivative depends on several underlying variables we can simulate paths for each of them in a risk-neutral world to calculate the values for the derivative.
Sampling Through the Tree

• Instead of sampling from the stochastic process we can sample paths randomly through a binomial or trinomial tree to value a derivative
3. Monte Carlo Simulation

3.7 Variance Reduction Techniques
Variance Reduction Procedures

• Usually, a very large value of $M$ is needed to estimate $V$ with reasonable accuracy.

• Variance reduction techniques lead to dramatic savings in computational time
  • Antithetic variable technique
  • Control variate technique
  • Importance sampling
  • Stratified sampling
  • Moment matching
  • Using quasi-random sequences
Consider to estimate $V = E\left(f(U)\right)$ where $U \sim N(0,1)$.

The standard Monte Carlo estimate:

$$V_{MC} = \frac{1}{M} \sum_{i=1}^{M} f(U_i) \text{ with i.i.d. } U_i \sim N(0,1).$$

The antithetic variate technique:

$$V_{AV} = \frac{1}{M} \sum_{i=1}^{M} \frac{f(U_i) + f(-U_i)}{2} \text{ with i.i.d. } U_i \sim N(0,1)$$

We can prove that

$$\text{var} \left( \frac{f(U_i) + f(-U_i)}{2} \right) = \frac{1}{2} \left[ \text{var} \left( f(U_i) \right) + \text{cov} \left( f(U_i), f(-U_i) \right) \right]$$

$$\leq \frac{1}{2} \text{var} \left( f(U_i) \right)$$

if $f$ is monotonic.
Antithetic Variable Technique

It involves calculating two values of the derivative.

- $V_1$: calculated in the usual way
- $V_2$: calculated by the changing the sign of all the random standard normal samples used for $V_1$

$V$ is average of the $V_1$ and $V_2$, the final estimate of the value of the derivative
Antithetic Variable Technique

Note:
1. Monte-Carlo works when the simulated variables "spread out" as closely as possible to the true distribution.
2. Antithetic variates relies upon finding samplings that are anticorrelated with the original random variable.
3. Works well when the payoff is monotonic w.r.t. $S$
4. Further reading:
   -- N. Madras, Lectures on Monte Carlo Methods, 2002
Example: European call

- Consider a European call on a stock with no-dividend payment. $S(0) = 50$, $K = 52$, $T = 5$ months, and annual risk-free interest rate $r = 10\%$ and a volatility $\sigma = 40\%$ per annum.

- Black-Schole Price: $5.1911$
- MC Price with 200,000 simulations: $5.1780$
  Confidence Interval $[5.1393, 5.2167]$
- MCAV Price with 200,000 simulations: $5.1837$
  Confidence Interval $[5.1615, 5.2058]$ (ratio: $1.75$)
Example: Butterfly Spread

- Consider a butterfly spread with $S(0) = 50$, $K_1 = 55$, $K_2 = 60$, $K_3 = 65$, $T = 5$ months, and annual risk-free interest rate $r = 10\%$ and a volatility $\sigma = 40\%$ per annum.
- Exact price: $0.6124$
- MC price with 100,000 simulations: $0.6095$
  Confidence Interval $[0.6017, 0.6173]$
- MCAV price with 50,000 simulations: $0.6090$
  Confidence Interval $[0.5982, 0.6198]$
Control Variate Technique

We can generalize the previous technique to

$$Z_\theta = V_A + \theta (E(V_B) - V_B)$$

for any $\theta \in \mathbb{R}$. Here, $V_B$ is "close" to $V_A$ with known mean $E(V_B)$.

In this case,

$$\text{var}(Z_\theta) = \text{var}(V_A - \theta V_B)$$

$$= \text{var}(V_A) - 2\theta \text{cov}(V_A, V_B) + \theta^2 \text{var}(V_B)$$

We can prove that $\text{var}(Z_\theta) < \text{var}(V_A)$ if and only if

$$0 < \theta < 2 \frac{\text{cov}(V_A, V_B)}{\text{var}(V_B)}.$$
Example: European call

• Consider a European call on a stock with no-dividend payment. \( S(0) = 50 \), \( K = 52 \), \( T = 5 \) months, and annual risk-free interest rate \( r = 10\% \) and a volatility \( \sigma = 40\% \) per annum.

- Black-Schole Price: $5.1911
- MC Price with 200,000 simulations: $5.1780
  Confidence Interval [5.1393, 5.2167]
- MCCV Price with 200,000 simulations: $5.1883
  Confidence Interval [5.1712, 5.2054] (ratio: 2.2645)
• Consider the option on a stock with no-dividend payment. $S(0) = $50, $K = 50$, $T = 5$ months, and annual risk-free interest rate $r = 10\%$ and a volatility $\sigma = 40\%$ per annum.

• We could use the sum of the asset prices as a control variate as we know its expected value and $Y$ is correlated to the option itself

\[
E\left[\sum_{i=0}^{N} S(t_i)\right] = \sum_{i=0}^{N} E[S(i\Delta t)] = S_0 \sum_{i=0}^{N} e^{ri\Delta t} = S_0 \frac{1-e^{r(N+1)\Delta t}}{1-e^{r\Delta t}}
\]

• Another choice of the control variate is the payoff of a geometric average option as this is known analytically

\[
\max\left\{\left(\prod_{i=1}^{N} S(t_i)\right)^{\frac{1}{N}} - K, 0\right\}
\]
Importance Sampling Technique

It is a way to distort the probability measure in order to sample from critical region. For example, in evaluating European call option:

- $F$: the unconditional probability distribution function for the asset price at time $T$
- $q$: the probability of the asset price $\geq K$ at maturity, known analytically
- $G = F / q$: the probability distribution of the asset conditional on the asset price $\geq K$, i.e., importance function

Instead of sampling from $F$, we sample from $G$. Then the value of the option is the average discounted payoff multiplied by $q$. 
Stratified Sampling Technique

Sampling representative values rather than random values from a probability distribution is usually more accurate.

It involves dividing the distribution into intervals and sampling from each interval (stratum) according to the distribution:

\[
E[X] = \sum_{j=1}^{m} p_j E[X \mid Y = y_j]
\]

where \( P\{Y = y_j\} = p_j \) for \( j = 1, \cdots, m \), is known.

Note:
1. For each stratum, we sample \( X \) conditioned on the even \( Y = y_j \)
2. The representative values are typically mean or median for each interval
Moment Matching

Moment matching involves adjusting the samples taken from a standard normal distribution so that the first, second, or possibly higher moments are matched.

For example, we want to sample from \( N(0,1) \).

Suppose that the samples are \( \varepsilon_i \) \((1 \leq i \leq n)\). To match the first two moments, we adjustify the samples by

\[
\varepsilon_i^* = \frac{\varepsilon_i - m}{s}
\]

where \( m \) and \( s \) are the mean and standard deviation of samples.

The adjusted samples \( \varepsilon_i^* \) has the correct mean 0 and standard deviation 1.

We then use the adjusted samples for calculation.
Quasi-Random Sequences

- Also called a low-discrepancy sequence is a sequence of representative samples from a probability distribution.
- At each stage of the simulation, the sampled points are roughly evenly distributed throughout the probability space.
- The standard error of the estimate is proportional to $\frac{1}{M}$. 
Further reading


Further reading

- [www.mcqmc.org](http://www.mcqmc.org)
- [www.mat.sbg.ac.at/~schmidw/links.html](http://www.mat.sbg.ac.at/~schmidw/links.html)
3. Monte Carlo Simulation
Asian Options (Chapter 22)

• Payoff related to average stock price
• Average Price options pay:
  – Call: $\max(S_{\text{ave}} - K, 0)$
  – Put: $\max(K - S_{\text{ave}}, 0)$
• Average Strike options pay:
  – Call: $\max(S_T - S_{\text{ave}}, 0)$
  – Put: $\max(S_{\text{ave}} - S_T, 0)$
Asian Options (Chapter 22)

- For arithmetic averaging:

\[ S_{\text{ave}} = \frac{1}{N} \sum_{i=1}^{N} S(t_i), \quad \text{where } T = N\Delta t \]

- For geometric averaging:

\[ S_{\text{ave}} = \sqrt[N]{S(t_1)S(t_2)\cdots S(t_N)} \]
Asian Options (geometric averaging)

- Closed form (Kemna & Vorst, 1990, J. of Banking and Finance, 14:113-129) because the geometric average of the underlying prices follows a lognormal distribution as well.

\[ c_G = S e^{(b-r)(T-t)} N(d_1) - X e^{-r(T-t)} N(d_2) \]

\[ p_G = X e^{-r(T-t)} N(-d_2) - S e^{(b-r)(T-t)} N(-d_1) \]

where \( N(x) \) is the cumulative normal distribution function and

\[ d_1 = \frac{\ln \left( \frac{S}{X} \right) + (b + 0.5 \sigma_A^2) T}{\sigma_A \sqrt{T}} \]

\[ d_2 = \frac{\ln \left( \frac{S}{X} \right) + (b - 0.5 \sigma_A^2) T}{\sigma_A \sqrt{T}} \]

\[ d_2 = d_1 - \sigma_A \sqrt{T} \]

\[ \sigma_A = \frac{\sigma}{\sqrt{3}} \]

\[ b = \frac{1}{2} \left( r - D - \frac{\sigma^2}{6} \right) \]
Asian Options (arithmetic averaging)

- No analytic solution
- Can be valued by assuming (as an approximation) that the average stock price is lognormally distributed
Consider an Asian Put on a stock with non-dividend payment. $S(0) = \$50$, $X = 50$, $T = 5$ months, and annual risk-free interest rate $r = 10\%$ and a volatility $\sigma = 40\%$ per annum.

- MC Price with 100,000 simulations: $\$3.9806$
  Confidence Interval $[3.9195, 4.0227]$
- MCAV Price with 100,000 simulations: $\$3.9581$
  Confidence Interval $[3.9854, 3.9309]$