Merton's Jump-Diffusion Model

- Empirically, stock returns tend to have fat tails, inconsistent with the Black-Scholes model's assumptions.
- Stochastic volatility and jump processes have been proposed to address this problem.
- Merton's jump-diffusion model is our focus.[a]

[a] Merton (1976).

Merton's Jump-Diffusion Model (continued)

- This model superimposes a jump component on a diffusion component.
- The diffusion component is the familiar geometric Brownian motion.
- The jump component is composed of lognormal jumps driven by a Poisson process.
  - It models the sudden changes in the stock price because of the arrival of important new information.

Merton's Jump-Diffusion Model (continued)

- Let S_t be the stock price at time t.
- The risk-neutral jump-diffusion process for the stock price follows

$$\frac{dS_t}{S_t} = (r - \lambda \bar{k})\, dt + \sigma\, dW_t + k\, dq_t. \tag{81}$$

- Above, σ denotes the volatility of the diffusion component.

Merton's Jump-Diffusion Model (continued)

- The jump event is governed by a compound Poisson process q_t with intensity λ, where k denotes the magnitude of the *random* jump.
  - The distribution of k obeys

$$\ln(1+k) \sim N\left(\gamma, \delta^2\right)$$

  with mean $\bar{k} \equiv E\left(k\right) = e^{\gamma+\delta^2/2} - 1$.

- The model with $\lambda = 0$ reduces to the Black-Scholes model.

Merton's Jump-Diffusion Model (continued)

- The solution to Eq. (81) on p. 700 is

$$S_t = S_0 e^{(r-\lambda\bar{k}-\sigma^2/2)\,t+\sigma W_t} U(n(t)), \tag{82}$$

where

$$U(n(t)) = \prod_{i=0}^{n(t)} (1+k_i)\,.$$

  - k_i is the magnitude of the ith jump with $\ln(1+k_i) \sim N(\gamma, \delta^2)$.
  - $k_0 = 0$.
  - $n(t)$ is a Poisson process with intensity λ.

Merton's Jump-Diffusion Model (concluded)

- Recall that $n(t)$ denotes the number of jumps that occur up to time t.
- As $k > -1$, stock prices will stay positive.
- The geometric Brownian motion, the lognormal jumps, and the Poisson process are assumed to be independent.

Tree for Merton's Jump-Diffusion Model[a]

- Define the S-logarithmic return of the stock price S' as

$$\ln(S'/S).$$

- Define the logarithmic distance between stock prices S' and S as

$$|\ln(S') - \ln(S)| = |\ln(S'/S)|.$$

[a] Dai (R86526008, D8852600), Wang (F95922018), Lyuu, and Liu (2010).

Tree for Merton's Jump-Diffusion Model (continued)

- Take the logarithm of Eq. (82) on p. 702:

$$M_t \equiv \ln\left(\frac{S_t}{S_0}\right) = X_t + Y_t, \tag{83}$$

  where

$$X_t \equiv \left(r - \lambda\bar{k} - \sigma^2/2\right) t + \sigma W_t, \tag{84}$$

$$Y_t \equiv \sum_{i=0}^{n(t)} \ln\left(1 + k_i\right). \tag{85}$$

- It decomposes the S_0-logarithmic return of S_t into the diffusion component X_t and the jump component Y_t.

Tree for Merton's Jump-Diffusion Model (continued)

- Motivated by decomposition (83) on p. 705, the tree construction divides each period into a diffusion phase followed by a jump phase.
- In the diffusion phase, X_t is approximated by the BOPM.
- Hence X_t can make an up move to $X_t + \sigma\sqrt{\Delta t}$ with probability p_u or a down move to $X_t - \sigma\sqrt{\Delta t}$ with probability p_d.

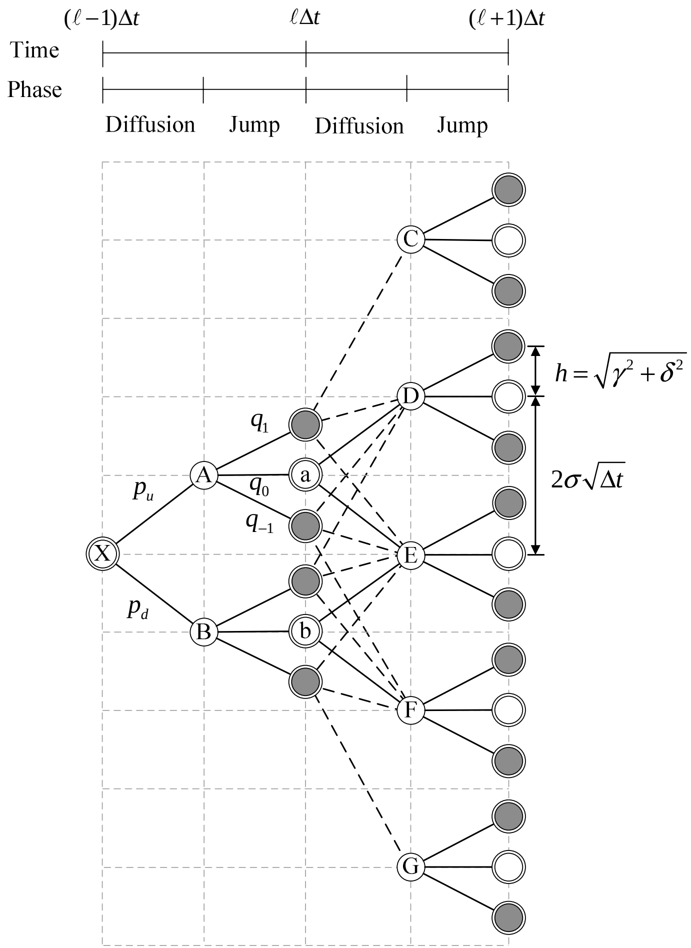
Tree for Merton's Jump-Diffusion Model (continued)

- According to BOPM,

$$
\begin{aligned}
p_u &= \frac{e^{\mu \Delta t} - d}{u - d}, \\
p_d &= 1 - p_u,
\end{aligned}
$$

  except that $\mu = r - \lambda \bar{k}$ here.

- The diffusion component gives rise to diffusion nodes.
- They are spaced at $2\sigma\sqrt{\Delta t}$ apart such as the white nodes A, B, C, D, E, F, and G on p. 708.



White nodes are diffusion nodes. Gray nodes are jump nodes. In the diffusion phase, the solid black lines denote the binomial structure of BOPM, whereas the dashed lines denote the trinomial structure. Here m is set to one here for simplicity. Only the double-circled nodes will remain after the construction. Note that a and b are diffusion nodes because no jump occurs in the jump phase.

Tree for Merton's Jump-Diffusion Model (concluded)

- In the jump phase, $Y_{t+\Delta t}$ is approximated by moves from *each* diffusion node to $2m$ jump nodes that match the first $2m$ moments of the lognormal jump.
- The m jump nodes above the diffusion node are spaced at h apart.
- The same holds for the m jump nodes below the diffusion node.
- The gray nodes at time $\ell \Delta t$ on p. 708 are jump nodes.
- After some work, the size of the tree is $O(n^{2.5})$.

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.
- Option on the best of two risky assets and cash has a terminal payoff of $\max(S_1(\tau), S_2(\tau), X)$.

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	$\max((S_1(\tau), \ldots, S_k(\tau))/k - X, 0)$
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 and 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
$$\begin{aligned} M_i &\equiv e^{r\Delta t}, \\ V_i &\equiv M_i^2(e^{\sigma_i^2 \Delta t} - 1). \end{aligned}$$
  - $S_i M_i$ 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, and 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 five distinct outcomes.

Correlated Trinomial Model (continued)

- The five-point probability distribution of the asset prices is (as usual, we impose $u_i d_i = 1$)

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

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

$$
\begin{aligned}
S_1^2 V_1 &= (p_1 + p_2)((S_1 u_1)^2 - (S_1 M_1)^2) + p_5 (S_1^2 - (S_1 M_1)^2) \\
&\quad + (p_3 + p_4)((S_1 d_1)^2 - (S_1 M_1)^2), \\
S_2^2 V_2 &= (p_1 + p_4)((S_2 u_2)^2 - (S_2 M_2)^2) + p_5 (S_2^2 - (S_2 M_2)^2) \\
&\quad + (p_2 + p_3)((S_2 d_2)^2 - (S_2 M_2)^2), \\
S_1 S_2 R &= (p_1 u_1 u_2 + p_2 u_1 d_2 + p_3 d_1 d_2 + p_4 d_1 u_2 + p_5)\, S_1 S_2.
\end{aligned}
$$

- The solutions are complex (see text).

Correlated Trinomial Model Simplified[a]

- Let $\mu'_i \equiv r - \sigma_i^2/2$ and $u_i \equiv e^{\lambda \sigma_i \sqrt{\Delta t}}$ for $i = 1, 2$.
- The following simpler scheme is 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, and Shefrin (1989).

Correlated Trinomial Model Simplified (continued)

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

$$
\begin{aligned}
-1+\lambda\sqrt{\Delta t}\left|\frac{\mu_1'}{\sigma_1}+\frac{\mu_2'}{\sigma_2}\right| \leq \quad & \rho \quad \leq 1-\lambda\sqrt{\Delta t}\left|\frac{\mu_1'}{\sigma_1}-\frac{\mu_2'}{\sigma_2}\right|, && (86)\\
1 \leq \quad & \lambda && (87)
\end{aligned}
$$

- 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] Kao (R98922093) (2011) and Kao (R98922093), Lyuu, and Wen (D94922003) (2014).

Correlated Trinomial Model Simplified (concluded)

- 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 Chang (B89704039, R93922034), Hsu (R7526001, D89922012), and Lyuu (2006) and Kao (R98922093), Lyuu and Wen (D94922003) (2014) for solutions.

[b] See Kao (R98922093), Lyuu, and Wen (D94922003) (2014) for one.

[c] Hull and White (1990) and Dai (R86526008, D8852600), Lyuu, and Wang (F95922018) (2012).

Extrapolation

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

$$f(n) = f + \frac{c}{n} + o\left(\frac{1}{n}\right). \tag{88}$$

- 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,

$$\begin{aligned} f(n_1) &= f + \frac{c}{n_1}, \\ f(n_2) &= f + \frac{c}{n_2}. \end{aligned}$$

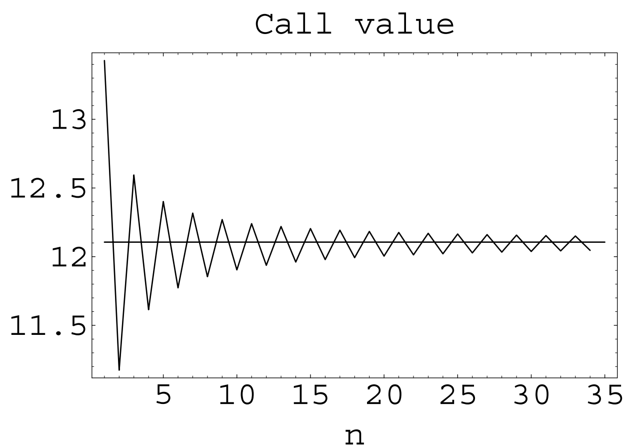
- A better approximation to the desired f is

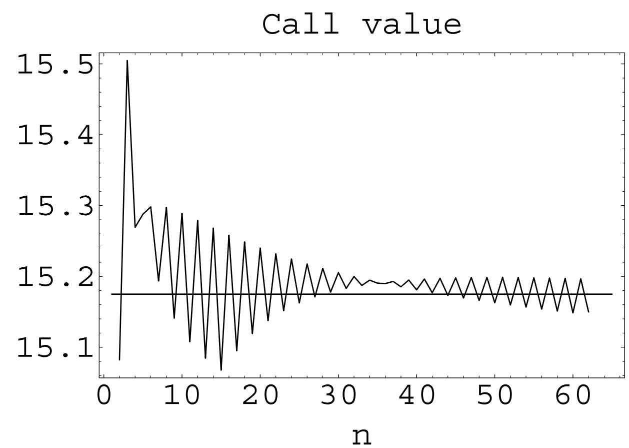
$$f = \frac{n_1 f(n_1) - n_2 f(n_2)}{n_1 - n_2}. \tag{89}$$

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

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 $1/n$, consistent with Eq. (88) on p. 720.
- But the plots on p. 282 (redrawn on next page) demonstrate that convergence to the true option value oscillates with n.
- Extrapolation is inapplicable at this stage.


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

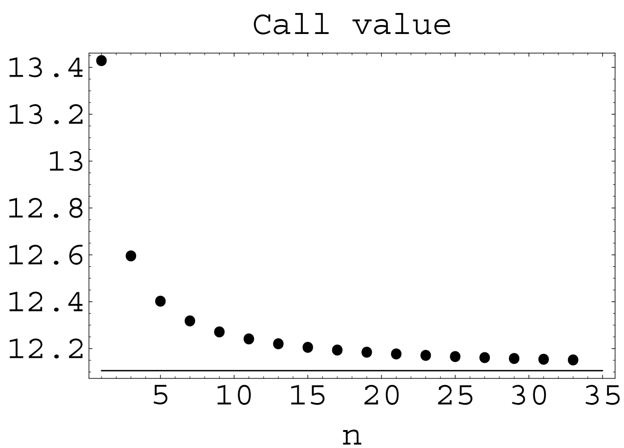

Call value
15.5
15.4
15.3
15.2
15.1
0 10 20 30 40 50 60
n

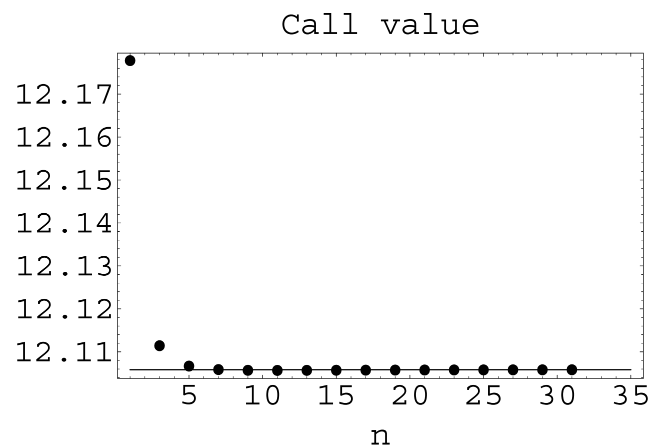
Improving BOPM with Extrapolation (concluded)

- Take the at-the-money option in the left plot on p. 723.
- The sequence with odd n turns out to be monotonic and smooth (see the left plot on p. 725).[a]
- Apply extrapolation (89) on p. 721 with $n_2 = n_1 + 2$, where n_1 is odd.
- Result is shown in the right plot on p. 725.
- The convergence rate is amazing.
- See Exercise 9.3.8 of the text (p. 111) for ideas in the general case.

---

[a]This can be proved; see Chang and Palmer (2007).


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

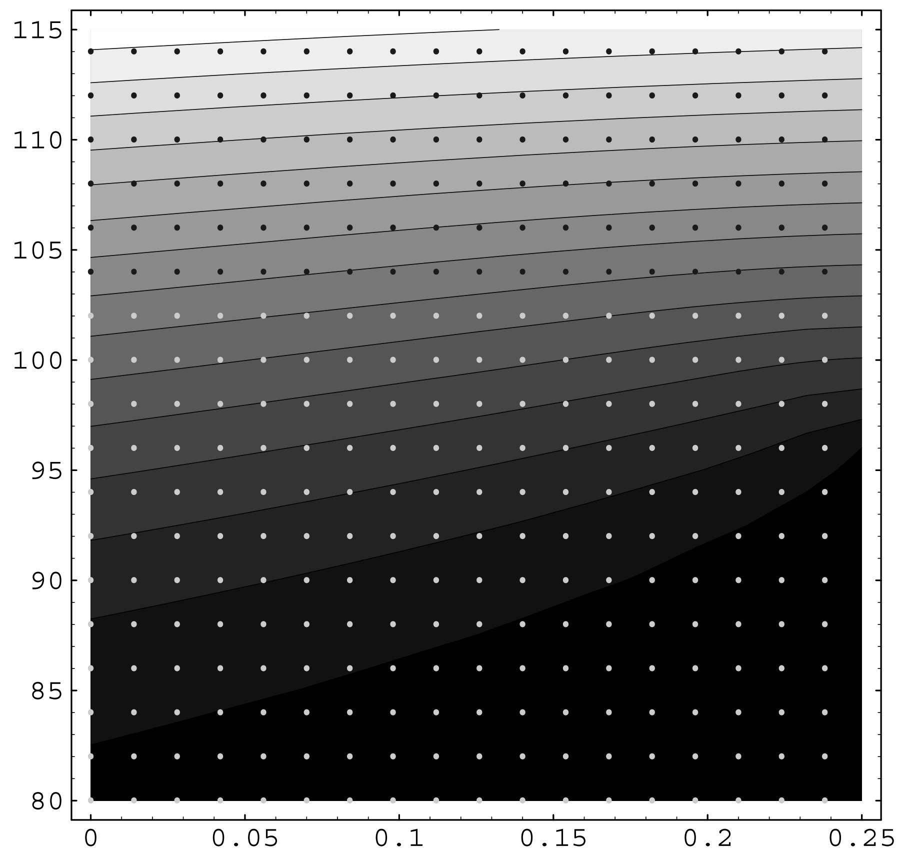

Call value
12.17
12.16
12.15
12.14
12.13
12.12
12.11
5 10 15 20 25 30 35
n

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. 729).
- 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 $\partial^2\theta/\partial x^2 + \partial^2\theta/\partial y^2 = -\rho(x, y)$.
- 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

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

Example: Poisson's Equation (concluded)

- In the above, $\Delta x \equiv x_i - x_{i-1}$ and $\Delta y \equiv y_j - y_{j-1}$ 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

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

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

$$\left.\frac{\partial\theta(x,t)}{\partial t}\right|_{t=t_j} = \frac{\theta(x,t_{j+1}) - \theta(x,t_j)}{\Delta t} + \cdots, \tag{90}$$

$$\left.\frac{\partial^2\theta(x,t)}{\partial x^2}\right|_{x=x_i} = \frac{\theta(x_{i+1},t) - 2\theta(x_i,t) + \theta(x_{i-1},t)}{(\Delta x)^2} + \cdots. \tag{91}$$

Explicit Methods (continued)

- Next, assemble Eqs. (90) and (91) 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. (91), we might as well use x_i for x in Eq. (90).
- Two choices are possible for t in Eq. (91).
- The first choice uses $t = t_j$ to yield the following finite-difference equation,

$$\frac{\theta_{i,j+1} - \theta_{i,j}}{\Delta t} = D\,\frac{\theta_{i+1,j} - 2\theta_{i,j} + \theta_{i-1,j}}{(\Delta x)^2}. \qquad (92)$$

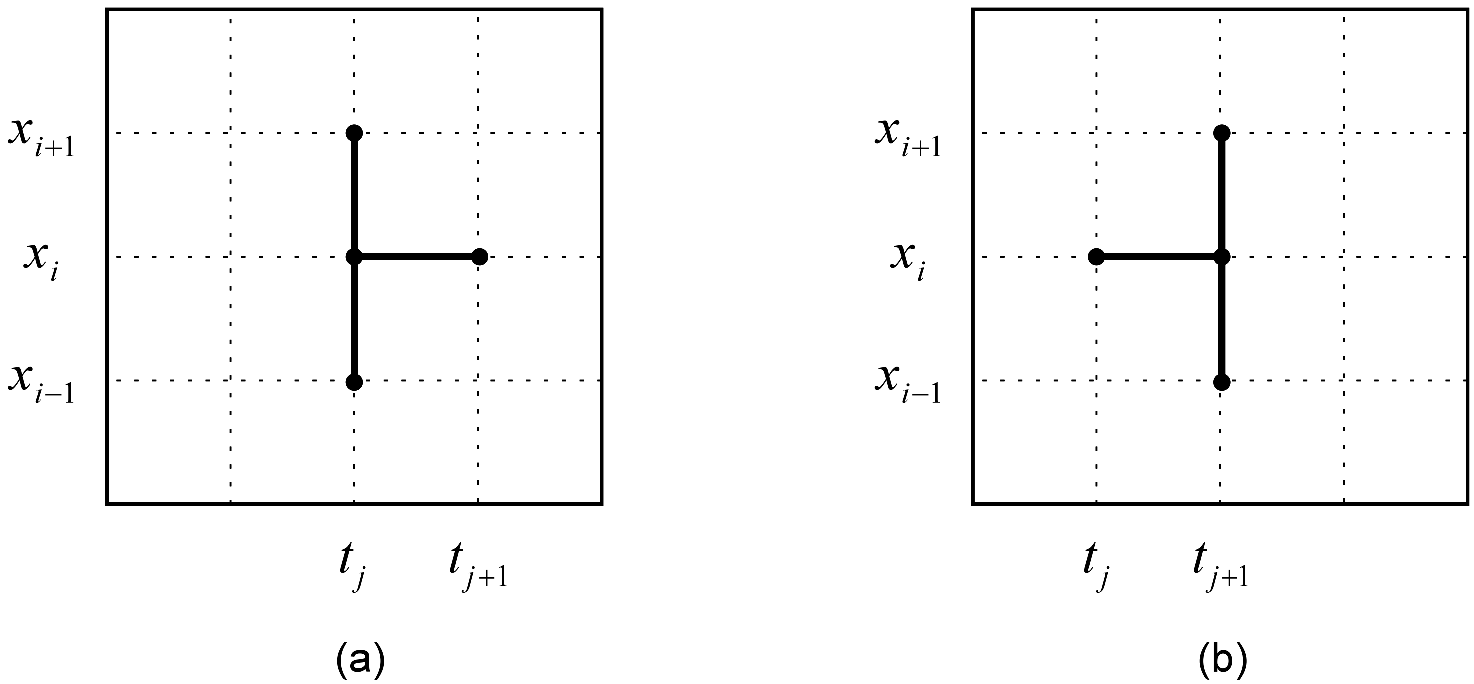
Explicit Methods (continued)

- 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. (92) on p. 733 as

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

- 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



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 that

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

- 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!
- The freedom in choosing Δx corresponds to similar freedom in the construction of trinomial trees.

Implicit Methods

- Suppose we use $t = t_{j+1}$ in Eq. (91) on p. 732 instead.
- The finite-difference equation becomes

$$\frac{\theta_{i,j+1} - \theta_{i,j}}{\Delta t} = D\, \frac{\theta_{i+1,j+1} - 2\theta_{i,j+1} + \theta_{i-1,j+1}}{(\Delta x)^2}. \tag{93}$$

- 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. 735.

Implicit Methods (continued)

- Equation (93) 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)

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

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

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 (92) on p. 733 and implicit method (93) on p. 739:

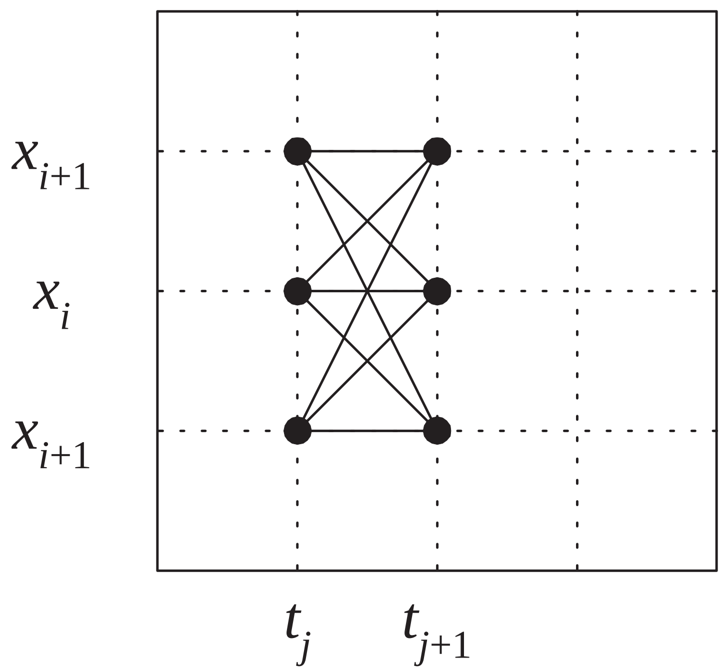
$$\frac{\theta_{i,j+1} - \theta_{i,j}}{\Delta t} = \frac{1}{2}\left(D\,\frac{\theta_{i+1,j} - 2\theta_{i,j} + \theta_{i-1,j}}{(\Delta x)^2} + D\,\frac{\theta_{i+1,j+1} - 2\theta_{i,j+1} + \theta_{i-1,j+1}}{(\Delta x)^2} \right).$$

- 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 (65) on p. 583

- 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 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 financial derivative only requires discrete sampling along the time dimension, 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 $\text{Var}[\,\overline{Y}\,] = \text{Var}[\,Y\,]/N$.
- 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 $c > 0$.
  - 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 is more efficient than alternative procedures for multivariate derivatives when n is 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

$$S_{i+1} = S_i e^{(\mu - \sigma^2/2)\, \Delta t + \sigma \sqrt{\Delta t}\; \xi}, \quad \xi \sim N(0,1). \tag{94}$$

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 (why?).
- 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 (pp. 807ff).[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.
- 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 gammas (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

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

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

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 $\partial^2 P(S_1, S_2, \dots)/(\partial S_1 \partial S_2)$ 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 (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.

Gamma (continued)

- In general, suppose

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

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

- Then formulas for the Greeks become integrals.
- As a result, we avoid ϵ, finite differences, and resimulation.

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) and Lyuu and 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

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

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), and Wang (2013).
[b] Lyuu, Teng (R91723054), and Tzeng (2014).
[c] Cao (1985); Ho and 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)

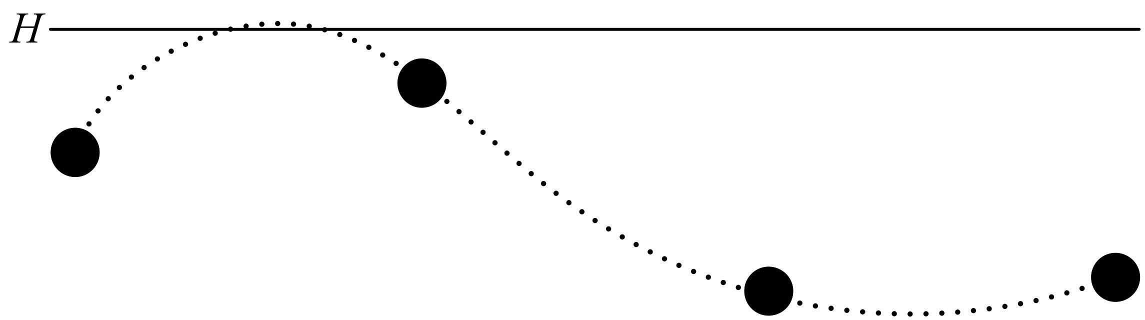
- If all of the sampled prices are below the barrier, this sample path pays $\max(S(t_n) - X, 0)$.
- Repeating these steps and averaging the payoffs yield a Monte Carlo estimate.

```
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};
        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).

[a] Shevchenko (2003).


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.