Section 5 Computational Methods

Financial Engineering · Fall 2026 · Zong-Wei Yeh · Department of Finance, CCU

Overview

Section 4 wrote the price of a claim as a risk-neutral expectation and as the cost of a replicating portfolio. This section computes those prices. §5.1 prices a two-period call with actual numbers. It takes the side of the seller, who hedges the call at every node. It then prices a three-period call. §5.2 repeats the argument in symbols, on one node and on two periods. The tree's risk-neutral probability appears there. The section then states and proves the rule for N periods. §5.3 chooses the tree parameters so that the tree imitates the Black–Scholes stock. It then proves that the tree price converges to the Black–Scholes price. §5.4 prices American options on the tree. §5.5 is Monte Carlo: the expectation estimated by simulation, with its error bars. It also covers correlated normal draws and two variance-reduction techniques. §5.6 is least-squares Monte Carlo, the simulation method for American options, worked through on eight paths with every number shown.

Two theorems are imported at book level. The first is the central limit theorem for triangular arrays, used in the convergence proof. The second is the theorem that the American recursion of §5.4 is the no-arbitrage price. One technical condition, the weak error of the Euler step in §5.5, is named and imported without proof. One convergence result, for the least-squares estimate of §5.6, is imported from a journal article named there.

The numbers of the examples, of Table 1 and of the numerical exercises are reproduced by the committed script `figures/crr_convergence.py`, and the figures by `figures/make_figures.py`. The page `web/index.html` runs every method here step by step.

5.1 Two Periods, Then Three

This subsection prices a call on a two-period tree with actual numbers, and then on a three-period tree. We take the side of the seller. The seller writes the call and holds shares against it. The number of shares is chosen to remove the risk. No arbitrage then fixes the call's price. §5.2 repeats the argument in symbols and extends it to N periods. Section 2 stated the multi-period fundamental theorem and left the binomial case to this section.

Start with the model. One risky asset (the stock) has price S at a node. A contract expires at time T, measured in years. We cut $[0, T]$ into N periods of equal length

$$\Delta t \;=\; \frac{T}{N}.$$

The symbol Δt is one object. It is not the hedge Δ times t. Over one period the stock moves to uS (up) or to dS (down), with $0 < d < u$. One riskless account pays the rate r per year. Over one period it grows by the factor $1 + r\Delta t$. Shares can be bought or shorted in any fraction, and the account lends and borrows at this rate.

The number Δt has a clear role. It is the length of one period, in years. The rate r is quoted

per year, so $r\Delta t$ is the interest for one period. For example, $r = 10\%$ and $\Delta t = \frac{1}{2}$ give the factor 1.05. The integer N counts the periods, and Δt gives their length. With T fixed, more periods mean a smaller Δt. §5.3 ties u and d to Δt and lets Δt shrink to zero. That step turns the tree into Black–Scholes.

The number $u - 1$ is the up return, $d - 1$ the down return, and $r\Delta t$ the riskless return, all per period. A shorter period has smaller moves. So u and d also depend on Δt, and §5.3 gives the formula.

With two periods there are three dates, 0, 1, 2. From date 0 the stock goes to uS_0 or dS_0. From each of those nodes it moves again, so date 2 has the prices u^2S_0, udS_0 and d^2S_0. Two paths reach the middle price: up then down, and down then up. So date 2 has three nodes, not four.

Now put numbers in. The stock is at 100 today, $u = 1.2$, $d = 0.8$, and the account pays $r = 10\%$ per year. A call with strike 100 expires at $T = 2$ years. With $N = 2$ periods, $\Delta t = 1$ year and $1 + r\Delta t = 1.1$. What is the call worth today? Nobody has told us how likely an up move is. The answer will not need it.

The method is the hedge you already know. Take a node whose two successors have known values. The seller is short one call and holds Δ shares. The seller chooses Δ so that the position is worth the same at both successors. A riskless position must earn the riskless rate, and that fixes the call's value at the node. At date 2 the values are the payoffs. So each date-1 node is a one-period problem. Once those are solved, date 0 is one too.

Example 5.1. Take $S_0 = 100$, $u = 1.2$, $d = 0.8$, $r = 10\%$ per year, $T = 2$ years, $N = 2$ (so $\Delta t = 1$), and the call with $K = 100$. The seller writes one call and holds Δ shares against it.

(a) Draw the stock tree and the call's payoffs at date 2.

(b) At each date-1 node, find the Δ that makes the seller's position riskless over the last period. What is the call worth at that node?

(c) Do the same at date 0. What is the call worth today?

Figure 1 shows the date-0 step of part (c). The seller sells the call at C and buys Δ shares, so the position costs $100\,\Delta - C$ today. At date 1 it is worth $120\,\Delta - 30$ or $80\,\Delta$. With $\Delta = \frac{3}{4}$ both numbers are 60.

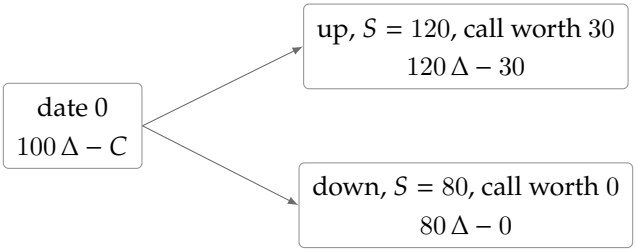


Figure 1: The seller's position at date 0: short one call, long Δ shares. The two outcomes at date 1 agree when $\Delta = \frac{3}{4}$. The position is then riskless, so it must earn the riskless rate, and that fixes C.

Two steps of the example need a reason. First, why must a sure 60 cost $60/1.1$? The position and the account are both riskless. If their returns differed, we would buy the one with the higher return and short the other. That is an arbitrage.

Second, why can the seller use the value 30 at date 1? The seller never buys the call back. The hedge needs no outside money. At date 0 the seller receives 20.4545, buys $\frac{3}{4}$ of a share for 75, and borrows the other 54.5455. After an up move the shares are worth 90 and the loan has grown to 60. So the seller has 30, exactly C_u.

The new hedge is $\frac{11}{12}$ of a share. The seller buys the extra $\frac{1}{6}$ of a share for 20 and borrows that 20. The loan is now 80, and it grows to 88 by date 2. After up-up the shares are worth 132. The seller sells them, repays 88 and pays the holder 44. After up-down the shares are worth 88. The seller sells them and repays the loan, and the call pays 0.

After a down move at date 1 the shares are worth 60. The seller sells them and repays the loan of 60. Nothing is left or owed. No outside money enters after date 0. That is self-financing in the sense of Section 2.

The work is the same at every node. Find Δ, find the sure amount, discount it one period, and subtract. With three periods the stock tree has four nodes at date 3: u^3S_0, u^2dS_0, ud^2S_0, d^3S_0. The claim is priced in three rounds: the three date-2 nodes, the two date-1 nodes, then date 0. That is six nodes.

Example 5.2. Same tree as Example 5.1: $S_0 = 100$, $u = 1.2$, $d = 0.8$, $r = 10\%$ per year, $\Delta t = 1$, $K = 100$, now with $T = 3$ years and $N = 3$. Price the European call by the seller's hedge, node by node.

5.2 Symbols, Then N Periods

This subsection repeats the argument of §5.1 in symbols: one node first, then the two-period tree, then N periods. The numbers of §5.1 do not show the pattern. The symbols show it.

Start with one node. Its stock price is S. After the move the price is uS or dS, and the claim is worth V^u or V^d there. The seller is short the claim and holds Δ shares. The position is riskless when

$$\Delta\, uS - V^u \;=\; \Delta\, dS - V^d, \qquad \text{that is,} \qquad \Delta \;=\; \frac{V^u - V^d}{(u-d)S}.$$

This is the spread of the successor values over the spread of the successor prices. In Example 5.1, $\Delta_u = (44-0)/48 = \frac{11}{12}$ and $\Delta_0 = (30-0)/40 = \frac{3}{4}$. The sure amount is

$$A \;=\; \Delta\, dS - V^d \;=\; \frac{d(V^u - V^d) - (u-d)V^d}{u-d} \;=\; \frac{dV^u - uV^d}{u-d}.$$

The position costs $\Delta S - V$ now. It pays A for sure after one period, so it must cost $A/(1+r\Delta t)$. Solve for V, and put the two terms over one denominator:

$$V \;=\; \Delta S - \frac{A}{1+r\Delta t} \;=\; \frac{V^u - V^d}{u-d} - \frac{dV^u - uV^d}{(u-d)(1+r\Delta t)} \;=\; \frac{\dfrac{1+r\Delta t-d}{u-d}\,V^u + \dfrac{u-1-r\Delta t}{u-d}\,V^d}{1+r\Delta t}.$$

Look at the two fractions that multiply V^u and V^d. They add up to 1. They depend on u, d, Δt and r, and not on the claim. Call the first one q:

$$q \;=\; \frac{1+r\Delta t-d}{u-d}, \qquad 1-q \;=\; \frac{u-1-r\Delta t}{u-d}.$$

Then

$$V \;=\; \frac{q\,V^u+(1-q)\,V^d}{1+r\Delta t}. \tag{1}$$

We call (1) the *roll-back rule*. It says: value equals the q-average of next period's values, discounted one period. The same q works at every node and for every claim. So to price a claim we need not find Δ at each node. Proposition 5.3 shows that $0 < q < 1$ exactly when the market has no arbitrage.

Check on the examples. There $1+r\Delta t = 1.1$, so $q = (1.1-0.8)/(1.2-0.8) = \frac{3}{4}$. In Example 5.1, $C_u = (\frac{3}{4}\times 44 + \frac{1}{4}\times 0)/1.1 = 30$ and $C_0 = (\frac{3}{4}\times 30 + \frac{1}{4}\times 0)/1.1 = 20.4545$. In Example 5.2, the node $S = 144$ gives $(\frac{3}{4}\times 72.8 + \frac{1}{4}\times 15.2)/1.1 = 58.4/1.1 = 53.0909$. This $\frac{3}{4}$ is not the hedge $\Delta_0 = \frac{3}{4}$ of Example 5.1. The two numbers agree by chance. The weights could not be seen in the numbers. We never used the probability of an up move, and we do not use it here.

Proposition 5.3 (The binomial specimen)**.**
In the binomial model with up factor u, down factor $d < u$, rate r and period length Δt: the market is arbitrage-free if and only if $d < 1 + r\Delta t < u$; in that case the equivalent martingale measure $\mathbf{Q}$ is unique, given at every node by the one-period probability

$$q \;=\; \frac{(1+r\Delta t)-d}{u-d},$$

and the market is complete: every claim is replicated by a self-financing strategy computed by backward induction.

The riskless return must lie strictly between the down return and the up return. Otherwise one asset dominates the other. Under that condition, exactly one reweighting of the up and down probabilities makes the discounted stock a martingale. That reweighting is q. It is not the probability of an up move in the world. It is the weight that makes prices come out as discounted averages. On the tree of Example 5.1, $q = \frac{3}{4}$.

The condition $d < 1 + r\Delta t < u$ has a direct reading. Suppose $1 + r\Delta t \le d$. Then the stock does at least as well as the account even after a down move. Borrow at r and buy the stock: the down state breaks even or better, and the up state gains. That is an arbitrage. Suppose $1 + r\Delta t \ge u$. Then the account does at least as well as the stock even after an up move. Short the stock and lend the proceeds. The up state breaks even or better, and the down state gains. Again this is an arbitrage. So q lies strictly between 0 and 1 exactly when neither asset dominates the other.

The proof writes the hedge as a portfolio (Δ, b): Δ shares and the amount b in the account. The seller's loan is the case $b < 0$, and $b = -A/(1 + r\Delta t)$.

Proof.

□

The cost at the end of the proof is the value in (1). So the portfolio that copies the claim and the seller's hedge give the same price.

Why are these weights called risk-neutral? Under them the stock's expected gross return is the riskless return. In the examples, $\frac{3}{4}(120) + \frac{1}{4}(80) = 110 = 100 \times 1.1$. The call has the same property: $\frac{3}{4}(30) + \frac{1}{4}(0) = 22.5 = \frac{225}{11} \times 1.1$. So under q every asset, the stock and any claim on it, earns the riskless rate on average. An investor who demanded no premium for risk would use these weights. That is why q is called the *risk-neutral probability*, and why pricing by the q-average is called risk-neutral valuation.

The real up probability p never enters. One could also average the two successor values with p. But then the discount rate would have to be the option's own required return, and nobody knows that number. With the risk-neutral weights we discount at the known rate r. Figure 2 shows the same tree under the two sets of weights.

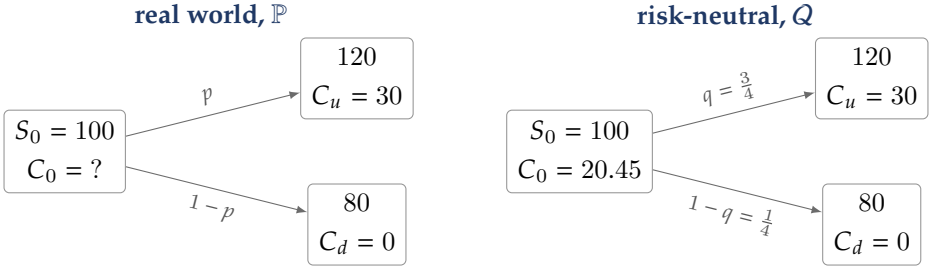


Figure 2: Date 0 of the same tree under two sets of weights. Under q the stock and the call both earn the riskless rate. The call's price is its discounted q-average, $22.5/1.1 = 20.45$.

Now the two-period tree in symbols. Write f for the claim's value at date 0, f_u and f_d for its values at date 1, and f_{uu}, f_{ud}, f_{dd} for its payoffs at date 2. The roll-back rule at the two date-1 nodes gives

$$f_u = \frac{q\, f_{uu} + (1-q)\, f_{ud}}{1 + r\Delta t}, \qquad f_d = \frac{q\, f_{ud} + (1-q)\, f_{dd}}{1 + r\Delta t},$$

and at date 0, after substituting,

$$f = \frac{q\, f_u + (1-q)\, f_d}{1 + r\Delta t} = \frac{q^2 f_{uu} + 2q(1-q)\, f_{ud} + (1-q)^2 f_{dd}}{(1 + r\Delta t)^2}.$$

The second form is a sum over the three terminal nodes. The weight q^2 is the probability of two up moves. The weight $2q(1-q)$ is the probability of one up and one down (two paths lead there). The weight $(1-q)^2$ is the probability of two down moves. And $(1 + r\Delta t)^2$ discounts two periods. Check on Example 5.1: $\big(\frac{9}{16}(44) + 2 \cdot \frac{3}{16}(0) + \frac{1}{16}(0)\big)/1.21 = 24.75/1.21 = 20.4545$.

The same two steps work for any number of periods. We apply the roll-back rule at every node. The q is always the same, because u, d, Δt and r are the same at every node. The date-0 price can also be written in one line. It is a sum over terminal nodes, and binomial coefficients count the paths. The next proposition states both for N periods.

Proposition 5.4 (Pricing on N periods)**.**
Consider the N-period binomial model with $d < 1 + r\Delta t < u$ and $q = (1 + r\Delta t - d)/(u - d)$. Write S_k for the stock price at date k, and let a claim pay $h(S_N)$ at date N.

(a) (Backward induction.) Its price V_k at every node of date k is given by $V_N = h(S_N)$ and

$$V_k \;=\; \frac{q\, V^u_{k+1} + (1-q)\, V^d_{k+1}}{1 + r\Delta t}, \qquad k = N-1, \ldots, 0,$$

where V^u_{k+1}, V^d_{k+1} are the values at the two successor nodes.

(b) (Closed form.) The date-0 price is

$$V_0 \;=\; \frac{1}{(1 + r\Delta t)^N} \sum_{j=0}^{N} \binom{N}{j} q^j (1-q)^{N-j}\, h\big(S_0\, u^j d^{\,N-j}\big).$$

(c) (Hedge.) At every node the replicating portfolio holds $\Delta = \dfrac{V^u_{k+1} - V^d_{k+1}}{(u-d)\, S_k}$ shares, and $b = V_k - \Delta S_k$ in the riskless account.

(d) (The call in two tails.) For the call, $h(S) = (S-K)^+$, let m be the smallest non-negative integer with $S_0 u^m d^{\,N-m} \geq K$. Then

$$C_0 \;=\; S_0\, \widetilde{\boldsymbol{Q}}(J \geq m) \;-\; K(1 + r\Delta t)^{-N}\, \boldsymbol{Q}(J \geq m),$$

where J is the number of up moves, $J \sim \text{Binomial}(N, q)$ under $\boldsymbol{Q}$, and $J \sim \text{Binomial}(N, \tilde{q})$ under $\widetilde{\boldsymbol{Q}}$ with $\tilde{q} = qu/(1 + r\Delta t)$. (If no such $m \leq N$ exists, the call is worth 0.)

Part (a) is the roll-back rule, applied $1 + 2 + \cdots + N$ times. Part (b) says the same price in one

formula. The number of up moves j is Binomial(N, q) under $\boldsymbol{Q}$. So the price is the discounted $\boldsymbol{Q}$-expectation of the payoff. Part (c) gives Δ. It is the spread of next period's values over the spread of next period's stock prices. Part (d) is part (b) for the call, rearranged. The call finishes in the money exactly when at least m of the N moves are up. The second term is the discounted strike times the risk-neutral probability of that event. The first term is the stock price times the probability of the same event under a second measure, $\widetilde{\boldsymbol{Q}}$. That measure weights each period by the stock's own gross return: $\tilde{q} = q \cdot u/(1 + r\Delta t)$.

Part (b) can be checked on Example 5.2 by counting paths. There are $2^3 = 8$ paths. A path with j up moves has $\boldsymbol{Q}$-probability $(\frac{3}{4})^j(\frac{1}{4})^{3-j}$. One path (up, up, up) ends at 172.8, with probability $(\frac{3}{4})^3 = \frac{27}{64}$. Three paths end at 115.2 (the down move can come first, second or third), each with probability $(\frac{3}{4})^2\frac{1}{4} = \frac{9}{64}$, together $\frac{27}{64}$. The other four paths pay nothing. Discounting three periods, $1.1^3 = 1.331$:

$$C_0 = \frac{\frac{27}{64}(72.8) + \frac{27}{64}(15.2)}{1.331} = \frac{30.7125 + 6.4125}{1.331} = \frac{37.125}{1.331} = 27.8926,$$

the value found node by node in Example 5.2.

Compare with Black–Scholes, $C = S_0\Phi(d_1) - Ke^{-rT}\Phi(d_2)$. The two tails of part (d) take the place of $\Phi(d_1)$ and $\Phi(d_2)$. Theorem 5.6 proves that they converge to them.

Proof.

□

Count the nodes. Date k has $k+1$ nodes. Only the number of up moves matters, not their order. An up-down and a down-up both land on udS_0. This is why the tree is called *recombining*. With N periods the tree has $(N+1)(N+2)/2$ nodes in total. For $N = 500$ that is $125{,}751$ nodes. A tree that did not recombine would have 2^{500} paths, and no computer could price on it. Figure 3 shows the shape.

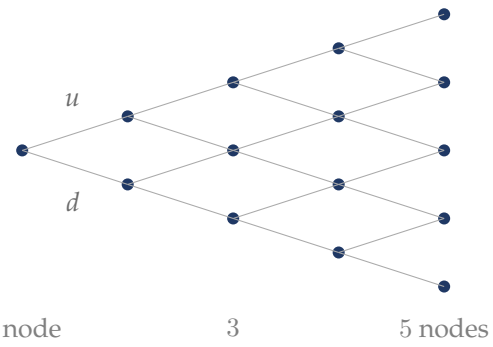


Figure 3: The recombining binomial lattice. The order of the moves does not matter, so date k has $k+1$ nodes rather than 2^k. This is what makes the tree a usable method.

5.3 Calibrating the Tree to Black–Scholes

So far u and d were given. This subsection chooses them so that the tree imitates the Black–Scholes stock of Section 4, with volatility σ and rate r. It then proves that the tree price converges to the Black–Scholes price as the number of periods grows.

Keep the N periods of length $\Delta t = T/N$ from §5.1. The tree has three unknowns: u, d and q. We choose them to match one period of the Black–Scholes stock. One period of the tree gets the same mean, and the same variance up to small terms. Under $\boldsymbol{Q}$, the GBM solution of §3.4 gives

$$S_{t+\Delta t} \;=\; S_t \exp\Big(\big(r - \tfrac{\sigma^2}{2}\big)\Delta t + \sigma\sqrt{\Delta t}\, Z\Big), \qquad Z \sim N(0,1) \text{ independent of } S_t.$$

Black–Scholes compounds continuously, so in §5.3 the growth factor is $e^{r\Delta t}$ in place of $1 + r\Delta t$. Propositions 5.3 and 5.4 use only that the account grows by a fixed factor each period. So they hold with this factor, and q is the one in Proposition 5.3 with that change. The no-arbitrage condition becomes $d < e^{r\Delta t} < u$. For the CRR tree of Proposition 5.5, this holds once Δt is small. The two factors are close: $e^{r\Delta t} = 1 + r\Delta t + O((\Delta t)^2)$. And for fixed T, $(1 + r\Delta t)^N = (1 + rT/N)^N \to e^{rT}$ as $\Delta t \to 0$.

This q fixes the mean. One more condition is $ud = 1$. It keeps the lattice centered at S_0. Then the variance fixes u to leading order. The next proposition carries this out. It is the main tool of the convergence proof.

Proposition 5.5 (Matching one period: the CRR choice).
Let $d < e^{r\Delta t} < u$ and $q = (e^{r\Delta t} - d)/(u - d)$. All moments are those of $S_{t+\Delta t}$ under $\boldsymbol{Q}$, conditional on S_t. Every $O(\cdot)$ is for $\Delta t \to 0$ with r and σ fixed.

(a) (Black–Scholes period.) The mean is $S_t e^{r\Delta t}$. The variance is $S_t^2 e^{2r\Delta t}(e^{\sigma^2\Delta t} - 1) = S_t^2\big(\sigma^2\Delta t + O((\Delta t)^2)\big)$.

(b) (Tree period.) The mean is $S_t e^{r\Delta t}$, exactly. The variance is $S_t^2\, q(1-q)(u-d)^2 = S_t^2\big[e^{r\Delta t}(u + d) - ud - e^{2r\Delta t}\big]$.

(c) (The choice.) If $ud = 1$ and the two variances, divided by S_t^2, agree up to $O((\Delta t)^2)$, then $u + 1/u = 2 + \sigma^2\Delta t + O((\Delta t)^2)$, so $\ln u = \sigma\sqrt{\Delta t} + O((\Delta t)^{3/2})$. The Cox–Ross–Rubinstein (CRR) choice drops the remainder:

$$u \;=\; e^{\sigma\sqrt{\Delta t}}, \qquad d \;=\; \frac{1}{u}, \qquad q \;=\; \frac{e^{r\Delta t} - d}{u - d},$$

with discounting at $e^{-r\Delta t}$ per period. Its variance is $S_t^2\big(\sigma^2\Delta t + O((\Delta t)^2)\big)$.

(d) (The weight.) As $\Delta t \to 0$, the CRR q satisfies

$$q \;=\; \frac{1}{2} + c\,\sqrt{\Delta t} + O(\Delta t), \qquad c = \frac{r - \sigma^2/2}{2\sigma}.$$

So the tree has the right mean in every period, exactly. It has the right variance up to $O((\Delta t)^2)$. There are $N = T/\Delta t$ periods. So the errors add up to $O(\Delta t)$. They vanish as $\Delta t \to 0$. The up factor is a one-standard-deviation move, $\sigma\sqrt{\Delta t}$ in logs, the $\sqrt{\Delta t}$ scaling of §3.1.

We can now state the role of Δt. The interest for one period is of order Δt. The stock's log

move in one period is of order $\sqrt{\Delta t}$, so its square is of order Δt. Over $N = T/\Delta t$ periods the interest adds up to about rT. The squared log moves add up to $\sigma^2 T$. So the interest stays finite, and the total variance stays finite and positive, as $\Delta t \to 0$. This scaling allows a Black–Scholes limit, and Theorem 5.6 proves the limit. With moves of order Δt, the total variance would vanish and no randomness would remain.

Proof.

□

CRR is not the only calibration. Jarrow and Rudd (1983, *Option Pricing*) fix $q = \frac{1}{2}$ and put the drift into the moves: $u = e^{(r-\sigma^2/2)\Delta t+\sigma\sqrt{\Delta t}}$ and $d = e^{(r-\sigma^2/2)\Delta t-\sigma\sqrt{\Delta t}}$. Both trees match the mean and the variance of one period up to $O((\Delta t)^2)$. Both converge to the same limit. This section uses CRR. Its $d = 1/u$ keeps the proof of Theorem 5.6 short.

The next theorem goes from one period to the whole tree. Proposition 5.4(d) wrote the tree price of the call as two tail probabilities. The proof shows that they converge to $\Phi(d_1)$ and $\Phi(d_2)$ separately. The first tail converges to $\Phi(d_1)$. It is computed under $\widetilde{Q}$, a discrete version of the share measure of §4.6.

Theorem 5.6 (CRR → Black–Scholes)**.**
Let C_N be the N-period CRR price of the European call (strike K, maturity T). Then

$$C_N \longrightarrow C_{\mathrm{BS}} = S_0\,\Phi(d_1) - Ke^{-rT}\,\Phi(d_2) \qquad (N \to \infty).$$

As the periods shrink, the tree price becomes the Black–Scholes price. The observed error is $O(1/N)$ (Table 1), so doubling the periods halves it.

Proof.

□

The theorem gives convergence. Table 1 shows how fast. It reports the at-the-money benchmark of §4.5 ($S_0 = K = 100$, $r = 0$, $\sigma = 0.2$, $T = 1$; Black–Scholes price 7.9656), computed by the committed script `figures/crr_convergence.py`.

Table 1: CRR convergence on the ATM benchmark.

N	C_N	error	$N\times$ error
10	7.7692	−0.1963	−1.96
25	8.0451	+0.0795	+1.99
50	7.9258	−0.0397	−1.99
100	7.9457	−0.0199	−1.99
250	7.9576	−0.0080	−1.99
500	7.9616	−0.0040	−1.99

Two features of the table matter. First, $N\times$error is essentially constant (≈ 2 in absolute value). So the error is $O(1/N)$, an order better than the $O(1/\sqrt{N})$ the central limit theorem alone would suggest. The symmetric CRR tree cancels the first-order error term. Second, the sign of the error flips with N (see $N = 25$). As the grid shifts, the discrete strike threshold a_N lands alternately above and below the kink of the payoff. This produces the well-known odd–even oscillation. The practical consequence: average two adjacent-N prices rather than simply increasing N. Figure 4 shows both features across the whole range of N.

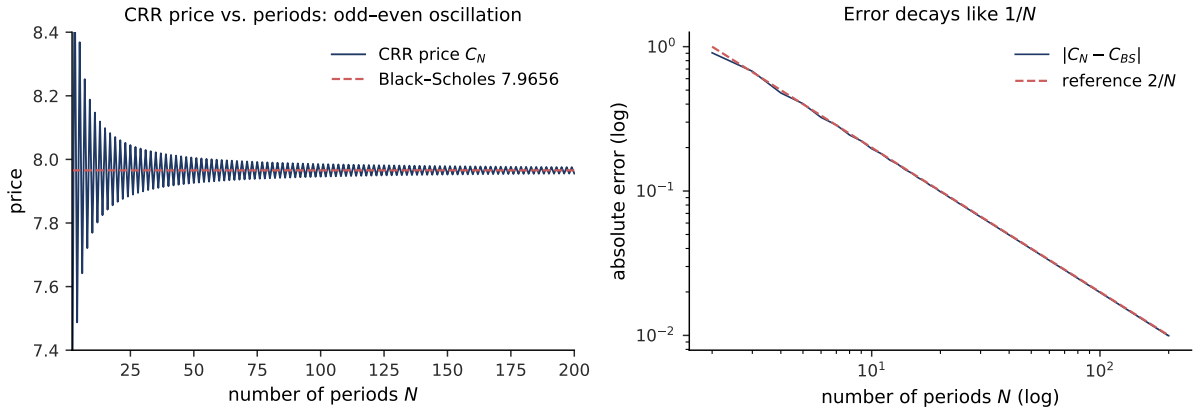


Figure 4: Actual CRR runs on the ATM benchmark (generated by `figures/make_figures.py`). Left: the price oscillates in a shrinking cone around the Black–Scholes value, the odd–even zigzag of the strike against the grid. Right: the absolute error on log–log axes lies on the $2/N$ reference line; slope -1 is the graphical signature of $O(1/N)$ convergence.

5.4 American Options on the Tree

This subsection prices options that can be exercised before maturity. The tree handles them with a one-line change to the roll-back rule. This is the main practical advantage of the lattice. The closed form of Section 4 cannot do it, and neither can the plain Monte Carlo method of §5.5.

A European option can be exercised only at date N. An American option can be exercised at any date $k = 0, 1, \ldots, N$, once. If the holder exercises at date k, she receives the *exercise value* $h(S_k)$: for a put, $(K - S_k)^+$; for a call, $(S_k - K)^+$. The holder will exercise only when that is better than keeping the option alive.

So at every node the holder faces a comparison: the exercise value now, against the value of holding the option one more period. The value of holding is the roll-back value of Proposition 5.4(a). We call it the *continuation value*. The option is worth the larger of the two.

Definition 5.7 (American option and its tree price)**.**
An American option with exercise value $h(\cdot)$ gives its holder the right to receive $h(S_k)$ at any one date $k \in \{0, 1, \dots, N\}$ of her choice. Its price on the N-period binomial tree is defined by backward induction:

$$V_N = h(S_N), \qquad V_k = \max\Big\{ h(S_k),\ \underbrace{\frac{q\,V^u_{k+1} + (1-q)\,V^d_{k+1}}{1 + r\Delta t}}_{\text{continuation value}} \Big\}, \qquad k = N-1, \dots, 0.$$

A node where $h(S_k)$ is strictly larger than the continuation value is an early-exercise node*: the holder exercises there.*

The recursion is the European roll-back rule with one extra operation at each node, the maximum. Its economic reading is the holder's decision. At every node, compare what exercising pays now with what holding is worth, and take the larger. The price is then the cost of the best policy.

The seller can hedge this price. At every node Proposition 5.4(c) replicates the continuation value. At an early-exercise node the seller's portfolio is worth at least $h(S_k)$, and she pays it out. The proof that this recursion gives the unique no-arbitrage price is imported at book level from Shreve, *Stochastic Calculus for Finance I*, Chapter 4. That proof has two halves: the seller's hedge and the buyer's optimal stopping rule. The argument needs stopping times, which this course does not build.

Early exercise comes from interest. Exercising a put early gives $K - S_k$ today instead of a payoff later. Receiving K today is worth more than receiving K later when $r > 0$. That is the gain. The cost is giving up the chance that the stock falls further. The maximum compares the two. Example 5.8 shows a node where the gain is larger than the cost. Exercise 5.6 shows that with $r = 0$ no such node exists on the same tree.

Example 5.8. Same tree as Example 5.1: $S_0 = 100$, $u = 1.2$, $d = 0.8$, $r = 10\%$ per year, $\Delta t = 1$, $T = 2$ years, $N = 2$. Price the European put and the American put with $K = 100$. Say at which node, if any, the American put is exercised early.

For the call the comparison always goes the other way. The next proposition proves it on the tree.

Proposition 5.9 (No early exercise of the call)**.**
In the N-period binomial model with $r \geq 0$ and a stock that pays no dividends, the tree price of the American call with strike K equals the tree price of the European call at every node. No node is an early-exercise node.

The holder of a call gains nothing by paying K early. Paying later is cheaper when $r \geq 0$, and holding keeps the chance that the stock rises further. So the right to exercise early is worth zero, and the American call price equals the European one.

Proof.

□

The same argument fails for the put at Step 3, because $K(1 + r\Delta t)^{-(N-k)} - S_k \le K - S_k$ runs the wrong way. That is where the extra value of the American put in Example 5.8 comes from.

5.5 Monte Carlo Methods

This subsection estimates the price $e^{-rT}\mathbb{E}^{Q}[h(S_T)]$ by averaging simulated payoffs. In §5.5 and §5.6, N counts simulated paths, not tree periods. Monte Carlo is the slowest accurate method. It is also the only method whose cost does not grow with the number of underlyings. Both facts follow from one theorem. The method entered option pricing with Boyle (1977), *Journal of Financial Economics*.

The two methods of this section are mirror images. The tree works backward from maturity. Monte Carlo works forward from today. The tree converges to Black–Scholes by the central limit theorem (Theorem 5.6). Monte Carlo converges by the law of large numbers. The tree handles early exercise with one maximum. Monte Carlo handles path-dependent payoffs and many underlyings, where a tree becomes too large to build. §5.6 shows how far simulation can be pushed toward early exercise.

Monte Carlo in one picture: draw N points uniformly in the unit square. The fraction that lands inside the quarter circle $x^2 + y^2 \le 1$ estimates its area, $\pi/4$. The fraction is a sample proportion, so its standard error is $\sqrt{p(1-p)/N}$ with $p = \pi/4 \approx 0.785$. At $N = 10{,}000$ that is about 0.004, which is about 0.016 on the estimate of π itself. Option pricing is the same idea with a different integrand. The discounted payoff takes the place of the indicator of the quarter circle.

Draw N independent copies $S_T^{(i)}$ of the terminal price under Q and form

$$\widehat{C}_N = e^{-rT}\,\frac{1}{N}\sum_{i=1}^{N} h\big(S_T^{(i)}\big).$$

By the law of large numbers (Statistics I), $\widehat{C}_N$ is consistent. By the central limit theorem its error

is asymptotically normal:

$$\widehat{C}_N \;\approx\; C \;\pm\; z_{\alpha/2}\,\frac{s}{\sqrt{N}},$$

where s is the sample standard deviation of the discounted payoffs. The rate $N^{-1/2}$ contains no dimension. Pricing on five underlyings costs the same order of work as on one. So Monte Carlo is the standard method for multi-asset and path-dependent claims. The cost is the exponent $1/2$: each extra decimal digit of accuracy needs a hundred times more paths.

For geometric Brownian motion no discretization is needed. Sample $Z \sim N(0,1)$ and set $S_T = S_0 \exp\bigl((r-\sigma^2/2)T+\sigma\sqrt{T}Z\bigr)$, which is *exact* (§3.4). For models without a closed-form terminal law (Heston, Section 7) one simulates the SDE on a time grid by the Euler–Maruyama step $X_{t+\Delta t} = X_t + \mu(t,X_t)\Delta t + \sigma(t,X_t)\sqrt{\Delta t}\,Z$. The discretization bias this introduces (the weak error, $O(\Delta t)$ under regularity conditions we import) then adds to the statistical error. (Remark: refinements of path construction such as the Brownian bridge are deliberately outside this course's scope. The name is recorded here so you can find the literature.)

For a European call the whole method is four steps.

1. Draw N independent standard normals $Z^{(1)}, \dots, Z^{(N)}$.
2. Turn each draw into a terminal price: $S_T^{(i)} = S_0 \exp\bigl((r-\sigma^2/2)T + \sigma\sqrt{T}\,Z^{(i)}\bigr)$.
3. Turn each price into a discounted payoff: $C^{(i)} = e^{-rT}\bigl(S_T^{(i)} - K\bigr)^+$.
4. Average: $\widehat{C}_N = \frac{1}{N}\sum_{i=1}^{N} C^{(i)}$. Report the standard error $s/\sqrt{N}$ next to it, where s is the sample standard deviation of the $C^{(i)}$.

Step 4 is the law of large numbers at work. A Monte Carlo estimate without its standard error is not a result. Exercise 5.11 runs these four steps on the web page.

Where do the normal draws come from? A computer produces *pseudo-random* numbers: a deterministic recursion started from a *seed*. Its output passes statistical tests for uniformity and independence. The same seed gives the same sequence. That is what makes a simulation reproducible. The recursion delivers Uniform$(0,1)$ draws. The Box–Muller transform turns two of them into two independent standard normals:

$$Z_1 = \sqrt{-2\ln U_1}\,\cos(2\pi U_2), \qquad Z_2 = \sqrt{-2\ln U_1}\,\sin(2\pi U_2)$$

(Exercise 5.14). Software hides these steps. The seed is yours to set, and a reported simulation should report it.

Two or more underlyings need one more ingredient. Suppose two stocks follow geometric Brownian motions. Their Brownian motions have correlation ρ. Each terminal price needs its own normal draw, ε_1 and ε_2, and the two draws must have correlation ρ. A random number generator gives independent normals Z_1 and Z_2 only. The fix is a linear combination:

$$\varepsilon_1 = Z_1, \qquad \varepsilon_2 = \rho\,Z_1 + \sqrt{1-\rho^2}\,Z_2.$$

Check: $\mathrm{Var}(\varepsilon_2) = \rho^2 + (1-\rho^2) = 1$, and $\mathrm{Cov}(\varepsilon_1,\varepsilon_2) = \rho\ \mathrm{Var}(Z_1) = \rho$. Both are linear in independent normals, so the pair is bivariate normal with the required correlation. This is the Cholesky factorization of the 2×2 correlation matrix. Exercise 5.12 does three assets, where the same idea needs one more line.

Example 5.10. The accuracy budget. For the ATM benchmark, compute $\mathrm{Var}^{Q}\left[(S_T - K)^+\right]$ analytically, and determine how many paths a 1-cent standard error requires. (Use $\Phi(0.3) = 0.617911$, $\Phi(0.1) = 0.539828$ and $e^{0.04} = 1.040811$; carry six decimals: the factor 10^4 uses up four of them, and the subtraction then cancels the two leading digits. Use the identity $\mathbb{E}^{Q}[S_T^2 \mathbf{1}\{S_T > K\}] = S_0^2 e^{(2r+\sigma^2)T}\,\Phi(d_1 + \sigma\sqrt{T})$, proved by the completing-the-square step of the Black–Scholes formula in §4.5 applied with $2\sigma\sqrt{T}$ in place of $\sigma\sqrt{T}$.)

Figure 5 (left) shows the budget in an actual simulation run. Its 95% confidence band shrinks like $1/\sqrt{N}$ around the Black–Scholes value. After 200,000 paths the estimate is 7.984 ± 0.058: still six cents on each side of the point estimate. The budget shows why variance reduction is needed. Two standard techniques almost always cut the variance by more than they cost to implement.

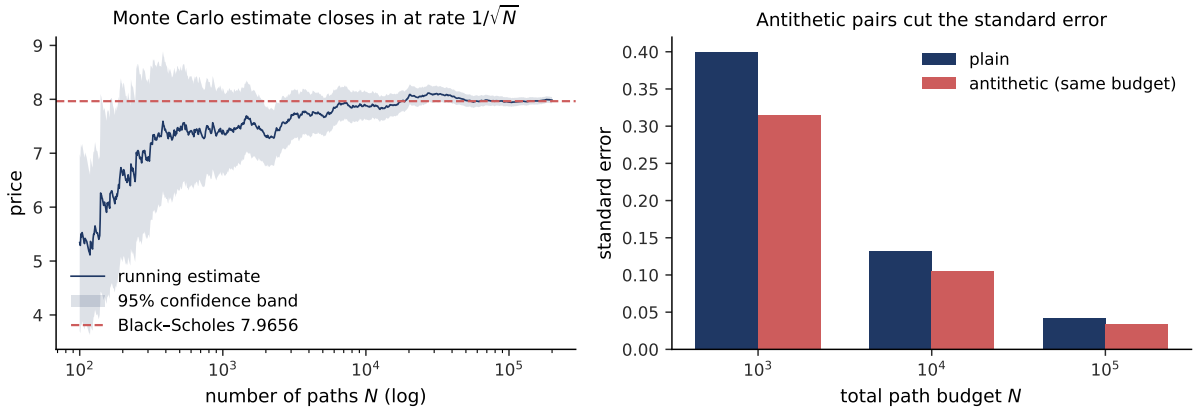


Figure 5: Actual Monte Carlo runs (seed 42; generated by `figures/make_figures.py`). Left: the running estimate and its 95% band close in on the Black–Scholes price at the $1/\sqrt{N}$ rate. Note the logarithmic path axis: each equal step rightward costs ten times more work. Right: at every budget, antithetic pairs deliver a smaller standard error than plain sampling. Here the error is about 21% smaller, a 37% saving in paths for equal accuracy.

Antithetic variates. Simulate in pairs: for each draw Z, also use $-Z$, and average the pair's payoffs. The pair estimator has variance

$$\operatorname{Var}\left(\frac{f(Z)+f(-Z)}{2}\right) = \frac{\operatorname{Var}(f)}{2}\big(1+\rho\big), \qquad \rho = \operatorname{Corr}\big(f(Z), f(-Z)\big),$$

against $\operatorname{Var}(f)/2$ for two independent draws. The technique wins exactly when $\rho < 0$. That holds for antithetic pairs and monotone payoffs: a call is high when Z is high and low when $-Z$ is high. The nearer the payoff is to linear in Z, the closer ρ is to -1 and the larger the gain (Exercise 5.8).

Control variates. Alongside the target payoff f, simulate a correlated quantity g whose expectation is known in closed form. For a path-dependent claim such as an Asian option, g can be the terminal price, with $\mathbb{E}[S_T] = S_0 e^{rT}$. It can also be the geometric-average Asian, which has a closed form. Then correct the estimate by the observed error of the control:

$$\widehat{C}^{\,\mathrm{cv}} = \bar{f} - \beta\big(\bar{g} - \mathbb{E}[g]\big), \qquad \beta^* = \frac{\operatorname{Cov}(f,g)}{\operatorname{Var}(g)},$$

which multiplies the variance by $1-\rho_{f,g}^2$ (Exercise 5.9). The optimal β^* is a regression coefficient. Variance reduction is a regression of the payoff on a quantity whose mean is known. A control variate with $\rho^2 = 0.9$ cuts the variance by a factor of ten. That is worth the same as ten times as many paths, at almost no extra cost.

5.6 Least-Squares Monte Carlo

This subsection prices American options by simulation. Plain Monte Carlo cannot do it, and the first paragraph says why. The method that can, least-squares Monte Carlo (LSM), replaces the missing ingredient by a regression. The subsection then works the method through on eight paths, with every number printed.

Recall what the tree needs at each node: the continuation value, a conditional expectation of next period's value given the current stock price. The tree has it because every node branches into the same two successors, so the average over the successors is available. A simulated path does not branch. At date t on path i there is one future, the rest of path i. The average over that one future is just the path's own later payoff, which is not an expectation at all. So a simulation gives no continuation value and no comparison can be made.

The idea of Longstaff and Schwartz (2001) is to estimate the continuation value by regression across paths. At date t, every path has a current stock price $X^{(i)} = S_t^{(i)}$ and a realized later payoff $Y^{(i)}$ (discounted back to t). The conditional expectation $\mathbb{E}^{Q}[Y \mid X = x]$ is a function of x. Regress Y on a few functions of X across the paths. Then read the fitted value at $X^{(i)}$ as the estimated continuation value on path i.

Why does a regression estimate a conditional expectation? Among all functions g of X, the conditional expectation has the smallest mean squared error:

$$\mathbb{E}\big[(Y-g(X))^2\big] = \mathbb{E}\big[(Y-\mathbb{E}[Y\mid X])^2\big] + \mathbb{E}\big[(\mathbb{E}[Y\mid X]-g(X))^2\big].$$

The cross term $2\,\mathbb{E}[(Y-\mathbb{E}[Y|X])(\mathbb{E}[Y|X]-g(X))]$ vanishes by the law of iterated expectations (condition on X first). Least squares does the same minimization with two substitutions. The

sample average replaces the expectation. And g is restricted to a small family, here the quadratic polynomials $a+bx+cx^2$. This is the regression of Statistics II, used as an estimator of a conditional expectation rather than as a model.

The algorithm, for an American put with strike K on exercise dates $t_1 < t_2 < \cdots < t_M = T$, with N simulated paths:

1. Simulate N paths of the stock under Q on the exercise dates.
2. At $t_M = T$, set the cash flow of path i to $(K - S_T^{(i)})^+$.
3. Go backward, $t_m = t_{M-1}, \ldots, t_1$. At each date:
   (a) Select the paths that are in the money at t_m, that is $K - S_{t_m}^{(i)} > 0$. Only these paths face a decision.
   (b) For each selected path set $X^{(i)} = S_{t_m}^{(i)}$ and $Y^{(i)}$ = the cash flow the path receives after t_m under the current rule, discounted to t_m.
   (c) Regress Y on $1, X, X^2$ by least squares. The fitted value $\hat{a}+\hat{b}X^{(i)}+\hat{c}(X^{(i)})^2$ is the estimated continuation value of path i.
   (d) If the exercise value $K - X^{(i)}$ exceeds the estimated continuation value, path i exercises at t_m. Set its cash flow at t_m to $K - X^{(i)}$ and set all its later cash flows to zero.
4. Each path now has at most one nonzero cash flow. The price is the average over paths of that cash flow discounted to date 0.

Only in-the-money paths enter the regression. On an out-of-the-money path there is no decision to make. Leaving those paths out gives a better fit where the fit matters. The regression uses realized later payoffs, not fitted ones, as Y. So the fitted curve is only used to make the decision, and the price is built from actual simulated cash flows.

Example 5.11. Price an American put with strike $K = 1.10$ on a stock with $S_0 = 1$, exercisable at dates $t = 1, 2, 3$, with rate $r = 6\%$ per year, $\Delta t = 1$, continuously compounded, so that one period of discounting is $e^{-0.06} = 0.94176$. Use the eight paths of the numerical example in Longstaff and Schwartz (2001):

path	$t = 0$	$t = 1$	$t = 2$	$t = 3$
1	1.00	1.09	1.08	1.34
2	1.00	1.16	1.26	1.54
3	1.00	1.22	1.07	1.03
4	1.00	0.93	0.97	0.92
5	1.00	1.11	1.56	1.52
6	1.00	0.76	0.77	0.90
7	1.00	0.92	0.84	1.01
8	1.00	0.88	1.22	1.34

The computation is printed in full below. It is the one example in the section that cannot be done on the board in the time available. Read it with a calculator at hand. Every step is one of the four steps of the algorithm.

Step 1: cash flows at $t = 3$. The payoff $(1.10 - S_3)^+$ on the eight paths:

path	1	2	3	4	5	6	7	8
S_3	1.34	1.54	1.03	0.92	1.52	0.90	1.01	1.34
cash flow at $t = 3$	0	0	0.07	0.18	0	0.20	0.09	0

Step 2: the decision at $t = 2$. The paths in the money at $t = 2$ are those with $S_2 < 1.10$: paths 1, 3, 4, 6, 7. For each, X is S_2 and Y is the $t = 3$ cash flow discounted one period, $Y = 0.94176 \times$ (cash flow at $t = 3$):

path	1	3	4	6	7
$X = S_2$	1.08	1.07	0.97	0.77	0.84
Y	0	0.0659	0.1695	0.1884	0.0848

Regress Y on $1, X, X^2$. The least-squares coefficients solve the three normal equations

$$n\,\hat{a} + \sum X\,\hat{b} + \sum X^2\,\hat{c} = \sum Y,$$
$$\sum X\,\hat{a} + \sum X^2\,\hat{b} + \sum X^3\,\hat{c} = \sum XY,$$
$$\sum X^2\,\hat{a} + \sum X^3\,\hat{b} + \sum X^4\,\hat{c} = \sum X^2Y,$$

with the sums over the five selected paths. The sums are exact (the X and Y values have at most four decimals), and they must be used exactly:

$$n = 5, \quad \sum X = 4.73, \quad \sum X^2 = 4.5507, \quad \sum X^3 = 4.446665, \quad \sum X^4 = 4.40597955,$$

$$\sum Y = 0.5086, \quad \sum XY = 0.451228, \quad \sum X^2Y = 0.4064687.$$

The three equations are badly conditioned. If the sums are rounded to four places, the coefficients move by several tenths, to -1.3535, 3.6033, -2.1463. No decision below changes. With the exact sums, any calculator or spreadsheet gives

$$\hat{a} = -1.0691, \qquad \hat{b} = 2.9817, \qquad \hat{c} = -1.8127,$$

so the estimated continuation value is $\widehat{\mathrm{Cont}}(X) = -1.0691 + 2.9817\,X - 1.8127\,X^2$. (Longstaff and Schwartz report -1.070, 2.983, -1.813. They regress on the unrounded Y; that regression gives the same five fitted values below to four decimals.) Now compare, path by path:

path	1	3	4	6	7
exercise value $1.10 - X$	0.02	0.03	0.13	0.33	0.26
continuation $\widehat{\mathrm{Cont}}(X)$	0.0367	0.0459	0.1175	0.1520	0.1564
decision	hold	hold	exercise	exercise	exercise

For instance on path 4, $-1.0691 + 2.9817(0.97) - 1.8127(0.97)^2 = -1.0691 + 2.8922 - 1.7056 = 0.1175 < 0.13$, so path 4 exercises at $t = 2$. (The table lists the fitted values from the unrounded coefficients. Recomputing from the four-place coefficients can differ by one unit in the fourth decimal.) Paths 4, 6, 7 exercise at $t = 2$ and their $t = 3$ cash flows are set to zero.

The cash-flow matrix is now:

path	1	2	3	4	5	6	7	8
$t = 2$	0	0	0	0.13	0	0.33	0.26	0
$t = 3$	0	0	0.07	0	0	0	0	0

Step 3: the decision at $t = 1$. The paths in the money at $t = 1$ are those with $S_1 < 1.10$: paths 1, 4, 6, 7, 8. Now Y is the path's later cash flow discounted to $t = 1$. A $t = 2$ cash flow is discounted one period. A $t = 3$ cash flow is discounted two periods ($e^{-0.12} = 0.88692$). Path 1 has no later cash flow, so $Y = 0$. The same holds for path 8.

path	1	4	6	7	8
$X = S_1$	1.09	0.93	0.76	0.92	0.88
Y	0	0.1224	0.3108	0.2449	0

(For path 4, $Y = 0.94176 \times 0.13 = 0.1224$; for path 6, $0.94176 \times 0.33 = 0.3108$; for path 7, $0.94176 \times 0.26 = 0.2449$.) The exact sums are

$$n = 5, \quad \sum X = 4.58, \quad \sum X^2 = 4.2514, \quad \sum X^3 = 3.998522, \quad \sum X^4 = 3.8093437,$$

$$\sum Y = 0.6781, \quad \sum XY = 0.575348, \quad \sum X^2 Y = 0.4926652,$$

and the normal equations, again solved with the exact sums, give

$$\hat{a} = 2.0377, \qquad \hat{b} = -3.3359, \qquad \hat{c} = 1.3567.$$

Compare:

path	1	4	6	7	8
exercise value $1.10 - X$	0.01	0.17	0.34	0.18	0.22
continuation $\widehat{\mathrm{Cont}}(X)$	0.0135	0.1088	0.2861	0.1170	0.1528
decision	hold	exercise	exercise	exercise	exercise

Paths 4, 6, 7, 8 exercise at $t = 1$. Their later cash flows are set to zero. Path 1 holds, and since its later cash flow is 0, it ends with nothing.

Step 4: the price. The final cash-flow matrix has at most one nonzero entry per path:

path	1	2	3	4	5	6	7	8
$t = 1$	0	0	0	0.17	0	0.34	0.18	0.22
$t = 2$	0	0	0	0	0	0	0	0
$t = 3$	0	0	0.07	0	0	0	0	0

Discount each cash flow to date 0 ($e^{-0.06} = 0.94176$ for $t = 1$, $e^{-0.18} = 0.83527$ for $t = 3$) and average over the eight paths:

$$\widehat{V}_0 = \frac{(0.17 + 0.34 + 0.18 + 0.22)(0.94176) + (0.07)(0.83527)}{8} = \frac{0.8570 + 0.0585}{8} = \frac{0.9155}{8} = 0.1144.$$

For comparison, the European put on the same eight paths pays only at $t = 3$:

$$\frac{(0.07 + 0.18 + 0.20 + 0.09)\,(0.83527)}{8} = \frac{0.54 \times 0.83527}{8} = 0.0564.$$

The early-exercise right doubles the value on these paths. Four of the eight paths exercise at $t = 1$. At those nodes the regression says that the exercise value is larger than the expected later payoff.

Three remarks close the subsection. First, eight paths give a rough answer. In practice N is in the tens of thousands, and the regression at each date then has thousands of points. Second, the choice of basis matters. More paths reduce the sampling noise, but they do not repair a basis that is too simple. Exercise 5.10 repeats Step 3 with a linear fit and gets a different decision on path 1.

In their large examples, Longstaff and Schwartz regress on weighted Laguerre polynomials, $e^{-X/2}$, $e^{-X/2}(1 - X)$, $e^{-X/2}(1 - 2X + X^2/2)$. These are the quadratics 1, $1 - X$, $1 - 2X + X^2/2$, each multiplied by the weight $e^{-X/2}$. Their span is $e^{-X/2}$ times a quadratic, which is a different family from the plain quadratics. With eight paths the plain quadratic basis is enough to show the method.

The estimate converges to the true price in two steps: for a fixed basis as the number of paths grows, and then as the number of basis functions grows. That proof is imported from Clément, Lamberton and Protter (2002, *Finance and Stochastics*).

Third, the method extends with no change to several underlyings and to path-dependent exercise values, because the regression simply takes more regressors. That is why LSM is the standard method for American-style claims whenever a tree is unavailable.

Exercises

Exercise 5.1. *In the one-period binomial model with parameters $u, d, r, \Delta t$, show directly (without quoting Proposition 5.3) that the discounted stock is a $\mathbf{Q}$-martingale if and only if $q = \frac{(1+r\Delta t)-d}{u-d}$, and that $q \in (0,1)$ if and only if $d < 1 + r\Delta t < u$. Interpret the two failure cases economically.*

Exercise 5.2. *Repeat Example 5.1 with $r = 0$, so that $1 + r\Delta t = 1$: on the same tree ($S_0 = 100$, $u = 1.2$, $d = 0.8$), price the two-period European call and the two-period European put with $K = 100$ by the roll-back rule. Check put–call parity $C_0 - P_0 = S_0 - K(1 + r\Delta t)^{-2}$ at date 0 and at both date-1 nodes. Why is the call worth less than in Example 5.1?*

Exercise 5.3. *On the three-period tree of Example 5.2 with $r = 0$ ($S_0 = 100$, $u = 1.2$, $d = 0.8$, $1 + r\Delta t = 1$), price the European call and the European put with $K = 100$ by the roll-back rule, and verify put–call parity at every node of the tree.*

Exercise 5.4. *For $\sigma = 0.2$, $r = 0.05$, $\Delta t = 0.01$, compute the CRR q exactly and compare with the approximation of Proposition 5.5(d). (Exact value: $q = 0.507502$.)*

Exercise 5.5. *Show that the CRR lattice is arbitrage-free for all sufficiently small Δt, by verifying $d < e^{r\Delta t} < u$ (Proposition 5.3's condition). For which Δt could it fail?*

Exercise 5.6. *On the tree of Exercise 5.2 ($S_0 = 100$, $u = 1.2$, $d = 0.8$, $r = 0$), price the two-period* American *put with $K = 100$ by Definition 5.7. Show that no node is an early-exercise node, so that the American price equals the European price of Exercise 5.2. Then explain in two sentences why Example 5.8, on the same tree with $r = 10\%$, does have an early-exercise node.*

Exercise 5.7. *Check Proposition 5.4(d) on the three-period call of Example 5.2 ($S_0 = 100$, $u = 1.2$, $d = 0.8$, $1 + r\Delta t = 1.1$, $K = 100$): find m and $\tilde{q}$, compute the two tail probabilities $\widetilde{\mathbf{Q}}(J \geq m)$ and $\mathbf{Q}(J \geq m)$, and verify that $S_0\,\widetilde{\mathbf{Q}}(J \geq m) - K(1 + r\Delta t)^{-3}\,\mathbf{Q}(J \geq m) = 27.8926$.*

Exercise 5.8. *Let $f(Z) = a + bZ$ be a linear payoff. Show the antithetic pair estimator has variance* zero*, and explain in one sentence why payoffs that are nearly linear in Z enjoy nearly this gain.*

Exercise 5.9. *Derive the optimal control-variate coefficient: show that $\beta = \mathrm{Cov}(f, g)/\mathrm{Var}(g)$ minimizes $\mathrm{Var}\big(f - \beta(g - \mathbb{E}g)\big)$ and that the minimized variance is $\mathrm{Var}(f)(1 - \rho_{f,g}^2)$. Where have you met this formula before?*

Exercise 5.10. *Repeat Step 3 of Example 5.11 (the decision at $t = 1$) with the linear basis $\{1, X\}$ in place of $\{1, X, X^2\}$, keeping the $t = 2$ decisions as they are. Report the fitted coefficients, the five continuation values, which decisions change, and the new price.*

Exercise 5.11. *Price a European call with $S_0 = 100$, $K = 80$, $r = 1.5\%$, $\sigma = 20\%$ and $T = 1$ year by Monte Carlo, following the four steps of §5.5. Use the Monte Carlo method of `web/index.html` (or any software), first with $N = 10{,}000$ paths and then with $N = 1{,}000{,}000$. Report the estimate and its standard error each time, and compare with the Black–Scholes price. How many paths would a standard error of one cent need?*

Exercise 5.12. *A rainbow option on three stocks needs three standard normal draws $\varepsilon_1, \varepsilon_2, \varepsilon_3$ with correlations ρ_{12}, ρ_{13}, ρ_{23}. The generator gives independent standard normals Z_1, Z_2, Z_3. Extend the construction of §5.5: write*

$$\varepsilon_1 = Z_1, \qquad \varepsilon_2 = \ell_{21} Z_1 + \ell_{22} Z_2, \qquad \varepsilon_3 = \ell_{31} Z_1 + \ell_{32} Z_2 + \ell_{33} Z_3,$$

and find the five coefficients in terms of the correlations. Check your answer on $\rho_{12} = 0.5$, $\rho_{13} = 0.3$, $\rho_{23} = 0.6$.

Exercise 5.13. *Calibrate a two-period CRR tree to $S_0 = 100$, $\sigma = 20\%$, $T = 1$ year and $r = 1\%$ per year, continuously compounded. The tree has $N = 2$ periods, so $\Delta t = \frac{1}{2}$. Compute u, d, the growth factor $e^{r\Delta t}$ and q. Price the European put with $K = 100$ by backward induction, then the American put. Compare the European price with the Black–Scholes put, 7.4383, and say why two periods are far off. (The page `web/index.html` carries this tree as an experiment card.)*

Exercise 5.14. *Let U_1 and U_2 be independent Uniform$(0,1)$. Show that $Z_1 = \sqrt{-2\ln U_1}\cos(2\pi U_2)$ and $Z_2 = \sqrt{-2\ln U_1}\sin(2\pi U_2)$ are independent standard normals (the Box–Muller transform of §5.5). Hint: write two independent standard normals in polar coordinates, and find the law of the radius and of the angle.*