Multivariate Contingent Claims

- They depend on two or more underlying assets.
- The basket call on m assets has the terminal payoff

$$\max\left(\sum_{i=1}^{m} \alpha_i S_i(\tau) - X, 0\right),$$

  where α_i is the percentage of asset i.
- Basket options are essentially options on a portfolio of stocks (or index options).[a]
- Option on the best of two risky assets and cash has a terminal payoff of $\max(S_1(\tau), S_2(\tau), X)$.

[a]Except that membership and weights do *not* change for basket options (Bennett, 2014).

Multivariate Contingent Claims (concluded)[a]

Name	Payoff
Exchange option	$\max(S_1(\tau) - S_2(\tau), 0)$
Better-off option	$\max(S_1(\tau), \ldots, S_k(\tau), 0)$
Worst-off option	$\min(S_1(\tau), \ldots, S_k(\tau), 0)$
Binary maximum option	$I\{\max(S_1(\tau), \ldots, S_k(\tau)) > X\}$
Maximum option	$\max(\max(S_1(\tau), \ldots, S_k(\tau)) - X, 0)$
Minimum option	$\max(\min(S_1(\tau), \ldots, S_k(\tau)) - X, 0)$
Spread option	$\max(S_1(\tau) - S_2(\tau) - X, 0)$
Basket average option	$\max((S_1(\tau) + \cdots + S_k(\tau))/k - X, 0)$
Multi-strike option	$\max(S_1(\tau) - X_1, \ldots, S_k(\tau) - X_k, 0)$
Pyramid rainbow option	$\max(\mid S_1(\tau) - X_1 \mid + \cdots + \mid S_k(\tau) - X_k \mid - X, 0)$
Madonna option	$\max(\sqrt{(S_1(\tau) - X_1)^2 + \cdots + (S_k(\tau) - X_k)^2} - X, 0)$

[a] Lyuu & Teng (R91723054) (2011).

Correlated Trinomial Model[a]

- Two risky assets S_1 and S_2 follow
$$\frac{dS_i}{S_i} = r\,dt + \sigma_i\,dW_i$$
in a risk-neutral economy, $i = 1, 2$.

- Let
$$\begin{aligned} M_i &\stackrel{\Delta}{=} e^{r\Delta t}, \\ V_i &\stackrel{\Delta}{=} M_i^2(e^{\sigma_i^2\Delta t} - 1). \end{aligned}$$
  - $S_i M_i$ is the mean of S_i at time Δt.
  - $S_i^2 V_i$ the variance of S_i at time Δt.

[a]Boyle, Evnine, & Gibbs (1989).

Correlated Trinomial Model (continued)

- The value of $S_1 S_2$ at time Δt has a joint lognormal distribution with mean $S_1 S_2 M_1 M_2 e^{\rho \sigma_1 \sigma_2 \Delta t}$, where ρ is the correlation between dW_1 and dW_2.
- Next match the 1st and 2nd moments of the approximating discrete distribution to those of the continuous counterpart.
- At time Δt from now, there are 5 distinct outcomes.

Correlated Trinomial Model (continued)

- The five-point probability distribution of the asset prices is

Probability	Asset 1	Asset 2
p_1	$S_1 u_1$	$S_2 u_2$
p_2	$S_1 u_1$	$S_2 d_2$
p_3	$S_1 d_1$	$S_2 d_2$
p_4	$S_1 d_1$	$S_2 u_2$
p_5	S_1	S_2

- As usual, impose $u_i d_i = 1$.

Correlated Trinomial Model (continued)

- The probabilities must sum to one, and the means must be matched:

$$
\begin{aligned}
1 &= p_1 + p_2 + p_3 + p_4 + p_5, \\
S_1 M_1 &= (p_1 + p_2)\, S_1 u_1 + p_5 S_1 + (p_3 + p_4)\, S_1 d_1, \\
S_2 M_2 &= (p_1 + p_4)\, S_2 u_2 + p_5 S_2 + (p_2 + p_3)\, S_2 d_2.
\end{aligned}
$$

Correlated Trinomial Model (concluded)

- Let $R \stackrel{\Delta}{=} M_1 M_2 e^{\rho\sigma_1\sigma_2\Delta t}$.
- Match the variances and covariance:

$$\begin{aligned}
S_1^2 V_1 &= (p_1+p_2)\left[(S_1u_1)^2-(S_1M_1)^2\right]+p_5\left[S_1^2-(S_1M_1)^2\right]\\
&\quad +(p_3+p_4)\left[(S_1d_1)^2-(S_1M_1)^2\right],\\
S_2^2 V_2 &= (p_1+p_4)\left[(S_2u_2)^2-(S_2M_2)^2\right]+p_5\left[S_2^2-(S_2M_2)^2\right]\\
&\quad +(p_2+p_3)\left[(S_2d_2)^2-(S_2M_2)^2\right],\\
S_1S_2R &= (p_1u_1u_2+p_2u_1d_2+p_3d_1d_2+p_4d_1u_2+p_5)\,S_1S_2.
\end{aligned}$$

- The solutions appear on p. 246 of the textbook.

Correlated Trinomial Model Simplified[a]

- Let $\mu'_i \stackrel{\Delta}{=} r - \sigma_i^2/2$ and $u_i \stackrel{\Delta}{=} e^{\lambda \sigma_i \sqrt{\Delta t}}$ for $i = 1, 2$.
- The following simpler scheme is often good enough:

$$
\begin{aligned}
p_1 &= \frac{1}{4}\left[\frac{1}{\lambda^2} + \frac{\sqrt{\Delta t}}{\lambda}\left(\frac{\mu'_1}{\sigma_1} + \frac{\mu'_2}{\sigma_2}\right) + \frac{\rho}{\lambda^2}\right],\\
p_2 &= \frac{1}{4}\left[\frac{1}{\lambda^2} + \frac{\sqrt{\Delta t}}{\lambda}\left(\frac{\mu'_1}{\sigma_1} - \frac{\mu'_2}{\sigma_2}\right) - \frac{\rho}{\lambda^2}\right],\\
p_3 &= \frac{1}{4}\left[\frac{1}{\lambda^2} + \frac{\sqrt{\Delta t}}{\lambda}\left(-\frac{\mu'_1}{\sigma_1} - \frac{\mu'_2}{\sigma_2}\right) + \frac{\rho}{\lambda^2}\right],\\
p_4 &= \frac{1}{4}\left[\frac{1}{\lambda^2} + \frac{\sqrt{\Delta t}}{\lambda}\left(-\frac{\mu'_1}{\sigma_1} + \frac{\mu'_2}{\sigma_2}\right) - \frac{\rho}{\lambda^2}\right],\\
p_5 &= 1 - \frac{1}{\lambda^2}.
\end{aligned}
$$

[a] Madan, Milne, & Shefrin (1989).

Correlated Trinomial Model Simplified (continued)

- All of the probabilities lie between 0 and 1 if and only if

$$-1 + \lambda\sqrt{\Delta t}\left|\frac{\mu_1'}{\sigma_1} + \frac{\mu_2'}{\sigma_2}\right| \leq \rho \leq 1 - \lambda\sqrt{\Delta t}\left|\frac{\mu_1'}{\sigma_1} - \frac{\mu_2'}{\sigma_2}\right|, \quad (116)$$

$$1 \leq \lambda. \quad (117)$$

- We call a multivariate tree (correlation-) optimal if it guarantees valid probabilities as long as

$$-1 + O(\sqrt{\Delta t}) < \rho < 1 - O(\sqrt{\Delta t}),$$

such as the above one.[a]

---

[a] W. Kao (R98922093) (2011); W. Kao (R98922093), Lyuu, & Wen (D94922003) (2014).

Correlated Trinomial Model Simplified (continued)

- But this model cannot price 2-asset 2-barrier options accurately.[a]
- Few multivariate trees are both optimal and able to handle multiple barriers.[b]
- An alternative is to use orthogonalization.[c]

---

[a]See Y. Chang (B89704039, R93922034), Hsu (R7526001, D89922012), & Lyuu (2006); W. Kao (R98922093), Lyuu, & Wen (D94922003) (2014) for solutions.

[b]See W. Kao (R98922093), Lyuu, & Wen (D94922003) (2014) for an exception.

[c]Hull & White (1990); Dai (B82506025, R86526008, D8852600), C. Wang (F95922018), & Lyuu (2013).

Correlated Trinomial Model Simplified (concluded)

- Suppose we allow each asset's volatility to be a function of time.[a]
- There are k assets.
- Can you build an optimal multivariate tree that can handle two barriers on each asset in time $O(n^{k+1})$?[b]

---

[a]Recall p. 315.

[b]See Y. Zhang (R05922052) (2019) for a complete solution.

Extrapolation

- It is a method to speed up numerical convergence.
- Say $f(n)$ converges to an unknown limit f at rate of $1/n$:

$$f(n) = f + \frac{c}{n} + o\left(\frac{1}{n}\right). \tag{118}$$

- Assume c is an unknown constant independent of n.
  - Convergence is basically monotonic and smooth.

Extrapolation (concluded)

- From two approximations $f(n_1)$ and $f(n_2)$ and ignoring the smaller terms,

$$f(n_1) = f + \frac{c}{n_1},$$
$$f(n_2) = f + \frac{c}{n_2}.$$

- A better approximation to the desired f is

$$f = \frac{n_1 f(n_1) - n_2 f(n_2)}{n_1 - n_2}. \qquad (119)$$

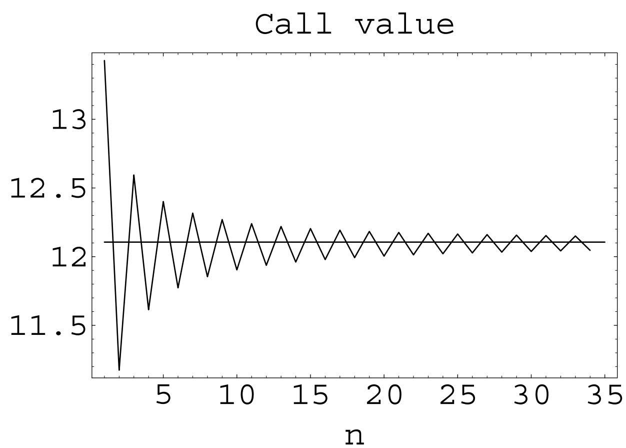
- This estimate should converge faster than $1/n$.[a]
- The Richardson extrapolation uses $n_2 = 2n_1$.

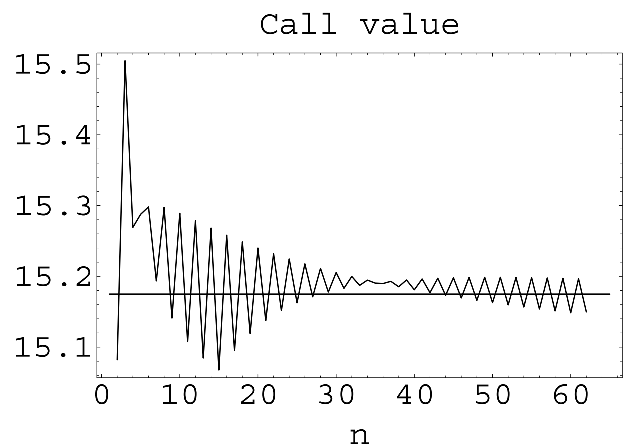
[a]It is identical to the forward rate formula (22) on p. 150!

Improving BOPM with Extrapolation

- Consider standard European options.
- Denote the option value under BOPM using n time periods by $f(n)$.
- It is known that BOPM convergences at the rate of $1/n$,[a] consistent with Eq. (118) on p. 830.
- The plots on p. 306 (redrawn on next page) show that convergence to the true option value oscillates with n.
- Extrapolation is inapplicable at this stage.

[a] L. Chang & Palmer (2007); F. Diener & M. Diener (2004).


Call value
13
12.5
12
11.5
5
10
15
20
25
30
35
n

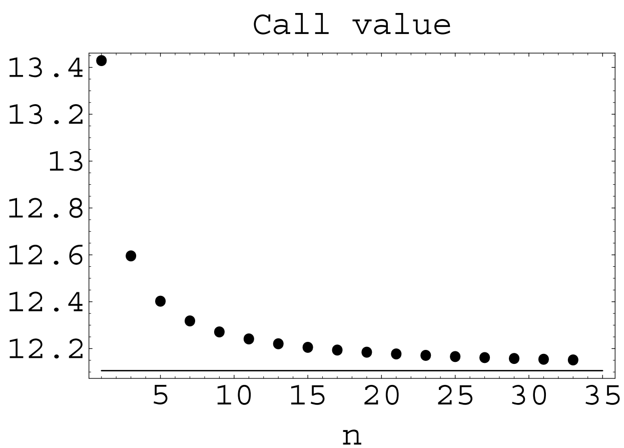

Call value
15.5
15.4
15.3
15.2
15.1
0
10
20
30
40
50
60
n

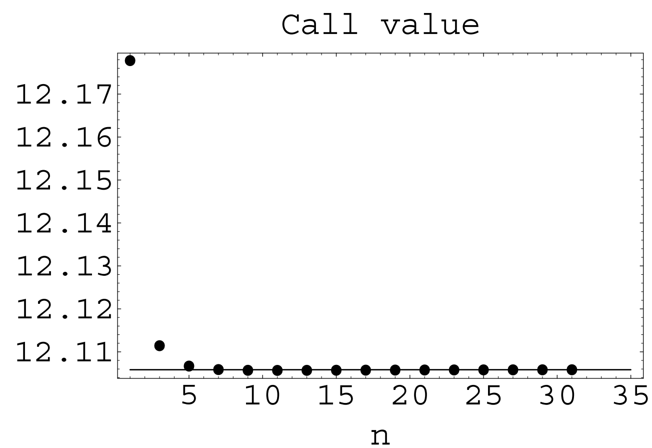
Improving BOPM with Extrapolation (concluded)

- Take the at-the-money option in the left plot on p. 833.
- The sequence with odd n turns out to be monotonic and smooth (see the left plot on p. 835).[a]
- Apply extrapolation (119) on p. 831 with $n_2 = n_1 + 2$, where n_1 is odd.
- Result is shown in the right plot on p. 835.
- The convergence rate is amazing.
- See Exercise 9.3.8 (p. 111) of the text for ideas in the general case.

---

[a]This can be proved (L. Chang & Palmer, 2007; F. Diener & M. Diener, 2004).


Call value
13.4
13.2
13
12.8
12.6
12.4
12.2
5 10 15 20 25 30 35
n

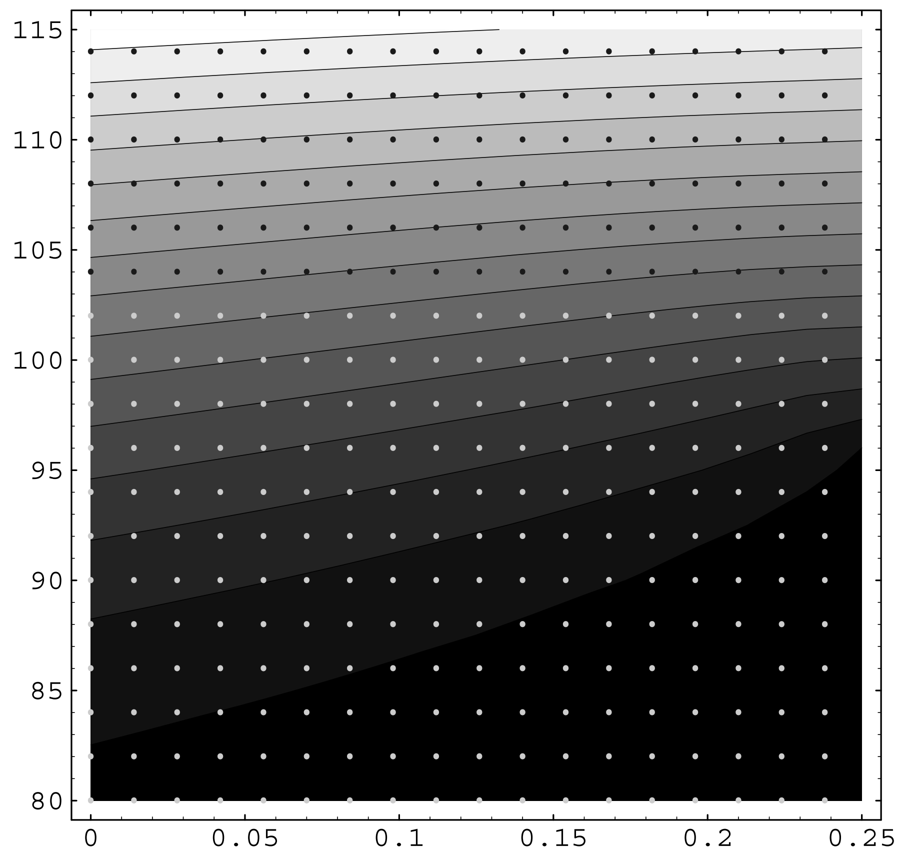

Call value
12.17
12.16
12.15
12.14
12.13
12.12
12.11
5 10 15 20 25 30 35
n

Numerical Methods

All science is dominated
by the idea of approximation.
— Bertrand Russell

Finite-Difference Methods

- Place a grid of points on the space over which the desired function takes value.
- Then approximate the function value at each of these points (p. 839).
- Solve the equation numerically by introducing difference equations in place of derivatives.


115
110
105
100
95
90
85
80
0
0.05
0.1
0.15
0.2
0.25

Example: Poisson's Equation

- It is $\partial^2\theta/\partial x^2 + \partial^2\theta/\partial y^2 = -\rho(x, y)$, which describes the electrostatic field.
- Replace second derivatives with finite differences through central difference.
- Introduce evenly spaced grid points with distance of Δx along the x axis and Δy along the y axis.
- The finite-difference form is

$$-\rho(x_i, y_j) = \frac{\theta(x_{i+1}, y_j) - 2\theta(x_i, y_j) + \theta(x_{i-1}, y_j)}{(\Delta x)^2} + \frac{\theta(x_i, y_{j+1}) - 2\theta(x_i, y_j) + \theta(x_i, y_{j-1})}{(\Delta y)^2}.$$

Example: Poisson's Equation (concluded)

- In the above, $\Delta x \triangleq x_i - x_{i-1}$ and $\Delta y \triangleq y_j - y_{j-1}$ for $i, j = 1, 2, \ldots$.
- When the grid points are evenly spaced in both axes so that $\Delta x = \Delta y = h$, the difference equation becomes

$$\begin{aligned} -h^2 \rho(x_i, y_j) &= \theta(x_{i+1}, y_j) + \theta(x_{i-1}, y_j) \\ &+ \theta(x_i, y_{j+1}) + \theta(x_i, y_{j-1}) - 4\theta(x_i, y_j). \end{aligned}$$

- Given boundary values, we can solve for the x_is and the y_js within the square $[\pm L, \pm L]$.
- From now on, $\theta_{i,j}$ will denote the finite-difference approximation to the exact $\theta(x_i, y_j)$.

Explicit Methods

- Consider the diffusion equation[a] $D(\partial^2\theta/\partial x^2) - (\partial\theta/\partial t) = 0$, $D > 0$.
- Use evenly spaced grid points (x_i, t_j) with distances Δx and Δt, where $\Delta x \stackrel{\Delta}{=} x_{i+1} - x_i$ and $\Delta t \stackrel{\Delta}{=} t_{j+1} - t_j$.
- Employ central difference for the second derivative and forward difference for the time derivative to obtain

$$\left.\frac{\partial\theta(x,t)}{\partial t}\right|_{t=t_j} = \frac{\theta(x, \boxed{t_{j+1}}) - \theta(x, \boxed{t_j})}{\Delta t} + \cdots, \tag{120}$$

$$\left.\frac{\partial^2\theta(x,t)}{\partial x^2}\right|_{x=x_i} = \frac{\theta(\boxed{x_{i+1}}, t) - 2\theta(\boxed{x_i}, t) + \theta(\boxed{x_{i-1}}, t)}{(\Delta x)^2} + \cdots \tag{121}$$

[a] It is a parabolic partial differential equation.

Explicit Methods (continued)

- Next, assemble Eqs. (120) and (121) into a single equation at (x_i, t_j).
- But we need to decide how to evaluate x in the first equation and t in the second.
- Since central difference around x_i is used in Eq. (121), we might as well use x_i for x in Eq. (120).
- Two choices are possible for t in Eq. (121).
- The first choice uses $t = t_j$ to yield the following finite-difference equation,

$$\frac{\theta_{i,j+1} - \theta_{i,j}}{\Delta t} = D\,\frac{\theta_{i+1,j} - 2\theta_{i,j} + \theta_{i-1,j}}{(\Delta x)^2}. \tag{122}$$

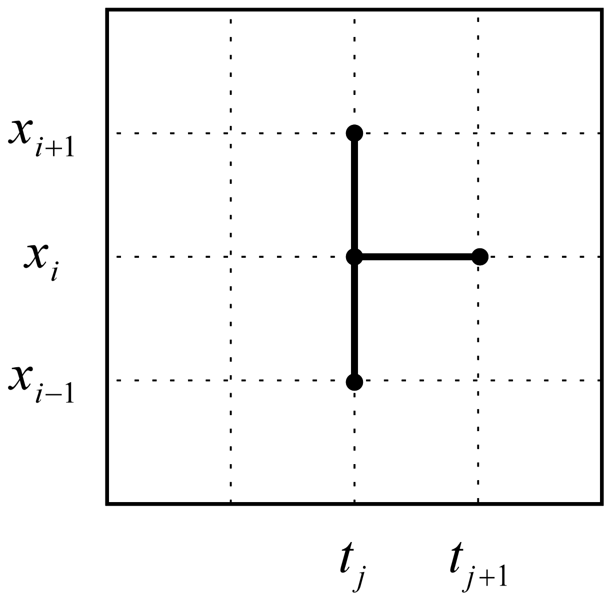
Explicit Methods (continued)

- The stencil of grid points involves four values, $\theta_{i,j+1}$, $\theta_{i,j}$, $\theta_{i+1,j}$, and $\theta_{i-1,j}$.
- Rearrange Eq. (122) on p. 843 as

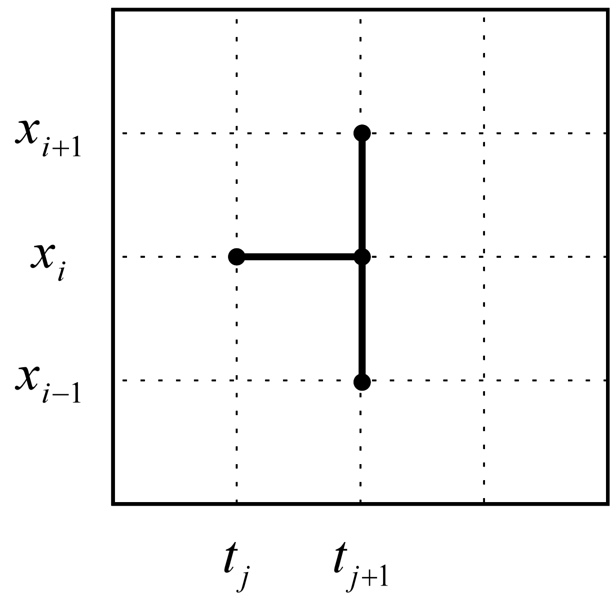
$$\theta_{i,j+1} = \frac{D\Delta t}{(\Delta x)^2}\,\theta_{i+1,j} + \left(1 - \frac{2D\Delta t}{(\Delta x)^2}\right)\theta_{i,j} + \frac{D\Delta t}{(\Delta x)^2}\,\theta_{i-1,j}. \tag{123}$$

- We can calculate $\theta_{i,j+1}$ from $\theta_{i,j}, \theta_{i+1,j}, \theta_{i-1,j}$, at the previous time t_j (see exhibit (a) on next page).

Stencils



(a)



(b)

Explicit Methods (concluded)

- Starting from the initial conditions at t_0, that is, $\theta_{i,0} = \theta(x_i, t_0)$, $i = 1, 2, \ldots$, we calculate

$$\theta_{i,1}, \quad i = 1, 2, \ldots .$$

- And then

$$\theta_{i,2}, \quad i = 1, 2, \ldots .$$

- And so on.

Stability

- The explicit method is numerically unstable unless

$$\Delta t \leq (\Delta x)^2/(2D).$$

  - A numerical method is unstable if the solution is highly sensitive to changes in initial conditions.

- The stability condition may lead to high running times and memory requirements.

- For instance, halving Δx would imply quadrupling $(\Delta t)^{-1}$, resulting in a running time 8 times as much.

Explicit Method and Trinomial Tree

- Recall Eq. (123) on p. 844:

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

- When the stability condition is satisfied, the three coefficients for $\theta_{i+1,j}$, $\theta_{i,j}$, and $\theta_{i-1,j}$ all lie between zero and one and sum to one.
- They can be interpreted as probabilities.
- So the finite-difference equation becomes identical to backward induction on trinomial trees!

Explicit Method and Trinomial Tree (concluded)

- The freedom in choosing Δx corresponds to similar freedom in the construction of trinomial trees.
- The explicit finite-difference equation is also identical to backward induction on a binomial tree.[a]
  - Let the binomial tree take 2 steps each of length $\Delta t/2$.
  - It is now a trinomial tree.

[a] Hilliard (2014).

Implicit Methods

- Suppose we use $t = t_{j+1}$ in Eq. (121) on p. 842 instead.
- The finite-difference equation becomes

$$\frac{\theta_{i,j+1} - \theta_{i,j}}{\Delta t} = D\, \frac{\theta_{i+1,j+1} - 2\theta_{i,j+1} + \theta_{i-1,j+1}}{(\Delta x)^2}. \tag{124}$$

- The stencil involves $\theta_{i,j}$, $\theta_{i,j+1}$, $\theta_{i+1,j+1}$, and $\theta_{i-1,j+1}$.
- This method is now 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. 845.

Implicit Methods (continued)

- Equation (124) 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 $\gamma \stackrel{\Delta}{=} (\Delta x)^2/(D\Delta t)$.

- 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)

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

where $a \stackrel{\Delta}{=} -2 - \gamma$.

Implicit Methods (concluded)

- Tridiagonal systems can be solved in $O(N)$ time and $O(N)$ space.
  - Never invert a matrix to solve a tridiagonal system.
- 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 (122) on p. 843 and implicit method (124) on p. 850:

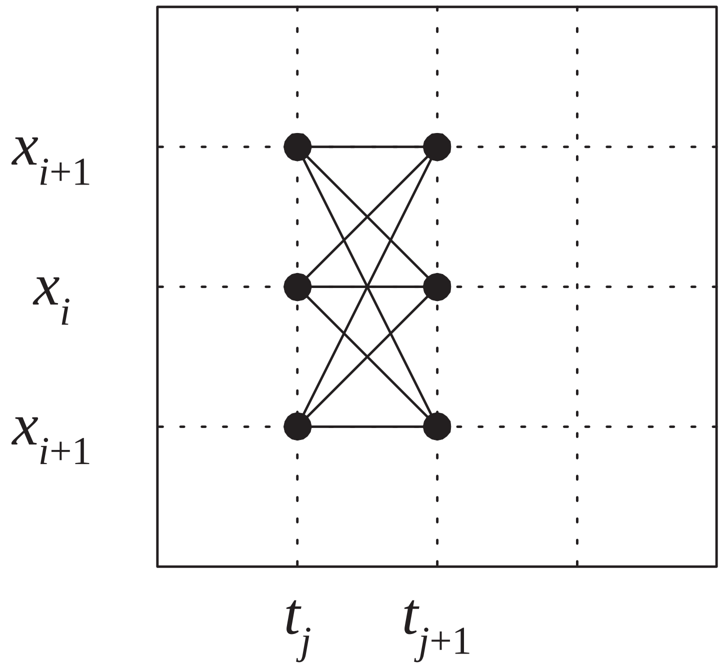
$$\frac{\theta_{i,j+1} - \theta_{i,j}}{\Delta t} = \frac{1}{2}\left(D\,\frac{\theta_{i+1,j} - 2\theta_{i,j} + \theta_{i-1,j}}{(\Delta x)^2} + D\,\frac{\theta_{i+1,j+1} - 2\theta_{i,j+1} + \theta_{i-1,j+1}}{(\Delta x)^2} \right).$$

- 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 (94) on p. 685

- See text.
- Brennan and Schwartz (1978) analyze the stability of the implicit method.

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 (Dongarra & Sullivan, 2000).

The Big Idea

- Assume $X_1, X_2, \ldots, X_n$ have a joint distribution.
- $\theta \stackrel{\Delta}{=} E[\, g(X_1, X_2, \ldots, X_n)\,]$ 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 $(X_1, X_2, \ldots, X_n)$.
- Output $\overline{Y} \stackrel{\Delta}{=} (1/N) \sum_{i=1}^{N} Y_i$, where

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

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

$$Y \stackrel{\Delta}{=} g(X_1, X_2, \ldots, X_n).$$

- 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 financial derivative only requires discrete sampling along time, such as the *discrete* barrier option.

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.[a]
- 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.

[a]The Berry-Esseen theorem.

Accuracy and Number of Replications (concluded)

- In contrast, classic numerical integration schemes have an error bound of $O(N^{-c/n})$ for some constant $c > 0$.
- 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 is more efficient than alternative procedures for multivariate derivatives for n large.

Monte Carlo Option Pricing

- For the pricing of European options on a dividend-paying stock, we may proceed as follows.
- Assume
$$\frac{dS}{S} = \mu\, dt + \sigma\, dW.$$
- Stock prices $S_1, S_2, S_3, \ldots$ at times $\Delta t, 2\Delta t, 3\Delta t, \ldots$ can be generated via

$$\begin{aligned} & S_{i+1} \\ = \; & S_i e^{(\mu-\sigma^2/2)\,\Delta t+\sigma\sqrt{\Delta t}\;\xi}, \quad \xi \sim N(0,1), \end{aligned} \qquad (125)$$

by Eq. (87) on p. 619.

Monte Carlo Option Pricing (continued)

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

$$S_{i+1} = S_i + S_i \mu\, \Delta t + S_i \sigma \sqrt{\Delta t}\, \xi.$$

- 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 (continued)

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

$C := 0$; {Accumulated terminal option value.}

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

  $P := S \times e^{(r - \sigma^2/2)\, T + \sigma \sqrt{T}\, \xi}$, $\xi \sim N(0, 1)$;

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

end for

`return` Ce^{-rT}/N;

Monte Carlo Option Pricing (concluded)

Pricing Asian options is also easy.

$C := 0;$
for $i = 1, 2, 3, \ldots, N$ **do**
  $P := S;\ M := S;$
  for $j = 1, 2, 3, \ldots, n$ **do**
    $P := P \times e^{(r - \sigma^2/2)(T/n) + \sigma \sqrt{T/n}\, \xi};$
    $M := M + P;$
  end for
  $C := C + \max(M/(n+1) - X, 0);$
end for
`return` $Ce^{-rT}/N;$

How about American Options?

- Standard Monte Carlo simulation is inappropriate for American options because of early exercise.
  - Given a sample path $S_0, S_1, \ldots, S_n$, how to decide which S_i is an early-exercise point?
  - What is the option price at each S_i if the option is not exercised?
- It is difficult to determine the early-exercise point based on one single path.
- But Monte Carlo simulation can be modified to price American options with small biases.[a]

---

[a] Longstaff & Schwartz (2001). See pp. 931ff.

Obtaining Profit and Loss of Delta Hedge[a]

- Profit and loss of delta hedge should be calculated under the real-world probability measure.[b]
- So stock prices should be sampled from

$$\frac{dS}{S} = \mu\, dt + \sigma\, dW.$$

- Suppose backward induction on a tree under the risk-neutral measure is performed for the delta.[c]

---

[a] Contributed by Mr. Lu, Zheng-Liang (D00922011) on August 12, 2021.

[b] Recall p. 711.

[c] Because, say, no closed-form formulas are available for the delta.

Obtaining Profit and Loss of Delta Hedge (concluded)

- Note that one needs a delta per stock price.
- So Nn trees are needed for the distribution of the profit and loss from N paths with $n + 1$ stock prices per path.
- These are a lot of trees!
- How to do it efficiently?[a]

---

[a] Hint: Eq. (43) on p. 299.

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.
- This is also called the bump-and-revalue method.

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 gamma.[a]

[a]For multivariate derivatives.

Problems with the Bump-and-Revalue Method

- Consider the binary option with payoff

$$\begin{cases} 1, & \text{if } S(T) > X, \\ 0, & \text{otherwise.} \end{cases}$$

- Then

$$P(S+\epsilon)-P(S-\epsilon) = \begin{cases} 1, & \text{if } S+\epsilon > X \text{ and } S-\epsilon < X, \\ 0, & \text{otherwise.} \end{cases}$$

- So the finite-difference estimate per run for the (undiscounted) delta is 0 or $O(1/\epsilon)$.
- This means high variance.

Problems with the Bump-and-Revalue Method (concluded)

- The price of the binary option equals

$$e^{-r\tau} N(x - \sigma\sqrt{\tau}).$$

  - It equals *minus* the derivative of the European call with respect to X.
  - It also equals $X\tau$ times the rho of a European call (p. 362).

- Its delta is

$$\frac{N'\left(x - \sigma\sqrt{\tau}\right)}{S\sigma\sqrt{\tau}}.$$

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 gamma $\partial^2 P(S_1, S_2, \ldots)/(\partial S_1 \partial S_2)$ 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 (continued)

- 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.[a]

[a] Aït-Sahalia & Lo (1998); Bondarenko (2003).

Gamma (continued)

- In general, suppose (in some sense)

$$\frac{\partial^i}{\partial \theta^i} e^{-r\tau} E[\, P(S)\,] = e^{-r\tau} E\left[\, \frac{\partial^i P(S)}{\partial \theta^i} \,\right]$$

  holds for all $i > 0$, where θ is a parameter of interest.[a]

  – A common requirement is Lipschitz continuity.[b]

- Then Greeks become integrals.

- As a result, we avoid ϵ, finite differences, and resimulation.

---

[a]The $\partial^i P(S)/\partial\theta^i$ within $E[\,\cdot\,]$ may not be partial differentiation in the classic sense.

[b]Broadie & Glasserman (1996).

Gamma (continued)

- This is indeed possible for a broad class of payoff functions.[a]
  - Roughly speaking, any payoff function that is equal to a sum of products of differentiable functions and indicator functions with the right kind of support.
  - For example, the payoff of a call is

$$\max(S(T)-X,0)=(S(T)-X)I_{\{\,S(T)-X\geq 0\,\}}.$$

  - The results are too technical to cover here (see next page).

[a] Teng (R91723054) (2004); Lyuu & Teng (R91723054) (2011).

Gamma (continued)

- Suppose $h(\theta, x) \in \mathcal{H}$ with pdf $f(x)$ for x and $g_j(\theta, x) \in \mathcal{G}$ for $j \in \mathcal{B}$, a finite set of natural numbers.
- Then

$$
\begin{aligned}
& \frac{\partial}{\partial \theta} \int_{\Re} h(\theta, x) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\theta,x)>0\}}(x)\, f(x)\, dx \\
= \quad & \int_{\Re} h_\theta(\theta, x) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\theta,x)>0\}}(x)\, f(x)\, dx \\
& + \sum_{l \in \mathcal{B}} \left[h(\theta, x) J_l(\theta, x) \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_j(\theta, x)>0\}}(x)\, f(x) \right]_{x=\chi_l(\theta)},
\end{aligned}
$$

where

$$
J_l(\theta, x) = \operatorname{sign}\left(\frac{\partial g_l(\theta, x)}{\partial x_k} \right) \frac{\partial g_l(\theta, x)/\partial \theta}{\partial g_l(\theta, x)/\partial x} \text{ for } l \in \mathcal{B}.
$$

Gamma (concluded)

- Similar results have been derived for Levy processes.[a]
- Formulas are also recently obtained for credit derivatives.[b]
- In queueing networks, this is called infinitesimal perturbation analysis (IPA).[c]

---

[a] Lyuu, Teng (R91723054), & S. Wang (2013).
[b] Lyuu, Teng (R91723054), Tseng, & S. Wang (2014, 2019).
[c] Cao (1985); Y. C. Ho & Cao (1985).

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 $\max(S(t_n) - X, 0)$.
- Repeat these steps and average the payoffs for a Monte Carlo estimate.

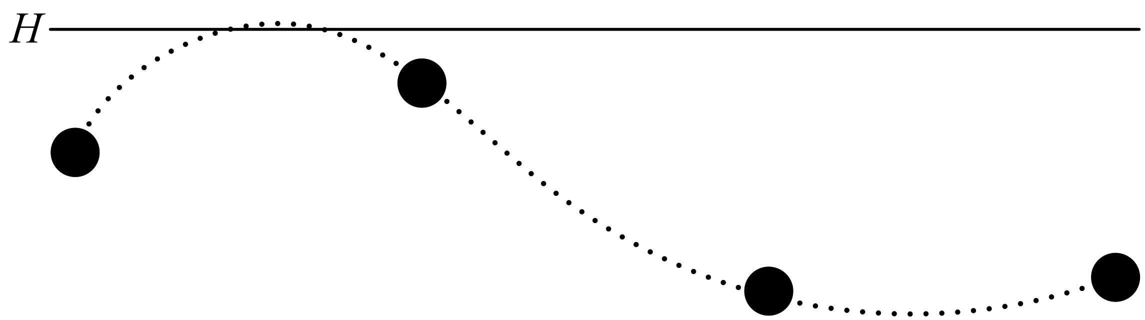
```
C := 0;
for i = 1, 2, 3, ..., N do
    P := S; hit := 0;
    for j = 1, 2, 3, ..., n do
        P := P × e^{(r − σ^2/2)(T/n) + σ sqrt((T/n)) ξ}; {By Eq. (125) on p. 863.}
        if P ≥ H then
            hit := 1;
            break;
        end if
    end for
    if hit = 0 then
        C := C + max(P − X, 0);
    end if
end for
return C e^{−rT} / N;
```

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

- This estimate is biased.[a]
  - 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).
  - Hence the knock-out probability is underestimated.

---

[a]Shevchenko (2003).


H

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

- The bias can be lowered by increasing the number of observations along the sample path.
  - For trees, the knock-out probability may *decrease* as the number of time steps is increased.
- 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 which can be calculated efficiently.
- 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.
  - Formally,

$$p \stackrel{\Delta}{=} \text{Prob}[\, S(t) < H, 0 \leq t \leq T \mid S(t_0), S(t_1), \ldots, S(t_n) \,].$$

- This methodology is called the Brownian bridge approach.

Brownian Bridge Approach to Pricing Barrier Options (continued)

- As a barrier is not hit over a time interval if and only if the maximum stock price over that period is at most H,

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

- 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 22 *Assume S follows $dS/S = \mu\, dt + \sigma\, dW$ and define*[a]

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

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

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

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

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

[a] Here, Δt is an arbitrary positive real number.

Brownian Bridge Approach to Pricing Barrier Options (continued)

- Lemma 22 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[a] call, choose $n = 1$.
- As a result,

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

[a]So $S(0) < H$ by definition.

Brownian Bridge Approach to Pricing Barrier Options (continued)

The following algorithm works for up-and-out *and* down-and-out calls.

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

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 $(t_i, t_{i+1}]$, $0 \leq i < m$.
- This option contains m barriers.
- Multiply the probabilities for the m time intervals to obtain the desired probability adjustment term.

Pricing Barrier Options without Brownian Bridge

- Let T_h denote the amount of time for a process X_t to hit h for the *first* time.
- It is called the first passage time or the first hitting time.
- Suppose X_t is a (μ, σ) Brownian motion:

$$dX_t = \mu\, dt + \sigma\, dW_t, \quad t \geq 0.$$

Pricing Barrier Options without Brownian Bridge (continued)

- The first passage time T_h follows the inverse Gaussian (IG) distribution with probability density function:[a]

$$\frac{|\,h - X(0)\,|}{\sigma t^{3/2}\sqrt{2\pi}}\, e^{-(h-X(0)-\mu x)^2/(2\sigma^2 x)}.$$

- For pricing a barrier option with barrier H by simulation, the density function becomes

$$\frac{|\,\ln(H/S(0))\,|}{\sigma t^{3/2}\sqrt{2\pi}}\, e^{-\left[\,\ln(H/S(0))-(r-\sigma^2/2)\,x\,\right]^2/(2\sigma^2 x)}.$$

---

[a] A. N. Borodin & Salminen (1996), with Laplace transform $E[\,e^{-\lambda T_h}\,] = e^{-|\,h-X(0)\,|\sqrt{2\lambda}}$, $\lambda > 0$.

Pricing Barrier Options without Brownian Bridge (concluded)

- Draw an x from this distribution.[a]
- If $x > T$, a knock-in option fails to knock in, whereas a knock-out option does not knock out.
- If $x \leq T$, the opposite is true.
- If the barrier option survives at maturity T, then draw an $S(T)$ to calculate its payoff.
- Repeat the above process and average the discounted payoff.

---

[a]The IG distribution can be very efficiently sampled (Michael, Schucany, & Haas, 1976).

Brownian Bridge Approach to Pricing Lookback Options[a]

- By Lemma 22(1) (p. 888),

$$F_{\max}(y) \stackrel{\Delta}{=} \text{Prob}\left[\max_{0\le t\le T} S(t) < y \,|\, S(0), S(T)\right] = 1 - \exp\left[-\frac{2\ln(y/S(0))\ln(y/S(T))}{\sigma^2 T}\right].$$

- So $F_{\max}$ is the conditional distribution function of the maximum stock price.

[a]El Babsiri & Noel (1998).

Brownian Bridge Approach to Pricing Lookback Options (continued)

- A random variable with that distribution can be generated by $F_{\max}^{-1}(x)$, where x is uniformly distributed over $(0, 1)$.[a]
- Note that

$$x = 1 - \exp\left[-\frac{2 \ln(y/S(0)) \ln(y/S(T))}{\sigma^2 T} \right].$$

[a]This is called the inverse-transform technique (see p. 259 of the textbook).

Brownian Bridge Approach to Pricing Lookback Options (continued)

- Equivalently,

$$
\begin{aligned}
& \ln(1-x) \\
= \quad & -\frac{2\ln(y/S(0))\ln(y/S(T))}{\sigma^2 T} \\
= \quad & -\frac{2}{\sigma^2 T}\{\,[\,\ln(y)-\ln S(0)\,]\,[\,\ln(y)-\ln S(T)\,]\,\}.
\end{aligned}
$$

Brownian Bridge Approach to Pricing Lookback Options (continued)

- There are two solutions for $\ln y$.
- But only one is consistent with $y \geq \max(S(0), S(T))$:

$$\begin{aligned} & \ln y \\ = \quad & \frac{\ln(S(0)\, S(T)) + \sqrt{\left(\ln \frac{S(T)}{S(0)}\right)^2 - 2\sigma^2 T \ln(1-x)}}{2}. \end{aligned}$$

Brownian Bridge Approach to Pricing Lookback Options (concluded)

The following algorithm works for the lookback put on the maximum.

$C := 0;$
for $i = 1, 2, 3, \ldots, N$ **do**
  $P := S \times e^{(r-q-\sigma^2/2)\,T+\sigma\sqrt{T}\;\xi(\,)}$; {By Eq. (125) on p. 863.}
  $Y := \exp\left[\frac{\ln(SP)+\sqrt{\left(\ln\frac{P}{S}\right)^2-2\sigma^2 T\ln[\,1-U(0,1)\,]}}{2}\right];$
  $C := C + (Y - P);$
end for
`return` $Ce^{-rT}/N;$

Pricing Lookback Options without Brownian Bridge

- Suppose we do not draw $S(T)$ in simulation.
- Now, the distribution function of the maximum logarithmic stock price is[a]

$$
\begin{aligned}
& \text{Prob}\left[\max_{0 \leq t \leq T} \ln \frac{S(t)}{S(0)} < y \right] \\
= \quad & 1 - N\left(\frac{-y + \left(r - q - \frac{\sigma^2}{2}\right) T}{\sigma\sqrt{T}} \right) \\
& - e^{\frac{2y\left(r - q - \frac{\sigma^2}{2}\right)}{\sigma^2}} N\left(\frac{-y - \left(r - q - \frac{\sigma^2}{2}\right) T}{\sigma\sqrt{T}} \right), \quad y \geq 0.
\end{aligned}
$$

- The inverse of that is much harder to calculate.

[a] A. N. Borodin & Salminen (1996).

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 work in this manner are called variance-reduction techniques.
- Such techniques become practical when the added costs are outweighed by the reduction in sampling.

Variance Reduction: Antithetic Variates

- We want to estimate $E[\,g(X_1, X_2, \ldots, X_n)\,]$.
- Let Y_1 and Y_2 be random variables with the same distribution as $g(X_1, X_2, \ldots, X_n)$.
- Then

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

  - $\mathrm{Var}[\,Y_1\,]/2$ is the variance of the Monte Carlo method with two *independent* replications.
- The variance $\mathrm{Var}[\,(Y_1 + Y_2)/2\,]$ is smaller than $\mathrm{Var}[\,Y_1\,]/2$ 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 first path's random numbers.
- 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 $X_1, X_2, \ldots, X_n$ on the sample path.
- Suppose one simulation run has realizations $\xi_1, \xi_2, \ldots, \xi_n$ for the normally distributed fluctuation term ξ.
- This generates samples $x_1, x_2, \ldots, x_n$.
- The first estimate is then $g(\boldsymbol{x})$, where $\boldsymbol{x} \stackrel{\Delta}{=} (x_1, x_2 \ldots, x_n)$.

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 $\boldsymbol{x}' \stackrel{\Delta}{=} (x'_1, x'_2 \ldots, x'_n)$ from $-\xi_1, -\xi_2, \ldots, -\xi_n$.
- Compute $g(\boldsymbol{x}')$.
- Output $(g(\boldsymbol{x}) + g(\boldsymbol{x}'))/2$.
- Repeat the above steps.

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
$$\text{Var}[\,E[\,X \mid Z\,]\,] \leq \text{Var}[\,X\,],$$
$E[\,X \mid Z\,]$ has a smaller variance than observing X directly.
- First, obtain a random observation z on Z.
- Then calculate $E[\,X \mid Z = z\,]$ as our estimate.
  - There is no need to resort to simulation in computing $E[\,X \mid Z = z\,]$.
- The procedure is repeated to reduce the variance.