Implicit Methods

- Suppose we use $t = t_{j+1}$ in Eq. (64) on p. 601 instead.

- The finite-difference equation becomes

$$\frac{\theta_{i,j+1} - \theta_{i,j}}{\Delta t} = D \frac{\theta_{i+1,j+1} - 2\theta_{i,j+1} + \theta_{i-1,j+1}}{(\Delta x)^2}. \tag{66}$$

- The stencil involves $\theta_{i,j}$, $\theta_{i,j+1}$, $\theta_{i+1,j+1}$, and $\theta_{i-1,j+1}$.

- This method is implicit:

  - The value of any one of the three quantities at t_{j+1} cannot be calculated unless the other two are known.

  - See exhibit (b) on p. 604.

Implicit Methods (continued)

- Equation (66) can be rearranged as

$$\theta_{i-1,j+1} - (2 + \gamma)\, \theta_{i,j+1} + \theta_{i+1,j+1} = -\gamma \theta_{i,j},$$

  where $\gamma \equiv (\Delta x)^2/(D\Delta t)$.

- This equation is unconditionally stable.

- Suppose the boundary conditions are given at $x = x_0$ and $x = x_{N+1}$.

- After $\theta_{i,j}$ has been calculated for $i = 1, 2, \ldots, N$, the values of $\theta_{i,j+1}$ at time t_{j+1} can be computed as the solution to the following tridiagonal linear system,

Implicit Methods (continued)

$$
\begin{bmatrix}
a & 1 & 0 & \cdots & \cdots & 0 \\
1 & a & 1 & 0 & \cdots & 0 \\
0 & 1 & a & 1 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & 1 & a & 1 \\
0 & \cdots & \cdots & \cdots & 1 & a
\end{bmatrix}
\begin{bmatrix}
\theta_{1,j+1} \\
\theta_{2,j+1} \\
\theta_{3,j+1} \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\theta_{N,j+1}
\end{bmatrix}
=
\begin{bmatrix}
-\gamma\theta_{1,j} - \theta_{0,j+1} \\
-\gamma\theta_{2,j} \\
-\gamma\theta_{3,j} \\
\vdots \\
\vdots \\
-\gamma\theta_{N-1,j} \\
-\gamma\theta_{N,j} - \theta_{N+1,j+1}
\end{bmatrix},
$$

where $a \equiv -2 - \gamma$.

Implicit Methods (concluded)

- Tridiagonal systems can be solved in $O(N)$ time and $O(N)$ space.

- The matrix above is nonsingular when $\gamma \geq 0$.

  – A square matrix is nonsingular if its inverse exists.

Crank-Nicolson Method

- Take the average of explicit method (65) on p. 602 and implicit method (66) on p. 608:
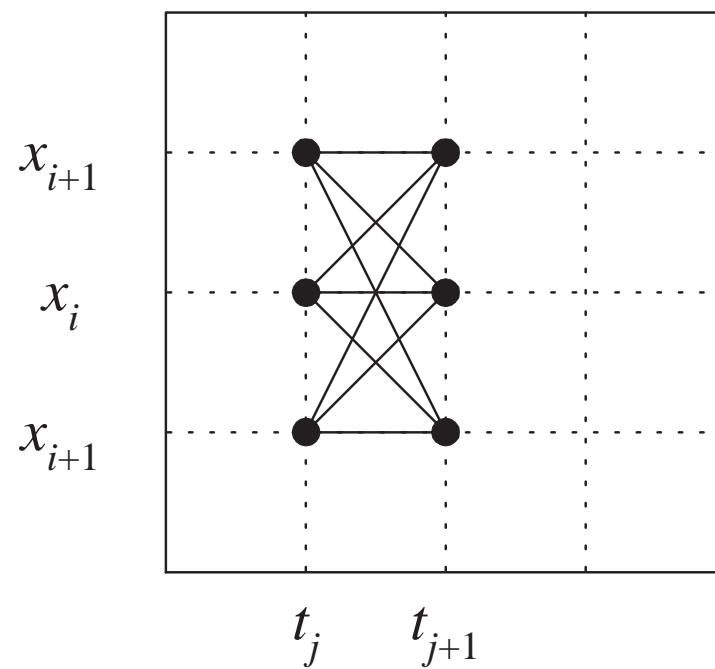
$$
\frac{\theta_{i,j+1} - \theta_{i,j}}{\Delta t}
$$

$$
= \quad \frac{1}{2} \left(D \, \frac{\theta_{i+1,j} - 2\theta_{i,j} + \theta_{i-1,j}}{(\Delta x)^2} + D \, \frac{\theta_{i+1,j+1} - 2\theta_{i,j+1} + \theta_{i-1,j+1}}{(\Delta x)^2} \right).
$$

- After rearrangement,

$$
\gamma \theta_{i,j+1} - \frac{\theta_{i+1,j+1} - 2\theta_{i,j+1} + \theta_{i-1,j+1}}{2} = \gamma \theta_{i,j} + \frac{\theta_{i+1,j} - 2\theta_{i,j} + \theta_{i-1,j}}{2}.
$$

- This is an unconditionally stable implicit method with excellent rates of convergence.

Stencil

Numerically Solving the Black-Scholes PDE

- See text.

Monte Carlo Simulation[a]

- Monte Carlo simulation is a sampling scheme.

- In many important applications within finance and without, Monte Carlo is one of the few feasible tools.

- When the time evolution of a stochastic process is not easy to describe analytically, Monte Carlo may very well be the only strategy that succeeds consistently.

---

[a]A top 10 algorithm according to Dongarra and Sullivan (2000).

The Big Idea

- Assume $X_1, X_2, \ldots, X_n$ have a joint distribution.

- $\theta \equiv E[\, g(X_1, X_2, \ldots, X_n)\,]$ for some function g is desired.

- We generate

$$\left(x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)} \right), \quad 1 \leq i \leq N$$

independently with the same joint distribution as $(X_1, X_2, \ldots, X_n)$.

- Set

$$Y_i \equiv g\left(x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)} \right).$$

The Big Idea (concluded)

- $Y_1, Y_2, \ldots, Y_N$ are independent and identically distributed random variables.

- Each Y_i has the same distribution as

$$Y \equiv g(X_1, X_2, \ldots, X_n).$$

- Since the average of these N random variables, $\overline{Y}$, satisfies $E[\overline{Y}] = \theta$, it can be used to estimate θ.

- The strong law of large numbers says that this procedure converges almost surely.

- The number of replications (or independent trials), N, is called the sample size.

Accuracy

- The Monte Carlo estimate and true value may differ owing to two reasons:

  1. Sampling variation.

  2. The discreteness of the sample paths.[a]

- The first can be controlled by the number of replications.

- The second can be controlled by the number of observations along the sample path.

---

[a]This may not be an issue if the derivative only requires discrete sampling along the time dimension.

Accuracy and Number of Replications

- The statistical error of the sample mean $\overline{Y}$ of the random variable Y grows as $1/\sqrt{N}$.
  - Because $\mathrm{Var}[\overline{Y}] = \mathrm{Var}[Y]/N$.

- In fact, this convergence rate is asymptotically optimal by the Berry-Esseen theorem.

- So the variance of the estimator $\overline{Y}$ can be reduced by a factor of $1/N$ by doing N times as much work.

- This is amazing because the same order of convergence holds independently of the dimension n.

Accuracy and Number of Replications (concluded)

- In contrast, classic numerical integration schemes have an error bound of $O(N^{-c/n})$ for some constant $c > 0$.

  - n is the dimension.

- The required number of evaluations thus grows exponentially in n to achieve a given level of accuracy.

  - The curse of dimensionality.

- The Monte Carlo method, for example, is more efficient than alternative procedures for securities depending on more than one asset, the multivariate derivatives.

Variance Reduction

- The statistical efficiency of Monte Carlo simulation can be measured by the variance of its output.

- If this variance can be lowered without changing the expected value, fewer replications are needed.

- Methods that improve efficiency in this manner are called variance-reduction techniques.

- Such techniques become practical when the added costs are outweighed by the reduction in sampling.

Monte Carlo Option Pricing

- For the pricing of European options on a dividend-paying stock, we may proceed as follows.

- Stock prices $S_1, S_2, S_3, \ldots$ at times $\Delta t, 2\Delta t, 3\Delta t, \ldots$ can be generated via

$$S_{i+1} = S_i e^{(\mu - \sigma^2/2)\,\Delta t + \sigma\sqrt{\Delta t}\,\xi}, \quad \xi \sim N(0, 1)$$

(67)

when $dS/S = \mu\,dt + \sigma\,dW$.

Monte Carlo Option Pricing (continued)

- If we discretize $dS/S = \mu \, dt + \sigma \, dW$, we will obtain

$$S_{i+1} = S_i + (\mu - \sigma^2/2) \, \Delta t + \sigma \sqrt{\Delta t} \, \xi.$$

- But this is locally normally distributed, not lognormally, hence biased.[a]

- In practice, this is not expected to be a major problem as long as Δt is sufficiently small.

---

[a]Contributed by Mr. Tai, Hui-Chin (R97723028) on April 22, 2009.

Monte Carlo Option Pricing (concluded)

- Non-dividend-paying stock prices in a risk-neutral economy can be generated by setting $\mu = r$.

  1: $C := 0;$

  2: **for** $i = 1, 2, 3, \ldots, m$ **do**

  3: $\quad P := S \times e^{(r-\sigma^2/2)\,T + \sigma\sqrt{T}\,\xi};$

  4: $\quad C := C + \max(P - X, 0);$

  5: **end for**

  6: return $Ce^{-rT}/m;$

- Pricing Asian options is easy (see text).

Pricing American Options

- Standard Monte Carlo simulation is inappropriate for American options because of early exercise (why?).

- It is difficult to determine the early-exercise point based on one single path.

- Monte Carlo simulation can be modified to price American options with small biases (p. 669ff).[a]

---

[a]Longstaff and Schwartz (2001).

Delta and Common Random Numbers

- In estimating delta, it is natural to start with the finite-difference estimate

$$e^{-r\tau} \frac{E[\,P(S+\epsilon)\,] - E[\,P(S-\epsilon)\,]}{2\epsilon}.$$

  - $P(x)$ is the terminal payoff of the derivative security when the underlying asset's initial price equals x.

- Use simulation to estimate $E[\,P(S+\epsilon)\,]$ first.

- Use another simulation to estimate $E[\,P(S-\epsilon)\,]$.

- Finally, apply the formula to approximate the delta.

Delta and Common Random Numbers (concluded)

- This method is not recommended because of its high variance.

- A much better approach is to use common random numbers to lower the variance:

$$e^{-r\tau} E\left[\frac{P(S+\epsilon) - P(S-\epsilon)}{2\epsilon} \right].$$

- Here, the *same* random numbers are used for $P(S+\epsilon)$ and $P(S-\epsilon)$.

- This holds for gamma and cross gammas (for multivariate derivatives).

Gamma

- The finite-difference formula for gamma is

$$e^{-r\tau} E \left[\frac{P(S + \epsilon) - 2 \times P(S) + P(S - \epsilon)}{\epsilon^2} \right].$$

- For a correlation option with multiple underlying assets, the finite-difference formula for the cross gammas $\partial^2 P(S_1, S_2, \dots)/(\partial S_1 \partial S_2)$ is:

$$e^{-r\tau} E \left[\frac{P(S_1 + \epsilon_1, S_2 + \epsilon_2) - P(S_1 - \epsilon_1, S_2 + \epsilon_2)}{4\epsilon_1\epsilon_2} \right.$$

$$\left. \frac{-P(S_1 + \epsilon_1, S_2 - \epsilon_2) + P(S_1 - \epsilon_1, S_2 - \epsilon_2)}{} \right].$$

Gamma (concluded)

- Choosing an ϵ of the right magnitude can be challenging.

  - If ϵ is too large, inaccurate Greeks result.

  - If ϵ is too small, unstable Greeks result.

- This phenomenon is sometimes called the curse of differentiation.

- Need formulas for Greeks which are integrals (thus avoiding ϵ, finite differences, and resimulation).[a]

---

[a]Lyuu and Teng (2008).

Biases in Pricing Continuously Monitored Options with Monte Carlo

- We are asked to price a continuously monitored up-and-out call with barrier H.

- The Monte Carlo method samples the stock price at n discrete time points $t_1, t_2, \ldots, t_n$.

- A sample path $S(t_0), S(t_1), \ldots, S(t_n)$ is produced.
  - Here, $t_0 = 0$ is the current time, and $t_n = T$ is the expiration time of the option.

Biases in Pricing Continuously Monitored Options with Monte Carlo (continued)

- If all of the sampled prices are below the barrier, this sample path pays $\max(S(t_n) - X, 0)$.

- Repeating these steps and averaging the payoffs yield a Monte Carlo estimate.

```
 1: C := 0;
 2: for i = 1, 2, 3, ..., m do
 3:     P := S; hit := 0;
 4:     for j = 1, 2, 3, ..., n do
 5:         P := P × e^{(r−σ²/2)(T/n)+σ√(T/n) ξ};
 6:         if P ≥ H then
 7:             hit := 1;
 8:             break;
 9:         end if
10:     end for
11:     if hit = 0 then
12:         C := C + max(P − X, 0);
13:     end if
14: end for
15: return Ce^{−rT}/m;
```

Biases in Pricing Continuously Monitored Options with Monte Carlo (continued)

- This estimate is biased.

  - Suppose none of the sampled prices on a sample path equals or exceeds the barrier H.

  - It remains possible for the continuous sample path that passes through them to hit the barrier between sampled time points (see plot on next page).

H
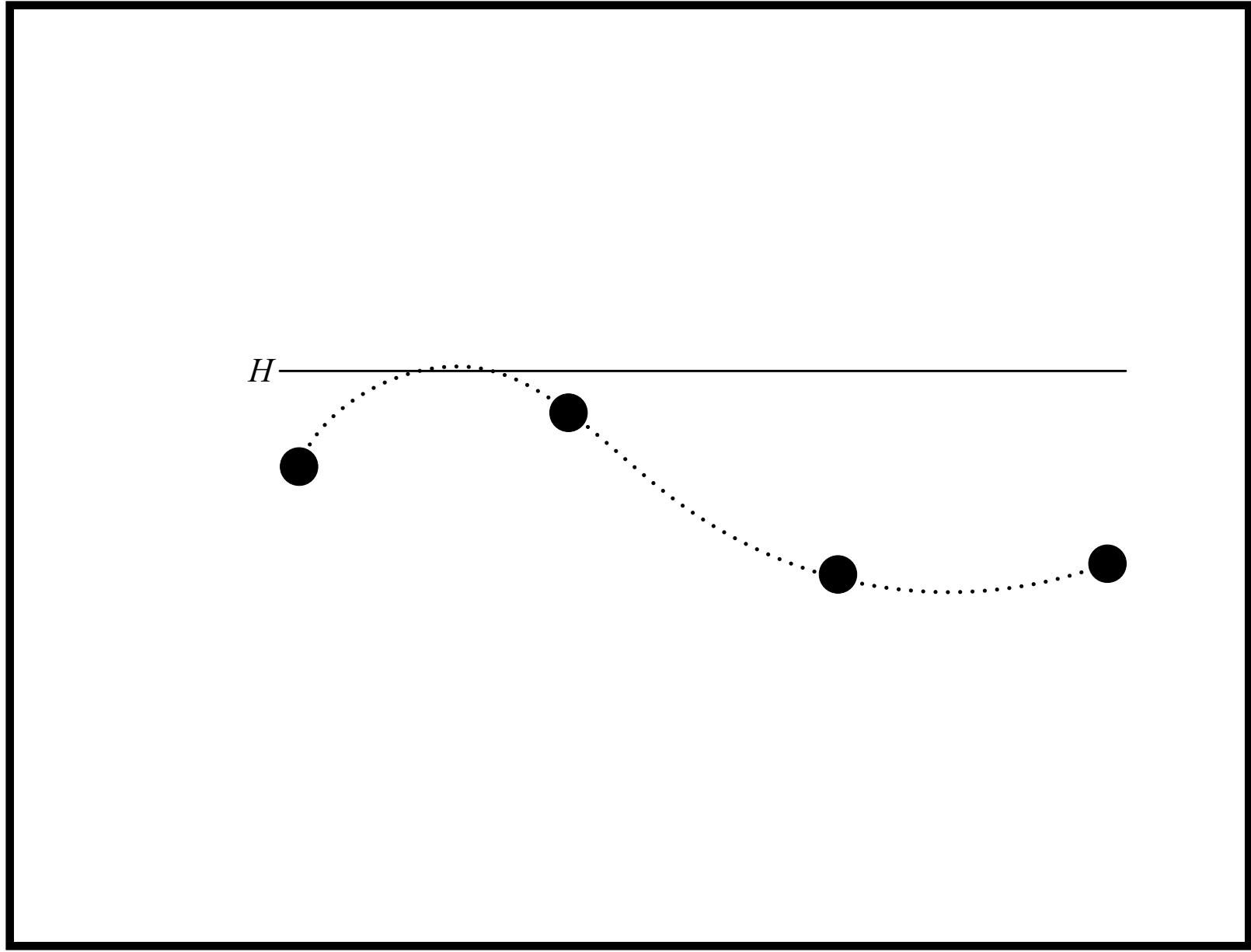
Biases in Pricing Continuously Monitored Options with Monte Carlo (concluded)

- The bias can certainly be lowered by increasing the number of observations along the sample path.

- However, even daily sampling may not suffice.

- The computational cost also rises as a result.

Brownian Bridge Approach to Pricing Barrier Options

- We desire an unbiased estimate efficiently.

- So the above-mentioned payoff should be multiplied by the probability p that a continuous sample path does not hit the barrier conditional on the sampled prices.

- This methodology is called the Brownian bridge approach.

- Formally, we have

$$p \equiv \mathrm{Prob}[\, S(t) < H, 0 \leq t \leq T \,|\, S(t_0), S(t_1), \ldots, S(t_n)\,].$$

Brownian Bridge Approach to Pricing Barrier Options (continued)

- As a barrier is hit over a time interval if and only if the maximum stock price over that period is at least H,

$$p = \mathrm{Prob}\left[\max_{0 \leq t \leq T} S(t) < H \mid S(t_0), S(t_1), \dots, S(t_n) \right].$$

- Luckily, the conditional distribution of the maximum over a time interval given the beginning and ending stock prices is known.

Brownian Bridge Approach to Pricing Barrier Options (continued)

Lemma 19 *Assume S follows $dS/S = \mu\,dt + \sigma\,dW$ and define*

$$\zeta(x) \equiv \exp\left[-\frac{2\ln(x/S(t))\ln(x/S(t+\Delta t))}{\sigma^2\Delta t} \right].$$

(1) If $H > \max(S(t), S(t+\Delta t))$, then

$$\mathrm{Prob}\left[\max_{t\leq u\leq t+\Delta t} S(u) < H \,\middle|\, S(t), S(t+\Delta t) \right] = 1 - \zeta(H).$$

(2) If $h < \min(S(t), S(t+\Delta t))$, then

$$\mathrm{Prob}\left[\min_{t\leq u\leq t+\Delta t} S(u) > h \,\middle|\, S(t), S(t+\Delta t) \right] = 1 - \zeta(h).$$

Brownian Bridge Approach to Pricing Barrier Options (continued)

- Lemma 19 gives the probability that the barrier is not hit in a time interval, given the starting and ending stock prices.

- For our up-and-out call, choose $n = 1$.

- As a result,

$$
p = \begin{cases} 1 - \exp\left[-\frac{2\ln(H/S(0))\ln(H/S(T))}{\sigma^2 T} \right], & \text{if } H > \max(S(0), S(T)), \\ 0, & \text{otherwise.} \end{cases}
$$

Brownian Bridge Approach to Pricing Barrier Options (continued)

1: $C := 0$;
2: **for** $i = 1, 2, 3, \ldots, m$ **do**
3:    $P := S \times e^{(r-q-\sigma^2/2)\,T + \sigma\sqrt{T}\,\xi(\,)}$;
4:    **if** $(S < H \text{ and } P < H)$ or $(S > H \text{ and } P > H)$ **then**
5:       $C := C + \max(P - X, 0) \times \left\{ 1 - \exp\left[-\frac{2\ln(H/S) \times \ln(H/P)}{\sigma^2 T} \right] \right\}$;
6:    **end if**
7: **end for**
8: **return** Ce^{-rT}/m;

Brownian Bridge Approach to Pricing Barrier Options (concluded)

- The idea can be generalized.

- For example, we can handle more complex barrier options.

- Consider an up-and-out call with barrier H_i for the time interval $(t_i, t_{i+1}]$, $0 \le i < n$.

- This option thus contains n barriers.

- It is a simple matter of multiplying the probabilities for the n time intervals properly to obtain the desired probability adjustment term.

Variance Reduction: Antithetic Variates

- We are interested in estimating $E[\,g(X_1, X_2, \ldots, X_n)\,]$, where $X_1, X_2, \ldots, X_n$ are independent.

- Let Y_1 and Y_2 be random variables with the same distribution as $g(X_1, X_2, \ldots, X_n)$.

- Then

$$\text{Var}\left[\frac{Y_1 + Y_2}{2}\right] = \frac{\text{Var}[\,Y_1\,]}{2} + \frac{\text{Cov}[\,Y_1, Y_2\,]}{2}.$$

  – $\text{Var}[\,Y_1\,]/2$ is the variance of the Monte Carlo method with two (independent) replications.

- The variance $\text{Var}[\,(Y_1 + Y_2)/2\,]$ is smaller than $\text{Var}[\,Y_1\,]/2$ when Y_1 and Y_2 are negatively correlated.

Variance Reduction: Antithetic Variates (continued)

- For each simulated sample path X, a second one is obtained by reusing the random numbers on which the first path is based.

- This yields a second sample path Y.

- Two estimates are then obtained: One based on X and the other on Y.

- If N independent sample paths are generated, the antithetic-variates estimator averages over $2N$ estimates.

Variance Reduction: Antithetic Variates (continued)

- Consider process $dX = a_t \, dt + b_t \sqrt{dt} \, \xi$.

- Let g be a function of n samples $X_1, X_2, \dots, X_n$ on the sample path.

- We are interested in $E[\, g(X_1, X_2, \dots, X_n)\,]$.

- Suppose one simulation run has realizations $\xi_1, \xi_2, \dots, \xi_n$ for the normally distributed fluctuation term ξ.

- This generates samples $x_1, x_2, \dots, x_n$.

- The estimate is then $g(\boldsymbol{x})$, where $\boldsymbol{x} \equiv (x_1, x_2 \dots, x_n)$.

Variance Reduction: Antithetic Variates (concluded)

- The antithetic-variates method does not sample n more numbers from ξ for the second estimate $g(\boldsymbol{x}')$.

- Instead, generate the sample path $\boldsymbol{x}' \equiv (x_1', x_2' \dots, x_n')$ from $-\xi_1, -\xi_2, \dots, -\xi_n$.

- Compute $g(\boldsymbol{x}')$.

- Output $(g(\boldsymbol{x}) + g(\boldsymbol{x}'))/2$.

- Repeat the above steps for as many times as required by accuracy.

Variance Reduction: Conditioning

- We are interested in estimating $E[X]$.

- Suppose here is a random variable Z such that $E[X \mid Z = z]$ can be efficiently and precisely computed.

- $E[X] = E[E[X \mid Z]]$ by the law of iterated conditional expectations.

- Hence the random variable $E[X \mid Z]$ is also an unbiased estimator of $E[X]$.

Variance Reduction: Conditioning (concluded)

- As
$$\text{Var}[\,E[\,X\,|\,Z\,]\,] \le \text{Var}[\,X\,],$$

  $E[\,X\,|\,Z\,]$ has a smaller variance than observing X directly.

- First obtain a random observation z on Z.

- Then calculate $E[\,X\,|\,Z=z\,]$ as our estimate.

  - There is no need to resort to simulation in computing $E[\,X\,|\,Z=z\,]$.

- The procedure can be repeated a few times to reduce the variance.

Control Variates

- Use the analytic solution of a similar yet simpler problem to improve the solution.

- Suppose we want to estimate $E[X]$ and there exists a random variable Y with a known mean $\mu \equiv E[Y]$.

- Then $W \equiv X + \beta(Y - \mu)$ can serve as a "controlled" estimator of $E[X]$ for any constant β.

  - However β is chosen, W remains an unbiased estimator of $E[X]$ as

$$E[W] = E[X] + \beta E[Y - \mu] = E[X].$$

Control Variates (continued)

- Note that

$$\text{Var}[\,W\,] = \text{Var}[\,X\,] + \beta^2 \,\text{Var}[\,Y\,] + 2\beta \,\text{Cov}[\,X, Y\,],$$

$$(68)$$

- Hence W is less variable than X if and only if

$$\beta^2 \,\text{Var}[\,Y\,] + 2\beta \,\text{Cov}[\,X, Y\,] < 0. \qquad (69)$$

Control Variates (concluded)

- The success of the scheme clearly depends on both β and the choice of Y.

- For example, arithmetic average-rate options can be priced by choosing Y to be the otherwise identical geometric average-rate option's price and $\beta = -1$.

- This approach is much more effective than the antithetic-variates method.

Choice of Y

- In general, the choice of Y is ad hoc, and experiments must be performed to confirm the wisdom of the choice.

- Try to match calls with calls and puts with puts.[a]

- On many occasions, Y is a discretized version of the derivative that gives μ.

  - Discretely monitored geometric average-rate option vs. the continuously monitored geometric average-rate option given by formulas (30) on p. 339.

- For some choices, the discrepancy can be significant, such as the lookback option.[b]

---

[a]Contributed by Ms. Teng, Huei-Wen (R91723054) on May 25, 2004.
[b]Contributed by Mr. Tsai, Hwai (R92723049) on May 12, 2004.

Optimal Choice of β

- Equation (68) on p. 649 is minimized when

$$\beta = -\mathrm{Cov}[\,X, Y\,]/\mathrm{Var}[\,Y\,],$$

  which was called beta in the book.

- For this specific β,

$$\mathrm{Var}[\,W\,] = \mathrm{Var}[\,X\,] - \frac{\mathrm{Cov}[\,X, Y\,]^2}{\mathrm{Var}[\,Y\,]} = \left(1 - \rho_{X,Y}^2\right)\mathrm{Var}[\,X\,],$$

  where $\rho_{X,Y}$ is the correlation between X and Y.

- The stronger X and Y are correlated, the greater the reduction in variance.

Optimal Choice of β (continued)

- For example, if this correlation is nearly perfect (± 1), we could control X almost exactly.

- Typically, neither $\mathrm{Var}[\,Y\,]$ nor $\mathrm{Cov}[\,X, Y\,]$ is known.

- Therefore, we cannot obtain the maximum reduction in variance.

- We can guess these values and hope that the resulting W does indeed have a smaller variance than X.

- A second possibility is to use the simulated data to estimate these quantities.

  - How to do it efficiently in terms of time and space?

Optimal Choice of β (concluded)

- Observe that $-\beta$ has the same sign as the correlation between X and Y.

- Hence, if X and Y are positively correlated, $\beta < 0$, then X is adjusted downward whenever $Y > \mu$ and upward otherwise.

- The opposite is true when X and Y are negatively correlated, in which case $\beta > 0$.

A Pitfall

- A potential pitfall is to sample X and Y independently.

- In this case, $\text{Cov}[\,X, Y\,] = 0$.

- Equation (68) on p. 649 becomes

$$\text{Var}[\,W\,] = \text{Var}[\,X\,] + \beta^2\,\text{Var}[\,Y\,].$$

- So whatever Y is, the variance is *increased*!

Problems with the Monte Carlo Method

- The error bound is only probabilistic.

- The probabilistic error bound of $\sqrt{N}$ does not benefit from regularity of the integrand function.

- The requirement that the points be independent random samples are wasteful because of clustering.

- In reality, pseudorandom numbers generated by completely deterministic means are used.

- Monte Carlo simulation exhibits a great sensitivity on the seed of the pseudorandom-number generator.