CHAPTER THIRTY-THREE

Answers to Selected Exercises

More questions may be easier to answer than just one question.
Imre Lakatos (1922–1974), *Proofs and Refutations*

CHAPTER 2

Exercise 2.2.2: (1) Recall that $\sum_{i=1}^{n} i = n(n+1)/2$. (2) Use $\sum_{i=1}^{n} i^2 = (2n^3 + 3n^2 + n)/6$. (3), (4) Use $\sum_{i=0}^{k} 2^i = 2^{k+1} - 1$. (5) Use Euler's summation formula [461, p. 18],

$$\int_a^{b+1} g(x)\,dx \le \sum_{i=a}^{b} g(i) \le \int_{a-1}^{b} g(x)\,dx.$$

CHAPTER 3

Exercise 3.1.1: It is sufficient to show that $g(m) \equiv (1 + \frac{1}{m})^m$ is an increasing function of m. Note that

$$g'(m) = g(m)\left[\ln\left(1 + \frac{1}{m}\right) - \frac{1}{m+1}\right].$$

We can show the expression within the brackets to be positive by differentiating it with respect to m.

An alternative approach is to expand $g(m)$ and $g(m+1)$ as polynomials of $x \equiv 1/m$ using the binomial expansion. It is not hard to see that every term in $g(m+1)$, except the one of degree $m+1$ (which $g(m)$ does not have), is at least as large as the term of the same degree in $g(m)$. This approach does not require calculus.

Exercise 3.1.2: Monthly compounding, i.e., 12 times per annum. This can be verified by noting that $18.70/12 = 1.5583$.

Exercise 3.1.3: (1) The computing power's growth function is $(1.54)^n$, where n is the number of years since 1987. The equivalent continuous compounding rate is 43.18%. Because the memory capacity has quadrupled every 3 years since 1977, the function is $4^{n/3} = (1.5874)^n$, where n is the number of years since 1977. The equivalent continuous compounding rate is 46.21%. (2) It is $(500{,}000/300{,}000)^{1/4} - 1 \approx 13.6\%$. Data are from [574].

Exercise 3.2.1:

$$\text{PV} = \sum_{i=0}^{nm-1} C\left(1 + \frac{r}{m}\right)^{-i} = C\,\frac{1 - \left(1 + \frac{r}{m}\right)^{-nm}}{(r/m)}\left(1 + \frac{r}{m}\right).$$

Exercise 3.3.1: We derived Eq. (3.8) by looking forward into the future. We can also derive the same relation by looking back into the past: Right after the kth payment, the remaining principal is the value of the original principal minus the value of all the payments made to date, which is exactly what the formula says. Mathematically,

$$C\,\frac{1-\left(1+\frac{r}{m}\right)^{-nm}}{\frac{r}{m}}\left(1+\frac{r}{m}\right)^{k}-\sum_{i=1}^{k}C\left(1+\frac{r}{m}\right)^{i-1}$$
$$=C\,\frac{\left(1+\frac{r}{m}\right)^{k}-\left(1+\frac{r}{m}\right)^{-nm+k}}{\frac{r}{m}}-C\,\frac{1-\left(1+\frac{r}{m}\right)^{k}}{\frac{r}{m}}$$
$$=C\,\frac{1-\left(1+\frac{r}{m}\right)^{-nm+k}}{\frac{r}{m}}.$$

Exercise 3.3.2: Note that the PV of an ordinary annuity becomes $\sum_{i=1}^{n} Ce^{ir} = C\,\frac{1-e^{-nr}}{e^{r}-1}$ under continuous compounding. Without loss of generality, assume that the PV of the original mortgage is \$1. The monthly payment is hence $\frac{e^{r}-1}{1-e^{-nr}}$. For an interest rate of $r-x$, the level payment is $D \equiv \frac{e^{r-x}-1}{1-e^{-n(r-x)}}$. The new instrument's cash flow is $De^{x}, De^{2x}, \dots, De^{nx}$ by definition. The PV is therefore

$$\sum_{i=1}^{n} D\,\frac{e^{ix}}{e^{ir}} = D\,\frac{1-e^{-n(r-x)}}{e^{r-x}-1} = 1.$$

Consult [330, p. 120].

Exercise 3.4.1: Apply Eq. (3.11) with $y = 0.0755$, $n = 3$, $C_1 = 1000$, $C_2 = 1000$, $C_3 = 1500$, and $P = 3000$.

Exercise 3.4.2: The FV is $C\,\frac{(1+r)^{n}-1}{r}$ from Eq. (3.4). To guarantee a return of y, the PV should be $C\,\frac{(1+r)^{n}-1}{r}\,\frac{1}{(1+y)^{n}}$.

Exercise 3.4.3: Proposal A's NPV is now \$2,010.014, and Proposal B's is \$1,773.384. Proposal A wins out under this scenario.

Exercise 3.4.4: The iteration is

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} = x_k - \frac{x_k^3 - x_k^2}{3x_k^2 - 2x_k} = \frac{2x_k^2 - x_k}{3x_k - 2}$$

by Eq. (3.15), and the desired sequence is 1.5, 1,2, 1.05, 1.004348, and 1.00003732. The last approximation is very close to the root 1.0.

Exercise 3.4.5: $F'(x) = \frac{f(x)\,f''(x)}{f'(x)^2}$; hence $F'(\xi) = 0$.

Exercise 3.4.6: See Fig. 3.5 [656].

Exercise 3.4.8: Let $y_f(x)$ be the y that makes $f(x, y) = 0$ given x. Similarly, let $y_g(x)$ be the y that makes $g(x, y) = 0$ given x. Assume that y_f and y_g are continuous functions. Then the bisection method is applicable to the function $\phi(x) \equiv y_f(x) - y_g(x)$ if it starts with x_1 and x_2 such that $\phi(x_1)\phi(x_2) < 0$. In so doing, a two-dimensional problem is reduced to the standard one-dimensional problem. Note that $\phi(x) = 0$ if and only if $y_f(x) = y_g(x)$. Unfortunately, this method has no obvious generalizations to $n > 2$ dimensions. Note that $y_f(x)$ and $y_g(x)$ may have to be numerically solved. See [447, p. 585].

Exercise 3.5.1: For each Δt time, it pays out $c\Delta t$ dollars, which is discounted at the rate of r [462].

Exercise 3.5.2: The price that guarantees the return if the bond is called is, from Eq. (3.18),

$$5 \times \frac{1-[\,1+(0.12/2)\,]^{-2\times 5}}{0.12/2} + \frac{100}{[\,1+(0.12/2)\,]^{2\times 5}} = 92.6399.$$

Similarly, the price should be

$$5 \times \frac{1-[\,1+(0.12/2)\,]^{-2\times 10}}{0.12/2} + \frac{100}{[\,1+(0.12/2)\,]^{2\times 10}} = 88.5301$$

dollars if the bond is held to maturity. Hence the price to pay is \$88.5301. (A more rigorous method yielding the same conclusion goes as follows. The formula for various call half-years $n = 10, \dots, 20$ is $(250/3) + (50/3) \times (1.06)^{-n}$, which is minimized at $n = 20$.)

Exercise 3.5.3: It is

$$\begin{aligned}\text{PV} &= \sum_{i=1}^{n} \frac{C(1-T)}{(1+r)^i} + \frac{F - \max((F-P)\,T_G, 0)}{(1+r)^n}\\ &= C(1-T)\,\frac{1-(1+r)^{-n}}{r} + \frac{F-\max((F-P)\,T_G,0)}{(1+r)^n}.\end{aligned}$$

See [827, p. 123].

Exercise 3.5.4: (1) Because $P = F/(1+r)^n$,

$$\frac{\partial P}{\partial n} = -\frac{F\ln(1+r)}{(1+r)^n}, \qquad \frac{\partial P}{\partial r} = -\frac{Fn}{(1+r)^{n+1}}.$$

(2) For $r = 0.04$ and $n = 40$, we have

$$\triangle P \approx -0.00817 \times F \times \Delta n, \qquad \triangle P \approx -8.011 \times F \times \Delta r.$$

Exercise 3.5.5:

$$\sum_{i=1}^{n} \frac{Fr}{(1+r)^i} + \frac{F}{(1+r)^n} = F.$$

Exercise 3.5.6: Intuitively, this is true because the accrued interest is not discounted, whereas the next coupon to be received by the buyer is. Mathematically,

$$F + Fc(1-\omega) = \frac{Fc}{(1+r)^{\omega}} + \frac{Fc}{(1+r)^{\omega+1}} + \frac{Fc}{(1+r)^{\omega+2}} + \cdots + \frac{Fc+F}{(1+r)^{\omega+n-1}}.$$

If $r \geq c$, then $1 + c(1-\omega) \leq \frac{1+r}{(1+r)^{\omega}}$, which can be shown to be impossible.

Another approach is to observe that the buyer should have paid $Fc[\,(1+r)^{1-\omega} - 1\,]$ instead of $Fc(1-\omega)$ and to prove that the former is smaller than the latter.

Exercise 3.5.7: The number of days between the settlement date and the next coupon date is calculated as $30+31+1 = 62$. The number of days in the coupon period being $30+30+31+30+31+31+1 = 184$, the accrued interest is $100 \times (0.1/2) \times (184-62)/184 = 3.31522$. The yield to maturity can be calculated with the help of Eq. (3.20). Note that the PV should be $116 + 3.31522$.

Exercise 3.5.9: Let P_0 be the bond price now and P_1 be the bond price one period from now after the coupon is paid. We are asked to prove that $(c + P_1 - P_0)/P_0 = y$. Without loss of generality, we assume that the par value is \$1. From Eq. (3.18),

$$P_0 = (c/y)[\,1-(1+y)^{-n}\,] + (1+y)^{-n} = (c/y) + (1+y)^{-n}(1-c/y),$$

$$P_1 = (c/y),[\,1-(1+y)^{-(n-1)}\,] + (1+y)^{-(n-1)} = (c/y) + (1+y)^{-(n-1)}(1-c/y).$$

So $P_1 - P_0 = (y-c)(1+y)^{-n}$. Also, $yP_0 = c + (y-c)(1+y)^{-n}$. Hence $c + P_1 - P_0 = yP_0$.

CHAPTER 4

Exercise 4.1.2: Equation (4.1) can be rearranged to become

$$-\frac{\partial P/P}{\partial y} = \frac{C\{-(n/y) + \frac{1}{y^2}[\,(1+y)^{n+1} - (1+y)\,]\} + nF}{C\{[\,(1+y)^{n+1} - (1+y)\,]/y\} + F(1+y)} = \frac{CA(y) + nF}{CB(y) + F(1+y)}.$$

To show that $-(\partial P/P)/\partial y$ increases monotonically if $y > 0$ as C decreases, it suffices to prove that

$$\frac{A(y)}{B(y)} < \frac{nF}{F(1+y)} = \frac{n}{(1+y)}.$$

We can verify that

$$\frac{A(y)}{B(y)} - \frac{n}{(1+y)} = \frac{(1+y)^2[\,(1+y)^n - 1 - ny(1+y)^{n-1}\,]}{B(y)(1+y)}.$$

The expression within the brackets is nonpositive because it is zero for $y = 0$ and its derivative is $-n(n-1)\,y(1+y)^{n-2} \leq 0$. (In fact, the expression within the brackets is strictly negative for $n > 1$.)

https://doi.org/10.1017/CBO9780511546839.034 Published online by Cambridge University Press

Exercise 4.1.3: (1) We observe that $-\frac{\partial P/P}{\partial y} = \frac{1-(1+y)^{-n}}{y}$, which is clearly a decreasing function of y. (2) We differentiate $-(\partial P/P)/\partial y$ with respect to yield y. After rearranging, we obtain

$$-(1+y)\left\{\frac{\sum_i i^2(1+y)^{-i}C_i}{\sum_i(1+y)^{-i}C_i} - \left[\frac{\sum_i i(1+y)^{-i}C_i}{\sum_i(1+y)^{-i}C_i}\right]^2\right\}.$$

The term within the braces can be interpreted in terms of probability theory: the variance of the random variable X defined by

$$\text{Prob}[\, X = j\,] = \frac{(1+y)^{-j}C_j}{\sum_i(1+y)^{-i}C_i}, \quad j \geq 1.$$

Hence it has to be positive. This shows that $-(\partial P/P)/\partial y$ is a decreasing function of yield. See [547, p. 318].

Exercise 4.2.1: 2.67 years.

Exercise 4.2.2: Let D be the modified duration as originally defined. We need to make sure that

$$100 \times D \times \Delta r = D_{\%} \times \Delta r_{\%},$$

where $\Delta r_{\%}$ denotes the rate change in percentage. Because $\Delta r = \Delta r_{\%}/100$, the above identity implies that $D_{\%} = D$.

Exercise 4.2.3: The cash flow of the bond is $C, C, \ldots, C, C+F$, and that of the mortgage is $M, M, \ldots, M$. The conditions imply that $M > C$. The bond hence has a longer duration as its cash flow is more tilted to the rear. Mathematically, we want to prove that

$$C\sum_{i=1}^{n} i(1+y)^{-i} + nF(1+y)^{-n} > M\sum_{i=1}^{n} i(1+y)^{-i}$$

subject to

$$C\sum_{i=1}^{n}(1+y)^{-i} + F(1+y)^{-n} = M\sum_{i=1}^{n}(1+y)^{-i}.$$

The preceding equality implies that $C - M = -\frac{F(1+y)^{-n}}{\sum_{i=1}^{n}(1+y)^{-i}}$. Hence

$$\begin{aligned}
&C\sum_{i=1}^{n} i(1+y)^{-i} + nF(1+y)^{-n} - M\sum_{i=1}^{n} i(1+y)^{-i}\\
&= (C-M)\sum_{i=1}^{n} i(1+y)^{-i} + nF(1+y)^{-n}\\
&= -\frac{F(1+y)^{-n}}{\sum_{i=1}^{n}(1+y)^{-i}}\sum_{i=1}^{n} i(1+y)^{-i} + nF(1+y)^{-n}\\
&= \frac{F}{(1+y)^n}\left[n - \frac{\sum_{i=1}^{n} i(1+y)^{-i}}{\sum_{i=1}^{n}(1+y)^{-i}}\right]\\
&> 0.
\end{aligned}$$

We remark that the market prices of the two instruments are not necessarily equal. To show this, assume the market price of the bond is some $a > 0$ times that of the mortgage. Express M as a function now of C, F, y, n, *and* a. Finally, substitute the expression for M into the formula of the MD for the mortgage.

Exercise 4.2.4: The MD is equal to $\sum_i \frac{iM}{(1+y)^i} / \sum_i \frac{M}{(1+y)^i}$, where M is the monthly payment. Note that M cancels out completely. The rest is just simple algebraic manipulation. See [348, p. 94].

Exercise 4.2.6: This is because convexity is additive. Check with Eq. (4.10) again and observe that it holds as long as each cash flow is nonnegative.

Exercise 4.2.7: Assume again that the liability is L at time m. The present value is therefore $L/(1+y)^m$. A coupon bond with an equal PV will grow to be exactly L at time m. In fact, it is not hard to show that at every point in time, the PV of the bond plus the cash incurred by reinvesting the coupon payments exactly matches the PV of the liability.

Exercise 4.2.8: Rearrange Eq. (4.8) to get

$$\frac{\partial \text{FV}}{\partial y} = (1+y)^{m-1} P \left[m + (1+y) \frac{\partial P/P}{\partial y} \right].$$

If m is less than the MD, the expression within the brackets is negative. This means that the FV decreases as y increases. The other cases can be handled similarly.

Exercise 4.2.9: Let Δy denote the rate change, which may be positive or negative. At time Δt, the PV of the liability will be $L/(1+y-\Delta y)^{m-\Delta t}$, whereas the PV of the bond plus any reinvestment of the interest will be $\sum_i C_i/(1+y-\Delta y)^{i-\Delta t}$. We are asked to prove that

$$\frac{L}{(1+y-\Delta y)^{m-\Delta t}} < \sum_i \frac{C_i}{(1+y-\Delta y)^{i-\Delta t}}.$$

After dividing both sides by $(1+y-\Delta t)^{\Delta t}$, we are left with the equivalent inequality

$$\frac{L}{(1+y-\Delta y)^{m}} < \sum_i \frac{C_i}{(1+y-\Delta y)^{i}}.$$

However, this must be true because we have shown in the text that any instantaneous change in the interest rate raises the bond's value over the liability because of convexity. Note that this conclusion holds for *any* $\Delta t < m$!

Exercise 4.2.10: Solve for ω_1 and ω_2 such that

$$\begin{aligned} 1 &= \omega_1 + \omega_2, \\ 3 &= \omega_1 + 4\omega_2. \end{aligned}$$

The results: $\omega_1 = 1/3$ and $\omega_2 = 2/3$. So 1/3 of the portfolio's market value should be put in bond one, and the remaining 2/3 in bond two.

Exercise 4.2.11: (1) Observe that

$$\begin{aligned} P(y') &= A_1 e^{a_1 y'} + A_2 e^{-a_2 y'} - L_t \\ &= A_1 e^{a_1 y'} + A_2 e^{-a_2 y'} - (A_1 e^{a_1 y} + A_2 e^{-a_2 y}) \\ &= A_1 e^{a_1 y} \left[e^{a_1 (y'-y)} + \frac{a_1}{a_2} e^{-a_2 (y'-y)} - \left(1 + \frac{a_1}{a_2} \right) \right] \end{aligned}$$

after $A_1 a_1 e^{a_1 y} - A_2 a_2 e^{-a_2 y} = 0$ is applied. It is easy to show that

$$g(x) = e^{ax} + \frac{a}{b} e^{-bx} - \left(1 + \frac{a}{b} \right) > 0$$

when $x \neq 0$. The expression within the brackets is hence positive for $y' \neq y$. See [547, pp. 406–407]. A more concise proof is this. It is easy to see that

$$A_1 = \frac{a_2 e^{-a_1 y}}{a_1 + a_2} L_t > 0, \qquad A_2 = \frac{a_1 e^{a_2 y}}{a_1 + a_2} L_t > 0.$$

So $P''(y) = A_1 a_1^2 e^{a_1 y} + A_2 a_2^2 e^{a_2 y} > 0$.

(2) It is not hard to see that a portfolio with cash inflows at t_1, t_2, t_3 can be constructed so that it is more valuable than one with cash inflows at $T < t_3$ after the shift [496, p. 635].

Exercise 4.2.12: To show that Eq. (4.12) is at most j, differentiate it with respect to y and prove that its first derivative is less than zero. Hence Eq. (4.12) is a decreasing function. Finally, show that it approaches j as $y \to 0$. An alternative approach uses the observation that $(1+y)^{-j} \geq 1 - jy$. Yet another alternative applies induction on j to show that Eq. (4.12) is at most j.

Exercise 4.3.1: Let C be the convexity in the original sense and $C_{\%}$ be the convexity in percentage terms. We need to make sure that

$$100 \times C \times (\Delta r)^2 = C_{\%} \times (\Delta r_{\%})^2,$$

where $\Delta r_{\%}$ denotes the rate change in percentage. Because $\Delta r = \Delta r_{\%}/100$, it follows that $C_{\%} = C/100$. See [490, p. 29].

Exercise 4.3.2: This is just a simple application of the chain rule. Recall that $\text{duration} = -\partial P/\partial y$ and $\text{convexity} = (\partial^2 P/\partial y^2)(1/P)$. See [547, p. 333].

Exercise 4.3.4: From Eq. (4.15) we can verify that convexity is equal to $\frac{CA(y)+n(n+1)\,y^3}{CB(y)+y^3(1+y)^2}$, where

$$A(y) = 2(y+1)^{n+2} - 2(y+1)^2 - 2ny(y+1) - n(n+1)\,y^2,$$
$$B(y) = y^2(1+y)^2[\,(1+y)^n - 1\,].$$

To prove the claim, it suffices to show that $A(y)/B(y) < n(n+1)/(1+y)^2$. This is equivalent to proving that

$$G(y) = 2(1+y)^{n+1} - y^2 n(n+1)(1+y)^{n-1} - 2(1+y) - 2ny < 0$$

after simplification. It is easy to show that $G(y)$ is concave for $y > 0$ because $G''(y) < 0$ for $y > 0$. Hence

$$G'(y) = 2(n+1)(1+y)^n - 2yn(n+1)(1+y)^{n-1} - y^2 n(n+1)(n-1)(1+y)^{n-2} - 2 - 2n$$

is less than zero for $y > 0$ as it is a decreasing function of y with $G'(0) = 0$. This leads directly to our conclusion because $G(0) = 0$.

Exercise 4.3.5: A proof is presented with only elementary mathematics. Suppose the universe consists of three kinds of zero-coupon bonds ($n = 3$). Note that $C_i = D_i(D_i + 1)/(1+y)^2$ by Exercise 4.3.3. The portfolio convexity, which is the objective function, is hence

$$\sum_{i=1}^{3} \omega_i C_i = (1+y)^{-2} \sum_{i=1}^{3} \omega_i \left(D_i^2 + D_i\right) = (1+y)^{-2}\left(D + \sum_{i=1}^{3} \omega_i D_i^2\right).$$

The original objective function can be replaced with the simpler

$$\sum_{i=1}^{3} \omega_i D_i^2. \tag{33.1}$$

It is not hard to show that, for distinct $i, j, k \in \{1, 2, 3\}$,

$$\omega_i = \frac{D - D_j + (D_j - D_k)\,\omega_k}{D_i - D_j},$$
$$\omega_j = \frac{D - D_i + (D_i - D_k)\,\omega_k}{D_j - D_i},$$

by manipulation of the two linear equality constraints. So only one variable ω_k remains. Objective function (33.1) becomes

$$\begin{aligned}
&\frac{D - D_j + (D_j - D_k)\,\omega_k}{D_i - D_j}\, D_i^2 + \frac{D - D_i + (D_i - D_k)\,\omega_k}{D_j - D_i}\, D_j^2 + \omega_k D_k^2\\
&= \frac{(D - D_j)\, D_i^2 - (D - D_i)\, D_j^2}{D_i - D_j} + \omega_k \left[\frac{(D_j - D_k)\, D_i^2 - (D_i - D_k)\, D_j^2}{D_i - D_j} + D_k^2\right]\\
&= D(D_i + D_j) - D_i D_j + \omega_k \left(D_i D_j - D_k(D_i + D_j) + D_k^2\right)\\
&= D^2 - (D - D_i)(D - D_j) + \omega_k (D_k - D_i)(D_k - D_j).
\end{aligned} \tag{33.2}$$

Equation (33.2) equals $D^2 - (D - D_1)(D - D_3)$ by picking $i = 1, k = 2, j = 3$, and $\omega_k = 0$. We can confirm that this is a valid choice by checking $\omega_1 = (D - D_3)/(D_1 - D_3) > 0$ and $\omega_3 = (D - D_1)/(D_3 - D_1) > 0$ because $D_1 < D < D_3$. Note that it is a barbell portfolio.

To verify that no other valid portfolios have as high a convexity, we prove that Eq. (33.2) is indeed maximized with $\omega_2 = 0$ as follows. From Eq. (33.2) we must choose $\omega_k = 0$ for the $k = 2$ case because $(D_2 - D_i)(D_2 - D_j) < 0$. Now we consider the objective function with $k \neq 2$. First, we consider $k = 3$. Without loss of generality, we assume that $D_1 = D_i < D_j = D_2$. The formulas for ω_1 and ω_2 dictate that

$$\frac{D_2 - D}{D_2 - D_3} \leq \omega_3 \leq \frac{D_1 - D}{D_1 - D_3}.$$

https://doi.org/10.1017/CBO9780511546839.034 Published online by Cambridge University Press

We plug the preceding second inequality into objective function (33.2) to obtain an upper bound of

$$\begin{aligned}&D^2-(D-D_1)(D-D_2)+\frac{D_1-D}{D_1-D_3}(D_3-D_1)(D_3-D_2)\\&\quad=D^2-(D-D_1)(D-D_2)+(D-D_1)(D_3-D_2)\\&\quad=D^2-(D-D_1)(D-D_3).\end{aligned}$$

Now we consider $k=1$. Without loss of generality, we assume that $D_2=D_i<D_j=D_3$. The formulas for ω_2 and ω_3 dictate that

$$\frac{D_2-D}{D_2-D_1}\leq\omega_1\leq\frac{D_3-D}{D_3-D_1}.$$

We plug the preceding second inequality into objective function (33.2) to obtain an upper bound of

$$\begin{aligned}&D^2-(D-D_2)(D-D_3)+\frac{D_3-D}{D_3-D_1}(D_1-D_2)(D_1-D_3)\\&\quad=D^2-(D-D_2)(D-D_3)+(D-D_3)(D_1-D_2)\\&\quad=D^2-(D-D_1)(D-D_3).\end{aligned}$$

Exercise 4.3.6: Let there be $n\geq 3$ kinds of zero-coupon bonds in the universe. Given a portfolio with more than two kinds of bonds, replace those with duration D, where $D_1<D<D_n$, with bonds with durations D_1 and D_n with a *matching* duration. By Exercise 4.3.5, the new portfolio has a higher convexity. We repeat the steps for each such D until we end up with a barbell portfolio consisting solely of bonds with durations D_1 and D_n.

CHAPTER 5

Exercise 5.2.1: Because $P=\sum_i C_i[\,1+S(i)\,]^{-i}$,

$$P\approx\sum_i\left\{C_i(y)+\frac{\partial C_i(y)}{\partial y}[\,S(i)-y\,]\right\}$$

when $C_i[\,1+S(i)\,]^{-i}$ is expanded in Taylor series at y. The preceding relation, when combined with $P=\sum_i C_i(y)$, leads to

$$\sum_i\frac{\partial C_i(y)}{\partial y}[\,S(i)-y\,]\approx 0.$$

Rearrange the above to obtain the result. See [38, pp. 23–24].

Exercise 5.3.1:

$$100=\sum_{i=1}^{19}\frac{8/2}{(1+0.1)^i}+\frac{(8/2)+100}{[\,1+S(20)\,]^{20}}.$$

Thus

$$66.54=\frac{104}{[\,1+S(20)\,]^{20}},$$

and $S(20)=2.258\%$.

Programming Assignment 5.4.1: See the algorithm in Fig. 33.1.

Exercise 5.5.1: (1) As always, assume $S(1)=y_1$ to start with; hence $S(1)\geq y_1$. From Eq. (5.1) and the definition of yield to maturity,

$$\begin{aligned}&\sum_{i=1}^{k-1}\frac{C}{(1+y_k)^i}+\frac{C+F}{(1+y_k)^k}=\sum_{i=1}^{k-1}\frac{C}{[\,1+S(i)\,]^i}+\frac{C+F}{[\,1+S(k)\,]^k}\\&\quad=\sum_{i=1}^{k-1}\frac{C}{(1+y_{k-1})^i}+\left\{\frac{F}{(1+y_{k-1})^{k-1}}-\frac{F}{[\,1+S(k-1)\,]^{k-1}}\right\}+\frac{C+F}{[\,1+S(k)\,]^k}.\end{aligned}\tag{33.3}$$

Static spread with the Newton-Raphson method:

input: $n, C, P, S[\,1..n\,]$;
real spread, price, priceD;
integer k;
spread := 0;
for ($k = 1$ to 10) {
  price := $\sum_{i=1}^{n} C/(1 + S[\,i\,] + \text{spread})^i + 100/(1 + S[\,n\,] + \text{spread})^n$;
  if [|price − P| < 0.000001] return spread;
  priceD := $-\sum_{i=1}^{n} iC/(1 + S[\,i\,] + \text{spread})^{i+1} - 100 \times n/(1 + S[\,n\,] + \text{spread})^{n+1}$;
  spread := spread − (price − P)/priceD;
}

Figure 33.1: Static spread with the Newton–Raphson method. P is the price (as a percentage of par) of the coupon bond maturing n periods from now, C is the coupon of the bond expressed as a percentage of par per period, and $S[\,i\,]$ denotes the i-period spot rate.

Because $S(k-1) \geq y_{k-1}$ by the induction hypothesis, the term inside the braces is nonnegative. The normality assumption about the yield curve further implies that

$$\sum_{i=1}^{k-1} \frac{C}{(1+y_k)^i} < \sum_{i=1}^{k-1} \frac{C}{(1+y_{k-1})^i}.$$

Therefore

$$\frac{C+F}{(1+y_k)^k} > \frac{C+F}{[\,1+S(k)\,]^k},$$

that is, $S(k) > y_k$, as claimed. (2) We can easily confirm the two claims by inspecting Eq. (33.3).

Exercise 5.5.2: Assume that the coupon rate is 100%, $y_1 = 0.05$, $y_2 = 0.5$, $y_3 = 0.6$, and $n = 3$. The spot rates are then $S(1) = 0.05$, $S(2) = 0.82093$, and $S(3) = 0.58753$.

Exercise 5.5.3: (1) The assumption means that we can use Eq. (33.3) with $F = 0$, which, after simplification, becomes

$$\sum_{i=1}^{k} \frac{1}{(1+y_k)^i} = \sum_{i=1}^{k} \frac{1}{[\,1+S(i)\,]^i}.$$

First, we assume that the spot rate curve is upward sloping but $y_{k-1} > y_k$. The preceding equation implies that

$$\sum_{i=1}^{k-1} \frac{1}{(1+y_{k-1})^i} + \frac{1}{(1+y_k)^k} = \sum_{i=1}^{k-1} \frac{1}{[\,1+S(i)\,]^i} + \frac{1}{(1+y_k)^k} < \sum_{i=1}^{k} \frac{1}{[\,1+S(i)\,]^i}.$$

Hence $y_k > S(k)$, a contradiction. (2) The given yield curve implies the following spot rates: $S(1) = 0.1$, $S(2) = 0.89242$, and $S(3) = 0.5036$. These spot rates are not increasing.

Exercise 5.6.1: $S(1) = 0.03$, $S(2) = 0.04020$, $S(3) = 0.04538$, $f(1,2) = 0.0505$, $f(1,3) = 0.0532$, and $f(2,3) = 0.0558$.

Exercise 5.6.2: One dollar invested for $b + c$ periods at the $(b + c)$-period forward rate starting from period a is the same as one dollar invested for b periods at the b-period forward rate starting from period a and reinvested for another c periods at the c-period forward rate starting from period $a + b$.

Exercise 5.6.3: The PV of \$1 at time T, or $d(T) = [\,1 + S(T)\,]^{-T}$, is the same as that of $1 + f(T, T+1)$ dollars at time $T + 1$. Mathematically,

$$[\,1+f(T,T+1)\,]\,d(T+1) = \frac{[\,1+S(T+1)\,]^{T+1}}{[\,1+S(T)\,]^T}\,[\,1+S(T+1)\,]^{-(T+1)} = d(T).$$

A shorter proof:

$$f(T, T+1) = \frac{[\,1 + S(T+1)\,]^{T+1}}{[\,1 + S(T)\,]^{T}} - 1 = \frac{[\,1 + S(T)\,]^{-T}}{[\,1 + S(T+1)\,]^{-(T+1)}} - 1 = \frac{d(T)}{d(T+1)} - 1.$$

Exercise 5.6.4: The 10-year spot rate is 5.174%. Now we move the bond price to \$60.6. The 10-year spot rate becomes $2 \times [\,(1/0.606)^{1/20} - 1\,] = 0.05072$. The percentage change is ∼1.97%. By Exercise 5.6.3, the forward rate in question equals $f(19, 20) = 2 \times [\,\frac{0.62}{d(20)} - 1\,]$. The multiplicative factor 2 converts the forward rate into a semiannual yield. Note that each year has two periods. Simple calculations show that the forward rate moves from 6.667% to 4.620%. The percentage change is therefore 30.7%!

Exercise 5.6.5: From relation (5.3), $f(j, j+1) > S(j+1) > S(j)$, where the spot rate is upward sloping. Similarly from relation (5.4), $f(j, j+1) < S(j+1) < S(j)$, where the spot rate curve is downward sloping. And $f(j, j+1) = S(j+1) = S(j)$, where the spot rate curve is flat. See also [147, p. 400].

Exercise 5.6.6: (1) Otherwise, there are arbitrage opportunities. See [198], [234, p. 385], and [746, p. 526]. For an uncertain world, the arbitrage argument no longer holds: Although we are still able to lock in the forward rate, there is no a priori reason any future rate has to be known today for sure. (2) Because they are all realized by today's spot rate for that period according to (1) [731, p. 165].

Exercise 5.6.7: (1) From Eq. (3.6), we solve

$$1000 = \frac{1}{0.0255} \times \left[1 - \frac{1}{(1 + 0.0255)^{100}}\right] \times 27 + \frac{F}{(1 + 0.0255)^{100}}$$

for F. The answer is $F = 329.1686$. (2) They are equivalent because their cash flows match exactly. (3) Verify with Eq. (3.6) again. See [302].

Exercise 5.6.8: The probability of default is $1 - (0.92/0.94) = 0.0213$. The forward probability of default, f, satisfies $(1 - 0.0213)(1 - f) = (0.84/0.87)$. Hence $f = 0.0135$.

Exercise 5.6.9: Let $S(i)$ denote the probability that the corporation survives past time i and let $d_c(\cdot)$ denote the discount factors obtained by corporate zeros. (1) By definition, $d_c(i) = d(i)\, S(i)$. (2) As in Exercise 5.6.3, $1 + f_c(i-1, i) = d_c(i-1)/d_c(i)$. The forward probability of default for period i is

$$\begin{aligned} \frac{S(i-1) - S(i)}{S(i-1)} &= 1 - \frac{S(i)}{S(i-1)} = 1 - \frac{d_c(i)\, d(i-1)}{d(i)\, d_c(i-1)} \\ &= 1 - \frac{1 + f(i-1, i)}{1 + f_c(i-1, i)} \approx f_c(i-1, i) - f(i-1, i). \end{aligned}$$

See [583, pp. 96–97]. The forward probability of default is also called the **hazard rate** [846].

Exercise 5.6.10: The forward rate is simply the f that equates $e^{i S(i)} e^{(j-i) f} = e^{j S(j)}$.

Exercise 5.6.11:

Period (n)	*Spot Rate % Per Period*	*One-Period Forward Rate % Per Period*
1	2.00	
2	2.50	3.00
3	3.00	4.00
4	3.50	5.00
5	4.00	6.00

Exercise 5.6.12: (1) Let $S(1) \to S(1) + \Delta y$ and, in general, $S(i) \to S(i) + \Delta y / i$. We confirm that this works by inspecting Eq. (5.9). (2) From Eq. (5.5), we know the n-period zero-coupon bond costs

$$\frac{1}{[\,1 + S(1)\,][\,1 + f(1, 2)\,] \cdots [\,1 + f(n-1, n)\,]}$$

now. By the assumption, none of the forward rates in the preceding formula changes when $S(1)$ does. So $-\frac{\partial P/P}{\partial y} = \frac{1}{1+S(1)}$ for zero-coupon bonds. The same equation holds for coupon bonds as well by similar arguments. See [231, p. 53].

Exercise 5.6.13: An investor is assured that \$1 will grow to be $1 + jS(j)$ at time j. Suppose a person invests \$1 in riskless securities for i periods and, at time i, invests the proceeds in riskless securities for another $j - i$ periods ($j > i$). The implied $(j-i)$-period forward rate at time i is the f that satisfies

$$[1 + iS(i)](1 + (j-i)f) = 1 + jS(j).$$

See [731, p. 7].

Exercise 5.7.1: An n-period zero-coupon bond fetches $[1 + S(n)]^{-n}$ today and $[1 + S(k,n)]^{-(n-k)}$ at time k. The return is hence

$$\left\{ \frac{[1 + S(k,n)]^{-(n-k)}}{[1 + S(n)]^{-n}} \right\}^{1/k} - 1.$$

If the forward rate is realized, then $S(k,n) = f(k,n)$ and $[1 + S(k,n)]^{n-k} = \frac{[1+S(n)]^n}{[1+S(k)]^k}$. After substitution, we arrive at the desired result.

Exercise 5.7.2: By Eq. (5.14),

$$\begin{aligned} &E[\{1 + S(1)\}]E[\{1 + S(1,2)\}] \cdots E[\{1 + S(n-1,n)\}] \\ &\quad = [1 + S(1)][1 + f(1,2)] \cdots [1 + f(n-1,n)], \end{aligned}$$

which equals $[1 + S(n)]^n$ by Eq. (5.5).

Exercise 5.7.3: (1) Rearrange definition (5.16) to be

$$\frac{E\left[\{1 + S(1,n)\}^{-(n-1)}\right]}{1 + S(1)} = [1 + S(n)]^{-n}. \tag{33.4}$$

It implies that

$$E\left[\frac{1}{\{1 + S(1)\}\{1 + S(1,n)\}^{n-1}}\right] = [1 + S(n)]^{-n}.$$

Now recursively apply Eq. (33.4) to $[1 + S(1,n)]^{-(n-1)}$. (2) Take $n = 2$. The local expectations theory implies that $E[\{1 + S(1)\}^{-1}\{1 + S(1,2)\}^{-1}] = [1 + S(2)]^{-2}$. However, this equals $[1 + S(1)]^{-1}[1 + f(1,2)]^{-1} = [1 + S(1)]^{-1}E[1 + S(1,2)]^{-1}$ under the unbiased expectations theory. We conclude that $E[\{1 + S(1,2)\}^{-1}] = E[1 + S(1,2)]^{-1}$, which is impossible unless there is no randomness by Jensen's inequality.

Exercise 5.7.4: If they were consistent, then

$$E[\{1 + S(1,2)\} \cdots \{1 + S(n-1,n)\}]\, E\left[\frac{1}{\{1 + S(1,2)\} \cdots \{1 + S(n-1,n)\}}\right] = 1$$

from Exercise 5.7.3(1). However, this is impossible from Jensen's inequality unless there is no randomness.

Exercise 5.7.5: From the assumption, $S(1) = S(2) = \cdots = s$. Hence

$$E[S(i, i+1)] < f(i, i+1) = s = S(i).$$

Exercise 5.8.1: (1) Suppose that the yields to maturity change by the same amount and that the spot rate curve shift is parallel. We prove that the spot rate curve is flat. Consider a security with cash flows $C_i \neq 0$ at times t_i, $i = 1, 2$. In the beginning, by definition,

$$\sum_{i=1}^{2} C_i e^{-S(t_i)\, t_i} = \sum_{i=1}^{2} C_i e^{-y t_i}. \tag{33.5}$$

Suppose that the spot rate curve witnesses a parallel shift by the amount of $\Delta \neq 0$. The yield to maturity must also shift by Δ because, in particular, a t_1-period zero-coupon bond's yield to maturity

https://doi.org/10.1017/CBO9780511546839.034 Published online by Cambridge University Press

is the t_1-period spot rate (hence (2) holds, incidentally). Therefore,

$$\sum_{i=1}^{2} C_i e^{-(S(t_i)+\Delta)\, t_i} = \sum_{i=1}^{2} C_i e^{-(y+\Delta)\, t_i}.$$

The preceding two equations give rise to

$$C_1 e^{-\Delta\times t_1}[\, e^{-yt_1} - e^{-S(t_1)\, t_1}\,] = C_2 e^{-\Delta\times t_2}[\, e^{-S(t_2)\, t_2} - e^{-yt_2}\,]. \tag{33.6}$$

Because Eq. (33.6) reduces to Eq. (33.5) for $\Delta = 0$, it cannot hold for any other Δ because of the different growing rates between $e^{-\Delta\times t_1}$ and $e^{-\Delta\times t_2}$ when $t_1 \neq t_2$ unless $y = S(t_i)$ for all i. See [496, Theorem 1] for a different proof.

Exercise 5.8.2:

$$-\lim_{\Delta y\to 0} \frac{\sum_i \frac{C_i}{[1+S(i)]^i} - \sum_i \frac{C_i}{\{[1+S(i)](1+\{\Delta y/[1+S(1)]\})\}^i}}{\Delta y \sum_i \frac{C_i}{[1+S(i)]^i}} = \frac{\sum_i \frac{iC_i}{[1+S(i)]^i}}{[\,1+S(1)\,]\sum_i \frac{C_i}{[1+S(i)]^i}}.$$

Exercise 5.8.3: See [424, p. 561].

CHAPTER 6

Exercise 6.1.1: See [273, p. 196].

Exercise 6.1.2: $E[\,XY\,] = E[\,E[\,XY \mid Y\,]\,] = E[\,Y\,E[\,X \mid Y\,]\,] = E[\,Y\,E[\,X\,]\,] = E[\,X\,]\,E[\,Y\,]$. By Eq. (6.3), $\text{Cov}[\,X, Y\,] = E[\,XY\,] - E[\,X\,]\,E[\,Y\,]$; thus $\text{Cov}[\,X, Y\,] = 0$.

Exercise 6.1.3: See [75, p. 3].

Exercise 6.1.4: Simple manipulations will do [30, p. 10].

Exercise 6.1.5: Easy corollary from Eq. (6.11). See also [147, p. 16].

Exercise 6.1.6: See [492, pp. 14–15].

Exercise 6.2.1: See [273, p. 463].

Exercise 6.3.1: The estimated regression line is $0.75 + 0.32\,x$. The coefficient of determination is $r^2 = 0.966038$. On the other hand, it is easy to show, by using Eq. (6.18), that $r = 0.982872$. Finally $0.982872 \times 0.982872 = 0.966038$.

Exercise 6.4.1: (1) Square both sides and take expectations to yield

$$E\left[\,Y^2\,\right] = \text{Var}[\,X_2\,] + \alpha^2 + \beta^2\,\text{Var}[\,X_1\,] - 2\beta\,\text{Cov}[\,X_1, X_2\,].$$

We minimize the preceding equation by setting $\alpha = 0$ and $\beta = \frac{\text{Cov}[\,X_1, X_2\,]}{\text{Var}[\,X_1\,]}$. See [642, p. 19]. (2) It is because

$$\begin{aligned}\text{Cov}[\,X_1, Y\,] &= E[\,(X_1 - E[\,X_1\,])(\{\,X_2 - E[\,X_2\,]\,\} - \beta\{\,X_1 - E[\,X_1\,]\,\})\,]\\ &= \text{Cov}[\,X_1, X_2\,] - \beta\,\text{Var}[\,X_1\,] = 0.\end{aligned}$$

Exercise 6.4.2: Differentiate Eq. (6.20) with respect to $\widehat{X}$ and set it to zero. The average emerges as an extremal value. To verify that it minimizes the function, check that the second derivative equals $2n > 0$. See [417, pp. 429–430].

Exercise 6.4.3: (1) This is because

$$\begin{aligned}(X-a)^2 &= \{\,(X - E[\,X\,]) - (E[\,X\,] - a)\,\}^2\\ &= (X - E[\,X\,])^2 + 2(E[\,X\,] - a)(X - E[\,X\,]) + (E[\,X\,] - a)^2.\end{aligned}$$

Now take expectations of both sides to obtain

$$E[\,(X-a)^2\,] = \text{Var}[\,X\,] + (E[\,X\,] - a)^2,$$

which is clearly minimized at the said a value. See [195, p. 319]. (2) For any a,

$$\begin{aligned} &E[\,(X_k - a)^2 \mid X_1, \dots, X_{k-1}\,] \\ &\quad = E[\,X_k^2 \mid X_1, \dots, X_{k-1}\,] - 2a\ E[\,X_k \mid X_1, \dots, X_{k-1}\,] + a^2 \\ &\quad = E[\,(X_k - E[\,X_k \mid X_1, \dots, X_{k-1}\,])^2 \mid X_1, \dots, X_{k-1}\,] \\ &\qquad + (a - E[\,X_k \mid X_1, \dots, X_{k-1}\,])^2, \end{aligned}$$

which we minimize by choosing $a = E[\,X_k \mid X_1, X_2, \dots, X_{k-1}\,]$. See [413, p. 173].

CHAPTER 7

Exercise 7.4.1: It is bullish and defensive.

Exercise 7.4.2: The maximum profit is derived as follows. The initial cash outflow is $\mathrm{PV}(X) - P$. If the option is not exercised, then the cash grows to be X, the final cash inflow.

Exercise 7.4.3: Insurance works by diversification of risk: Fires do not burn down all insured houses at the same time. In contrast, when the market goes down, all diversified portfolios take a nose dive. See [64, p. 271].

Exercise 7.4.4: The payoff 1 year from now is clearly $100 + \alpha \times \max(S - X, 0)$, where α is the number of calls purchased. This is because the initial cost is \$100, and the fund in the money market will grow to be \$100, thus guaranteeing the preservation of capital. Because 90% of the money is to be put into the money market and the fund 1 year from now should be just sufficient to exercise the call, we have to make sure that $100 = 90 \times (1 + r)$. So $r = 11\%$.

Exercise 7.4.5: Consider the butterfly spread $C_{X-\Delta X} - 2C_X + C_{X+\Delta X}$ with the strike prices in subscript. Because the area under the spread's terminal payoff equals $(1/2)(2\Delta X)\,\Delta X = (\Delta X)^2$,

$$\frac{C_{X-\Delta X} - 2C_X + C_{X+\Delta X}}{(\Delta X)^2}$$

has area one at the expiration date. As $\Delta X \to 0$, the area is maintained at one, and the payoff function approaches the Dirac delta function. See [157].

CHAPTER 8

Exercise 8.1.1: Short the high-priced bonds and long the low-priced ones.

Exercise 8.1.2: Let $p \in \boldsymbol{R}^n$ denote the prices of the n securities. One necessary condition for arbitrage freedom is that a portfolio of securities has a nonnegative market value if it has a nonnegative payoff in every state; in other words, $p^{\mathrm{T}}\gamma \geq 0$ if $D\gamma \geq 0$. This is called **weakly arbitrage-free** [289, p. 71]. The equivalent condition for this property to hold depends on **Farka's lemma** [67, 289, 581]. Another necessary condition for arbitrage freedom is $p^{\mathrm{T}}\gamma > 0$ if $D\gamma > 0$, which says that a portfolio of securities must have a positive market value if it has a positive payoff in every state. Interestingly, the two conditions combined are equivalent to the existence of a vector $\theta > 0$ satisfying $D^{\mathrm{T}}\theta = p$. This important result is called the **arbitrage theorem**, and the m elements of θ are called the state prices [692, p. 39]. In the preceding, $v > 0$ means every element of the vector v is positive, and $v \geq 0$ means that every element of the vector v is nonnegative.

Exercise 8.2.1: We want to show that $\mathrm{PV}(X_2 - X_1) < C_{X_1} - C_{X_2}$. Suppose it is not true. Then we can generate arbitrage profits by buying C_{X_2} and shorting C_{X_1}. This generates a positive cash flow. We deposit $\mathrm{PV}(X_2 - X_1)$ in a riskless bank account. At expiration, the funds will be sufficient to cover the calls because (1) $S < X_1$: the payoff is $X_2 - X_1 > 0$; (2) $X_1 \leq S < X_2$: the payoff is $X_2 - X_1 - (S - X_1) = X_2 - S > 0$; (3) $X_2 \leq S$: the payoff is $X_2 - X_1 - (X_2 - X_1) = 0$.

Another proof is to use the put–call parity:

$$C_{X_1} = P_{X_1} + S - \mathrm{PV}(X_1),$$
$$C_{X_2} = P_{X_2} + S - \mathrm{PV}(X_2).$$

We subtract to obtain

$$C_{X_1} - C_{X_2} = P_{X_1} - P_{X_2} - \mathrm{PV}(X_1) + \mathrm{PV}(X_2) > \mathrm{PV}(X_2 - X_1).$$

Exercise 8.2.2: It is $P \leq X$ if we consider a cash-secured put. An alternative is to observe that, because noone will put money in the bank, the interest rate can be treated as if it were zero and $\mathrm{PV}(X) = X$. Lemma 8.2.4 is thus applicable, and we have $P \leq X$.

Exercise 8.3.1: (1) The put–call parity shows the threshold interest to be $r = 12.905\%$ because it satisfies $95 \times e^{-r/3} = 3 + 94 - 6 = 91$. Because $r > 10\%$, $C - P > S - \text{PV}(X)$. We can create an arbitrage profit by shorting the call, buying the put, buying the stock, and lending $\text{PV}(X)$. (2) Apply the put–call parity $\text{PV}(X) = P + S - C$. See [346].

Exercise 8.3.2: Consider the payoffs from the following two portfolios.

		Value at Expiration Date	
	Initial Investment	$S_1 > X$	$S_1 \leq X$
Buy a call	$-C$	$S_1 - X$	0
Buy bonds	$-\text{PV}(X)$	X	X
Total	$-C - \text{PV}(X)$	S_1	X
Buy stock	$-S_0$	S_1	S_1

(S_0 and S_1 denote the stock prices now and at the expiration date, respectively.) The table shows that whichever case actually happens, the first portfolio is worth at least as much as the second. It therefore cannot cost less, and $C + \text{PV}(X) \geq S_0$. (A simpler alternative is to use the put–call parity.) American calls on a non-dividend-paying stock cannot be worth less than European ones. See [317, p. 577].

Exercise 8.3.3: The $C \geq S - \text{PV}(X)$ inequality is derived under the no-arbitrage condition. Apparently, not every stock price series is arbitrage free. For instance, $S > \text{PV}(X)$ is not for, otherwise, we can sell short the stock and invest the proceeds in riskless bonds, and at the option's expiration date, we close out the short position with X from the bonds. Margin requirements ignored, this is doable because we already assume that the stock price at expiration is less than the strike price.

Exercise 8.3.4: Consider the following portfolio: one short call, one long put, one share of stock, a loan of $\text{PV}(X)$ maturing at time t, and a loan of $D^* d(t_1)$ maturing at time t_1. The initial cash flow is $C - P - S + \text{PV}(X) + D^* d(t_1)$. The loan amount $D^* d(t_1)$ will be repaid by the dividend. The rest of the argument replicates that for the put–call parity at expiration. See [746, p. 148].

Exercise 8.3.5: It is equivalent to a long European call with strike price X and a short European call with exercise H, a vertical spread in short [111].

Exercise 8.3.6: Let the payoff function be

$$F(S) \equiv \begin{cases} 0, & \text{if } S < 0 \\ \alpha_i S + \beta_i, & \text{if } S_i \leq S < S_{i+1} \quad \text{for } 0 \leq i < n, \\ \alpha_n S + \beta_n, & \text{if } S_n \leq S \end{cases}$$

where $0 = S_0 < S_1 < \cdots < S_n$ are the breakpoints, $\alpha_{i-1} S_i + \beta_{i-1} = \alpha_i S_i + \beta_i$ for continuity, and $\beta_0 = 0$ for origin crossing. Clearly, $F(0) = 0$ and $F(S_i) = \sum_{j=1}^{i} \alpha_{j-1}(S_j - S_{j-1})$ for $i > 0$.

A generalized option can be replicated by a portfolio of α_0 European calls with strike price $S_0 = 0$, $\alpha_1 - \alpha_0$ European calls with strike price S_1, $\alpha_2 - \alpha_1$ European calls with strike price S_2, and so on, all with the same expiration date. When the stock price S finishes between S_i and S_{i+1}, the option has the payoff

$$F(S) = F(S_i) + \alpha_i(S - S_i) = \sum_{j=1}^{i} \alpha_{j-1}(S_j - S_{j-1}) + \alpha_i(S - S_i).$$

Among the options in the package, only those with the strike price not exceeding S_i finish in the money. The payoff is thus

$$\alpha_0(S - S_0) + \sum_{j=1}^{i} (\alpha_j - \alpha_{j-1})(S - S_j) = \alpha_i(S - S_i) + \sum_{j=1}^{i} \alpha_{j-1}(S_j - S_{j-1}),$$

the same as that of the generalized option.

For a payoff function that does not pass through the origin, say with an intercept of β, we add zero-coupon bonds with a total obligation in the amount of β at the expiration date to the portfolio. In general, when the payoff function could be any continuous function, we can use a piecewise linear function with enough breakpoints to approximate the payoff to the desired accuracy. See [236, pp. 371ff].

Exercise 8.4.1: From Exercise 8.3.2, the option value is at least $45 - 40 \times e^{-0.08/6} > 45 - 40 = 5$ dollars per share. Because the stock is considered overpriced and there is no dividend, the intrinsic value of the option will never exceed \$5. Furthermore, other things being equal, the call becomes less valuable as its maturity approaches. Hence the investor should sell the option.

Exercise 8.4.2: In short, $X - \mathrm{PV}(X) > D$. Take a date just before an ex-dividend date. If a call holder exercises the option, the holdings just after that date will be worth $S - X + D'$ with D' denoting the dividend for holding the stock through the ex-dividend date. Note that $D' \leq D$. If the holder chooses not to exercise the call, on the other hand, the holdings will then be worth by definition C after the dividend date. From Eq. (8.2) we conclude that

$$C \geq S - \mathrm{PV}(X) - (D - D') > S - X + D'.$$

Hence it is better to sell the call than to exercise it just before an ex-dividend date. Combine this conclusion with Theorem 8.4.2 to prove the result. See [236, p. 140].

Exercise 8.4.3: Because it is worse than exercising the option just after the ex-dividend date. More formally, let S be the stock price immediately before an ex-dividend date, let S' be the stock price immediately after an ex-dividend date, and let D be the amount of the dividend. S' should be $S - D$ within so short a time interval. So exercise immediately after an ex-dividend date fetches $X - S' = X - S + D$, and exercise immediately before an ex-dividend date fetches $X - S$. The interest gained from $X - S$ in such a short period of time can be ignored. The late-exercise strategy clearly dominates [236, p. 251].

Exercise 8.4.4: This inequality says that exercising this option now and investing the proceeds in riskless bonds fetches a terminal value exceeding X. However, Lemma 8.2.4 says that the American put can never be worth more than X. Exercising it is hence better.

Exercise 8.4.5: (1) The covered call strategy guarantees arbitrage profits otherwise. (2) Combine (1) with the put–call parity. (3) Lemma 8.2.4 says that $P \leq X$. The case of $P < X$ can never happen because it would imply that the put is selling at less than its intrinsic value.

Exercise 8.4.6: Assume that $C - P < S - X$. Write the put, buy the call, sell the stock short (hence the need for the no-dividend assumption), and place X in a bank account. This generates a positive net cash flow. If the short put is exercised before expiration, withdraw the money from the bank account to pay for the stock, which is then used to close out the short position.

Exercise 8.6.1: Although the floor is the same, the portfolio of options offers a higher payoff when the terminal stock prices are such that some, but not all, put options finish in the money.

CHAPTER 9

Exercise 9.2.1: Suppose that $R > u$. Selling the stock short and investing the proceeds in riskless bonds for one period creates a pure arbitrage profit. A similar argument can be made for the $d > R$ case.

Exercise 9.2.2: Let P denote B's current price. Consider a portfolio consisting of one unit of A and h units of B, worth $100 + Ph$ now. It can fetch either $160 + (50 \times h)$ or $80 + (60 \times h)$ in a period. Pick $h = 8$ to make them both equal to 560. Note that h is simply the delta, because A's price in a period has a range of $160 - 80 = 80$ vs. B's $60 - 50 = 10$, a ratio of eight. Now that this particular portfolio's FV is no longer random, its PV must be $560/e^{0.1} = 506.71$. Hence $P = (506.71 - 100)/8 \approx 50.84$. The above methodology is apparently general.

Exercise 9.2.4: Suppose that $k > 0$ (the other case is symmetric). Sell short M/k options and use the proceeds to buy $(M/k)\,h$ shares of stock and $(M/k)\,B$ dollars in riskless bonds. This transaction nets a current value of

$$\frac{M}{k}\,(hS + B + k) - \frac{Mh}{k}\,S - \frac{MB}{k} = M.$$

The obligations after one period are nil because, modulo a multiplicative factor of M/k, the preceding levered position above is h shares of stock and M dollars in riskless bonds, which replicate the option. See [289, p. 7].

Exercise 9.2.5: The expected value of the call in a risk-neutral economy one period from now is $pC_u + (1-p)\,C_d$. After being discounted by the riskless interest rate, the call value now is $\frac{pC_u + (1-p)\,C_d}{R}$, which is equal to $C = hS + B$ by Eq. (9.3).

Exercise 9.2.6: If $Su \leq X$, then $S - X < 0$ but $C = 0 > S - X$. If $X \leq Sd$, then

$$C = \frac{p(Su - X) + (1-p)(Sd - X)}{R} = S - \frac{X}{R} > S - X.$$

Finally, if $Sd < X < Su$, then $C = p(Su - X)/R$, which exceeds $S - X$ because

$$C = \frac{(R-d)(Su - X)}{R(u-d)} = \frac{Ru - ud}{Ru - Rd} S - \frac{R-d}{Ru - Rd} X > S - X.$$

See [236, p. 173]. A more compact derivation is to observe that, from Eq. (9.4),

$$hS + B = \frac{pC_u + (1-p)\,C_d}{R} \geq \frac{p(Su - X) + (1-p)(Sd - X)}{R} = S - \frac{X}{R} > S - X.$$

Exercise 9.2.7: This fact can be easily seen from Eq. (9.1) and the inductive steps in deriving the deltas.

Exercise 9.2.8: The stock price should grow at the riskless rate to SR^n at expiration in a risk-neutral economy. As $u > R > d$, we must have $R \to d$ as well. Now the formula says that $C = 0$ if $\text{PV}(X) = XR^{-n} > S$. This makes sense because $X > Sd^n$ and the call will finish out of the money. On the other hand, the formula says that $C = S - \text{PV}(X)$ if $\text{PV}(X) \leq S$. This also makes sense as the call will have a terminal payoff of $SR^n - X$.

Exercise 9.2.10: (1) Consider the butterfly spread with strike prices X_L, X_M, and X_H such that

$$\begin{aligned} Su^{i-1}d^{n-i+1} &< X_L < Su^i d^{n-i}, \\ X_M &= Su^i d^{n-i}, \\ Su^i d^{n-i} &< X_H < Su^{i+1}d^{n-i-1}, \end{aligned}$$

with $2X_M - X_H - X_L = 0$. This portfolio pays off $Su^i d^{n-i} - X_L$ dollars when the stock price reaches $Su^i d^{n-i}$. Furthermore, its payoff is zero if the stock price finishes at other prices. (2) The claim holds because calls can be replicated by continuous trading. Note that the continuous trading strategy described in Subsection 9.2.1 does *not* require that the strike price be one of the possible $n+1$ stock prices.

Exercise 9.2.11: For p to be a risk-neutral probability, both securities must earn an expected return equal to R, or

$$p = \frac{R - d_1}{u_1 - d_1} = \frac{R - d_2}{u_2 - d_2}.$$

It is not hard to pick R, u_1, d_1, u_2, and d_2 such that the preceding identity does not hold. See [836, p. 91].

Programming Assignment 9.2.14: The limited precision of digital computers dictates that we compute and store $\ln b(j; n, p)$ instead of $b(j; n, p)$. The needed changes to the algorithm are

$$\begin{aligned} &\cdots \\ &b := \ln n! - \ln a! - \ln(n-a)! + a \times \ln p + (n-a) \times \ln(1-p); \\ &\cdots \\ &1.\ b := b + \ln p + \ln(n - j + 1) - \ln(1-p) - \ln j; \\ &\cdots \\ &3.\ C := C + e^b \times (D - X)/R; \\ &\cdots \end{aligned}$$

Exercise 9.3.1: (1) This approach, which is due to Jarrow and Rudd [690], differs from the one in the text in that we do not derive the risk-neutral probability [346, p. 546]. See [346, pp. 103–104, p. 544], [289, p. 146], and [290, p. 197]. (2) Theoretically, the risk-neutral probability formula $p \equiv (R-d)/(u-d)$ with the new choices of u and d should be used as is required by the replication argument. In fact, Lemma 9.3.3 will imply that the risk-neutral probability simplifies to

$$\frac{e^{\sigma^2(\tau/n)/2} - e^{-\sigma\sqrt{\tau/n}}}{e^{\sigma\sqrt{\tau/n}} - e^{-\sigma\sqrt{\tau/n}}} \to \frac{1}{2}.$$

However, if we treat the discrete-time economy as an approximation to its continuous-time limit and are concerned about its behavior only as $n \to \infty$, either p or $1/2$ should work in the limit.

Exercise 9.3.2: Use Eq. (9.15) with $n \to \infty$.

Exercise 9.3.4: Observe that

$$p = \frac{\left[1 + \frac{r\tau}{n} + \left(\frac{r\tau}{n}\right)^2/2! + \cdots\right] - \left[1 - \sigma\sqrt{\frac{\tau}{n}} + \left(\sigma\sqrt{\frac{\tau}{n}}\right)^2/2! - \cdots\right]}{\left[1 + \sigma\sqrt{\frac{\tau}{n}} + \left(\sigma\sqrt{\frac{\tau}{n}}\right)^2/2! + \cdots\right] - \left[1 - \sigma\sqrt{\frac{\tau}{n}} + \left(\sigma\sqrt{\frac{\tau}{n}}\right)^2/2! - \cdots\right]}.$$

Exercise 9.3.5: The probability for $S_\tau \geq X$ equals that for $\ln(S_\tau/S) \geq \ln(X/S)$. By Lemma 9.3.3, this event occurs with probability

$$\begin{aligned} 1 - N\left(\frac{\ln(X/S) - (r - \sigma^2/2)\,\tau}{\sigma\sqrt{\tau}}\right) &= N\left(-\frac{\ln(X/S) - (r - \sigma^2/2)\,\tau}{\sigma\sqrt{\tau}}\right) \\ &= N\left(-\frac{\ln(S/X) + (r - \sigma^2/2)\,\tau}{\sigma\sqrt{\tau}}\right). \end{aligned}$$

Multiply it by $e^{-r\tau}$ to get the desired result. See [495].

Exercise 9.3.6: The put–call parity says that $C - P = S - \mathrm{PV}(X)$. The exercise clearly holds because the right-hand side equals $S - Xe^{-rT}$. See also [853].

Exercise 9.3.7: Lemma 9.2.1 says that the call value equals $e^{-r\tau}\, E^\pi[\max(S_\tau - X)]$, where $\ln S_\tau$ is normally distributed with mean $\ln S + (r - \sigma^2/2)\,\tau$ and variance $\sigma^2\tau$ by Lemma 9.3.3. The expectation for European calls is

$$\begin{aligned} & e^{-r\tau} \int_X^\infty (S - X)\, f(S)\, dS \\ &= e^{-r\tau} \int_X^\infty S\, f(S)\, dS - e^{-r\tau} X \int_X^\infty f(S)\, dS \\ &= e^{-r\tau} e^{\ln S + (r - \sigma^2/2)\,\tau + \sigma^2\tau/2} N\left(\frac{\ln S + (r - \sigma^2/2)\,\tau - \ln X}{\sigma\sqrt{\tau}} + \sigma\sqrt{\tau}\right) \\ &\quad - e^{-r\tau} XN\left(\frac{\ln S + (r - \sigma^2/2)\,\tau - \ln X}{\sigma\sqrt{\tau}}\right) \\ &= SN\left(\frac{\ln(S/X) + (r + \sigma^2/2)\,\tau}{\sigma\sqrt{\tau}}\right) - e^{-r\tau} XN\left(\frac{\ln(S/X) + (r - \sigma^2/2)\,\tau}{\sigma\sqrt{\tau}}\right), \end{aligned}$$

as desired. The second equality was due to Exercise 6.1.6 and Eq. (6.12).

Exercise 9.3.8: (1) This choice has the correct mean stock price $Se^{r\tau/n}$. The second moment also converges to that under the standard choice in the text. (2) When n is even, $Su^{n/2}d^{n/2} = X$, which places the strike price at the center of the tree at maturity. When n is odd, $Su^{(n+1)/2}d^{(n-1)/2} = Xe^{\sigma\sqrt{\tau/n}}$ and $Sd^{(n+1)/2}u^{(n-1)/2} = Xe^{-\sigma\sqrt{\tau/n}}$; the strike price is therefore between the two middle nodes of the tree at maturity. See [555, 589].

Exercise 9.4.3: (1) There are $m+2$ equations: the discounted expected payoffs of the options (m equations), the discounted expected stock price (one equation), and the summing of the terminal probabilities to one (one equation). So we can divide the time into $m+1$ periods, creating $m+2$ terminal nodes. (2) We solve for the terminal nodes' probabilities first by using the above-mentioned $m+2$ equations. Because each path leading up to the same node is equally likely, we divide each nodal probability by the number of paths leading up to that node for the path probability. Finally, we use backward induction to solve for the rest of the probabilities on the tree as follows. We assume that node A is followed by nodes B and C, which have path probabilities p_{B} and p_{C}, respectively. The path probability for node A is $p_{\mathrm{B}} + p_{\mathrm{C}}$, and the transition probability from node A to B is $p_{\mathrm{B}}/(p_{\mathrm{B}} + p_{\mathrm{C}})$. See [502, 503]. An algorithm for constructing the implied binomial tree under a more general setting is considered in [269].

Exercise 9.6.1: From the text we know the value can be computed by replacing the current stock price S with $(1-\delta)^m S$. But this effectively implies a payoff function of

$$\max\left((1-\delta)^m S - X\right) = (1-\delta)^m \times \max(S - (1-\delta)^{-m} X).$$

See [879, p. 97].

Exercise 9.6.2: No. When the American call is exercised, there is always a corresponding European call to exercise with the same payoff. Because there may be European calls remaining after the American call is exercised, the package is potentially more valuable than the American call.

Exercise 9.6.5: We know that $C = P + S - D - \text{PV}(X)$ from Eq. (8.2). Now, $S - D = Se^{-q\tau}$ because the stock price would be $S^{-q\tau}$ today to reach the same terminal price if there were no dividends. See [894, p. 72] for the formula.

Exercise 9.6.6: Just follow the same steps as before in setting up the replicating portfolio except that, now,

$$hSue^{\widehat{q}} + RB = C_u, \qquad hSde^{\widehat{q}} + RB = C_d.$$

The reason is that the stockholder is getting new shares at a rate of q per year. (Throughout this exercise, $\widehat{q} \equiv q\Delta t$ means the dividend yield per period.) Equations (9.1) and (9.2) are hence replaced with

$$h = \frac{C_u - C_d}{(Su - Sd)\, e^{\widehat{q}}}, \qquad B = \frac{uC_d - dC_u}{(u-d)\, R}.$$

After substitution and rearrangement,

$$hS + B = \left(\frac{Re^{-\widehat{q}} - d}{u - d}\, C_u + \frac{u - Re^{-\widehat{q}}}{u - d}\, C_d \right) \Big/ R$$

in place of Eq. (9.3). Finally, Eq. (9.4) becomes $hS + B = [\, pC_u + (1-p)\, C_d\,]/R$, where $p \equiv (Re^{-\widehat{q}} - d)/(u - d)$.

Another way to look at it is by observing that in a risk-neutral economy, the per-period return of holding the stock should be R. Now, because the stock is paying a dividend yield of $\widehat{q}$, the return of the stock price net of the dividends should be $Re^{-\widehat{q}}$.

Exercise 9.6.7: (1) The total *wealth* is growing at μ, not $\mu - q$! (2) It retains the binomial tree's backward-induction structure except that each S at j periods from now should be replaced with $Se^{-qj\Delta t}$. Of course, the result is not exactly the same as using Eq. (9.21) and pretending there were no dividends. However, it should converge to the same value. Interestingly, this is the same as retaining $(e^{r\Delta t} - d)/(u - d)$ and the original algorithm as if there were no dividends but with u and d multiplied by $e^{q\Delta t}$.

Exercise 9.6.8: Suppose there is one period to expiration and $Sue^{-\widehat{q}} < X$, where $\widehat{q}$ is the dividend yield per period. Clearly the option has zero value at present as it will not be exercised at expiration. However, if $S > X$, the option has positive intrinsic value, which means it should be exercised now. These two inequalities imply that $X < S < Xu^{-1}e^{\widehat{q}}$, which is possible when $u < e^{\widehat{q}}$.

Exercise 9.7.1: Consult [154, 156], [575, Subsection 4.1.4], or [783].

Exercise 9.7.2: Suppose the option is not exercised at price S but it is optimal to exercise it at the same price two periods earlier. It must hold that

$$\begin{aligned} e^{r\Delta t}(X - S) &\le pP_u + (1-p)\, P_d, \\ e^{r\Delta t}(X - S) &> pP'_u + (1-p)\, P'_d. \end{aligned}$$

(Primed symbols are for values two periods earlier.) Hence,

$$pP'_u + (1-p)\, P'_d < pP_u + (1-p)\, P_d.$$

However, $P_d \le P'_d$ and $P_u \le P'_u$ by Lemma 8.2.1 because of the longer maturity. That lemma works whether there are dividends or not. We now have the contradiction $pP'_u + (1-p)P'_d < pP'_u + (1-p)P'_d$.

CHAPTER 10

Exercise 10.1.1: Note that $N'(y) = e^{-(y^2/2)}/\sqrt{2\pi}$ and $x' \equiv \partial x/\partial S$ in the following equation:

$$\begin{aligned} \partial C/\partial S &= N(x) + SN'(x)\, x' - Xe^{-r\tau} N'(x - \sigma\sqrt{\tau})\, x' \\ &= N(x) + SN'(x)\, x' - Xe^{-r\tau} N'(x)\, e^{x\sigma\sqrt{\tau} - \sigma^2\tau/2} x' \\ &= N(x) + SN'(x)\, x' - Xe^{-r\tau} N'(x)\, e^{\ln(S/X) + r\tau} x' = N(x). \end{aligned}$$

Exercise 10.1.2: Note that $N'(y) = e^{-(y^2/2)}/\sqrt{2\pi}$ and $x' \equiv \partial x/\partial X$ in the following equation:

$$\begin{aligned}\partial P/\partial X &= e^{-r\tau} N(-x+\sigma\sqrt{\tau}) - Xe^{-r\tau} N'(-x+\sigma\sqrt{\tau})\, x' + SN'(-x)\, x' \\ &= e^{-r\tau} N(-x+\sigma\sqrt{\tau}) - Xe^{-r\tau} N'(-x)\, e^{x\sigma\sqrt{\tau}-\sigma^2\tau/2}\, x' + SN'(-x)\, x' \\ &= e^{-r\tau} N(-x+\sigma\sqrt{\tau}) - Xe^{-r\tau} N'(-x)\, e^{\ln(S/X)+r\tau}\, x' + SN'(-x)\, x' \\ &= e^{-r\tau} N(-x+\sigma\sqrt{\tau}).\end{aligned}$$

Exercise 10.1.3: We have to prove that the strike price X that maximizes the option's time value is the current stock price S. Recall that the time value is defined as $V \equiv C - \max(S - X, 0)$. It is not hard to verify that $\partial C/\partial X = -e^{-r\tau} N(x - \sigma\sqrt{\tau})$. So $\partial V/\partial X = \partial C/\partial X < 0$ if $X > S$, and $\partial V/\partial X = (\partial C/\partial X) + 1 > 0$ if $X \leq S$. The time value is hence maximized at $X = S$. The case of puts is similar.

Exercise 10.1.4: It is $\frac{r\tau + \sigma^2\tau/2 - \ln(S/X)}{2\sigma\tau\sqrt{\tau}}$ [894, p. 78].

Exercise 10.1.5: (1) It is $S = Xe^{(r+\sigma^2/2)\tau}$ (not $S = X$). To derive it, note that

$$\begin{aligned}\frac{\partial\Theta}{\partial S} &= -\frac{N'(x)\,\sigma + SN''(x)\,x'\sigma}{2\sqrt{\tau}} - r\,Xe^{-r\tau} N'(x - \sigma\sqrt{\tau})\, x' \\ &= -\frac{N'(x)\,\sigma + SN''(x)\,x'\sigma}{2\sqrt{\tau}} - r\,SN'(x)\, x'.\end{aligned}$$

The last equality above takes advantage of $Xe^{-r\tau} N'(x - \sigma\sqrt{\tau})\, x' = SN'(x)\, x'$, which can be verified with Eq. (10.1) and the Black–Scholes formula for the European call. With $N''(x) = -x N'(x)$ and $x' = 1/(S\sigma\sqrt{\tau})$, it is not hard to see that $\partial\Theta/\partial S = 0$ if and only if $-\sigma^2 + (x\sigma/\sqrt{\tau}) - 2r = 0$. From here, our claim follows easily. (2) This is by virtue of Lemma 8.2.1.

Exercise 10.1.6: Vega's derivative with respect to σ is $-xx'\Lambda$. Note that x can be expressed as $(A/\sigma) + B\sigma$. Hence $x'x = -2(A^2/\sigma^3) + 2B^2\sigma$. Clearly, $x'x = 0$ has a positive solution at $\sigma = \sigma^* \equiv \sqrt{|\,A/B\,|}$. It is not hard to see that Λ' begins at ∞ for $\sigma = 0$, penetrates the x axis at σ^* into the negative domain, and converges to zero at positive infinity. This confirms the unimodality of vega.

Exercise 10.1.7: (1) $-\frac{x+\sigma\sqrt{\tau}}{S\sigma\sqrt{\tau}}\,\Gamma$ [894, p. 78]. (2) $-\frac{\sigma\sqrt{\tau}+(\ln(X/S)+(r+\sigma^2/2)\,\tau)\,x}{2\sigma^2\tau^2 S}\, N(x)$. See [894, p. 78].

Exercise 10.2.7: The standard finite-difference scheme approximates the second derivative by using the function values at the three equally spaced stock prices $S - \Delta S$, S, and $S + \Delta S$. It purports to improve the accuracy by varying ΔS. Scheme (10.2) instead approximates the second derivative at the *fixed* prices Suu, S, and Sdd, and it improves the accuracy by varying n.

CHAPTER 11

Exercise 11.1.1: The guarantee generates a cash flow to the bondholders of $-\min(0, V^* - B)$, which is equal to $\max(0, B - V^*)$ [660, p. 630].

Exercise 11.1.3: Leibniz's rule might be useful [448]: If $F(x) \equiv \int_{a(x)}^{b(x)} f(x, z)\, dz$, then

$$F'(x) = \int_{a(x)}^{b(x)} \frac{\partial f(x, z)}{\partial x}\, dz + f(x, b(x))\, b'(x) - f(x, a(x))\, a'(x).$$

(1) The higher the firm value, the higher the bond price is. Intuitively, the higher the firm value, the less likely the firm is to default. (2) The more the firm borrows, the higher the bond price. However, note that the increase in the total bond value is less than the net increase in the face value. (3) The longer the time to maturity, the lower the bond price is. This is because the PV decreases and the default premium increases. See [746, p. 396].

Exercise 11.1.4: Any other bond price will lead to arbitrage profits by trading Merck stock, Merck calls, and XYZ.com's bonds.

Exercise 11.1.5: The stockholders gain $[\,35000 \times (5/35)\,] + 9500 - 15250 = -750$ dollars. The original bondholders lose the equal amount, $29250 - [\,(30/35) \times 35000\,] = -750$. So the bondholders gain.

Exercise 11.1.6: The results are tabulated below.

Promised Payment to Bondholders X	*Current Market Value of Bonds* B	*Current Market Value of Stock* nS	*Current Total Value of Firm* V	*Stockholders' Gains from Issuing* $\frac{X-30000}{1000}$ *Bonds*
55,000	54,375	16,750	71,125	0
60,000	59,375	11,750	71,125	−52.083
65,000	64,125	7,000	71,125	115.385
70,000	68,000	3,125.5	71,125	946.429

Exercise 11.1.7: Dividends reduce the total value by the equal amount. However, the stock, as an option on the total value, decreases by less than the dividend amount because of the convexity of option value by Lemma 8.5.1. See [424, p. 531].

Programming Assignment 11.1.8: (2) Guess a $V > nS$ and build the binomial tree for V/n. Price an American call with a strike price of X on that tree. Each warrant W is then $n/(n+m)$ of the call value. Stop if $V \approx nS + mW$. Otherwise, choose a new $V > nS$ and repeat the procedure.

Exercise 11.1.9: It is

$$V - C(X) + \lambda C(X/\lambda), \tag{33.7}$$

where $C(Y)$ represents a European call with a strike price of Y.

Exercise 11.1.10: (1) From expression (33.7) and that $V - C(X)$ is equivalent to a zero-coupon bond with a face value of X, a CB is basically a zero-coupon bond plus a call on λ times the total value of the corporation with a strike price of X/λ [746, p. 401]. (2) Warrants are calls (see Subsection 11.1.2). (3) By analyzation of the payoff at maturity, the terminal value is easily seen to be at least λV^* and sometimes more. On the other hand, conversion grants the owner a fraction $1/\lambda$ of the firm. The former strategy therefore dominates the latter. See [491, Theorem 1], [492, p. 430], and [697].

Exercise 11.1.11: In all three cases, the PVs are either less than or equal to P:

$$V < \text{PV}(X) \le P,$$
$$\text{PV}(X) \le X \le P,$$
$$\lambda V < P.$$

Exercise 11.2.1: (1) See [470, p. 463]. (2) See [470, p. 464].

Programming Assignment 11.2.5: Assume that the barrier option is a call in the following discussions. Our approach computes down-and-in and down-and-out options simultaneously, noting that they sum to the standard European call value at each node by the in–out parity.

Working from the terminal nodes toward the root, backward induction calculates the option value at each node *if the derivative is issued at that node*. Each node keeps two values: The in-value records the value of the down-and-in option and the out-value records the value of the down-and-out option. Each terminal node above the strike price starts with the payoff of the standard call option in the out-value and zero in the in-value, whereas each terminal node at or below the strike price starts with zeros in both values. Inductively, a node not on the barrier takes the expected PV of the in-values of its two successor nodes and that of the out-values of its two successor nodes and puts them into its respective value cells. For a node on the barrier, we do the same thing and then set the out-value to zero and the in-value to the sum of the in-value and out-value. The overall space requirement is linear, and the time complexity is quadratic. When there is a rebate for the down-and-out call, the rebate's cash flow has to be calculated *separately*. See Fig. 33.2 for the algorithm.

Exercise 11.2.6: It can be replicated as

$$D(X_0, H_1) + \sum_{i=1}^{n-1} [\, D(X_i, H_{i+1}) - D(X_i, H_i)\,].$$

Here $D(X, H)$ denotes a down-and-out option with strike price X and barrier H. This can be justified as follows. If the nearest barrier H_1 is never hit, then the first down-and-out option provides the necessary payoff and each term in the summation is zero because it contains two options with

Binomial tree algorithm for pricing down-and-out and down-and-in calls on a non-dividend-paying stock:

```
input:    S, u, d, X, H (H < X, H < S), n, r̂, K;
real      R, p, Co[n+1], Ci[n+1], Cr[n+1];
integer   i, j, h;
R := e^r̂; p := (R − d)/(u − d);
h := ⌊ln(H/S)/ln u⌋; H := Su^h;
for (i = 0 to n) {
          Co[i] := max(0, Su^(n−i) d^i − X);
          Ci[i] := 0;
          Ci[r] := 0;
}
if [n − h is even and 0 ≤ (n − h)/2 ≤ n]   {//A hit.
          Ci[(n − h)/2] := Co[(n − h)/2];
          Co[(n − h)/2] := 0;
          Cr[(n − h)/2] := K;
}
for (j = n − 1 down to 0) {
          for (i = 0 to j) {
             Co[i] := (p × Co[i] + (1 − p) × Co[i + 1])/R;
             Ci[i] := (p × Ci[i] + (1 − p) × Ci[i + 1])/R;
             Cr[i] := (p × Cr[i] + (1 − p) × Cr[i + 1])/R;
          }
          if [j − h is even and 0 ≤ (j − h)/2 ≤ j]   {// A hit.
             Ci[(j − h)/2] := Ci[(j − h)/2] + Co[(j − h)/2];
             Co[(j − h)/2] := 0;
             Cr[(j − h)/2] := K;
          }
}
return Co[0] + Cr[0], Ci[0];
```

Figure 33.2: Binomial tree algorithm for barrier calls on a non-dividend-paying stock. Because H may not correspond to a legal stock price, we lower it to Su^h, the highest stock price not exceeding H. The new barrier corresponds to $C_x[(j-h)/2]$ at times $j = n, n-1, \ldots, h$, where $x \in \{$"o", "i", "r"$\}$. The knock-out option provides a rebate K when the barrier is hit.

identical strike price. If H_1 is hit, then $D(X_0, H_1)$ and $D(X_1, H_1)$ are rendered worthless. The remaining portfolio is

$$D(X_1, H_2) + \sum_{i=2}^{n-1} [\, D(X_i, H_{i+1}) - D(X_i, H_i)\,].$$

The rest follows inductively. See [158, p. 16].

Exercise 11.4.1: See [95, p. 288].

Exercise 11.5.1: Buy a call with strike X_H and sell a put with strike X_L [346, p. 297].

Exercise 11.5.2: Buy a call with strike $X+p$, sell a put with strike $X+p$, and buy a put with strike X [346, p. 299].

Exercise 11.5.3: Buy a call with strike X and sell α puts with strike X [346, p. 302].

Exercise 11.5.4: (1) Use Eq. (9.21) [346, p. 165]. (2) Suppose we hold $\$h$ foreign riskless bond (denominated in foreign currency) and $\$B$ domestic riskless bond. The portfolio has the value $hS+B$ in domestic currency. Here S denotes the current domestic/foreign exchange rate. The same argument as that in Subsection 9.2.1 leads to the following equations:

$$he^{\widehat{r}\Delta t}Su + RB = C_u, \qquad he^{\widehat{r}\Delta t}Sd + RB = C_d.$$

The $e^{\widehat{r}\Delta t}$ term arises from foreign interest income. Solving them gives the result. See [514, p. 321].

Exercise 11.5.5: Use $N(x) = 1 - N(-x)$ and $r = \widehat{r}$.

Exercise 11.5.6: It holds when $S \gg X$ [746, p. 345].

Exercise 11.5.7: (1) See [746, p. 332] and Eq. (8.1). (2) See [746, p. 333] and Lemma 8.3.2.

Exercise 11.6.1: The stockholders, by paying the first interest payment, acquire the option to own the firm by making future interest and principal payments. See [424, p. 531] and [746, pp. 398–400].

Exercise 11.6.2: Because it can be exercised only when the contract is awarded [346, p. 305].

Exercise 11.6.3: See [879, p. 201].

Exercise 11.7.1: Construct a portfolio consisting of a lookback call on the minimum and a lookback put on the maximum [388].

Exercise 11.7.2: Let $r_k \equiv S_k/S_{k-1}$. Then $r_k \in \{u, d\}$, depending on whether the kth move is up or down. As $S_i = S_0 \prod_{j=1}^{i} r_j$, we have

$$\prod_{i=0}^{n} S_i = S_0^{n+1} \prod_{i=1}^{n} \prod_{j=1}^{i} r_j = S_0^{n+1} \prod_{i=1}^{n} r_i^{n-i+1}.$$

Our problem thus amounts to counting the distinct numbers (call it N) the expression $\prod_{i=1}^{n} r_i^{n-i+1}$ can take for $r_i \in \{u, d\}$; equivalently, we can ask the same question of $\sum_{i=1}^{n}(n-i+1)\ln r_i$. This implies that N equals the number of distinct sums of integers drawn from $\{1, 2, \ldots, n\}$. Because it is easy to see that every integer from 1 to $n(n+1)/2$ can be represented uniquely as the sum of distinct integers drawn from $\{1, 2, \ldots, n\}$, we conclude that $N = n(n+1)/2$. The preceding analysis did not assume that $ud = 1$.

Exercise 11.7.3: See [147, p. 382].

Exercise 11.7.4: Let the historical average from m prices be A as of time zero. The terminal payoff for a call is then

$$\begin{aligned}
&\max\left(\frac{mA + \sum_{i=0}^{n} S_i}{m+n+1} - X, 0\right)\\
&\quad = \max\left(\frac{\sum_{i=0}^{n} S_i}{m+n+1} - \left(X - \frac{mA}{m+n-1}\right), 0\right)\\
&\quad = \frac{n+1}{m+n+1} \times \max\left(\frac{\sum_{i=0}^{n} S_i}{n+1} - \frac{m+n+1}{n+1}\left(X - \frac{mA}{m+n-1}\right), 0\right).
\end{aligned}$$

So it becomes $\frac{n+1}{m+n+1}$ option with strike price $\frac{m+n+1}{n+1}\,(X - \frac{mA}{m+n-1})$.

Exercise 11.7.5: (1) In a risk-neutral economy, $E[\,S_i\,] = S_0 e^{\widehat{r}i}$. So the expected average price is

$$\frac{S_0}{n+1}\sum_{i=0}^{n} e^{\widehat{r}i} = \begin{cases} \frac{S_0}{n+1}\frac{1-e^{\widehat{r}(n+1)}}{1-e^{\widehat{r}}} & \text{if } \widehat{r} \neq 0\\ S_0 & \text{if } \widehat{r} = 0 \end{cases}.$$

(2) Assume that $\widehat{r} \neq 0$. The value of the running sum implies that the call will be in the money because any path extending that initial path will end up with an arithmetic average of at least $X + a/(n+1) \geq X$. The expected terminal value is thus $a/(n+1)$ plus $E[\,\frac{1}{n+1}\sum_{i=k}^{n} S_i\,]$, which is equal to

$$\frac{S_k}{n+1}\,\frac{1-e^{\widehat{r}(n-k+1)}}{1-e^{\widehat{r}}}$$

by (1).

Exercise 11.7.6: We prove the claim for the $i < n/2$ case. Assume $S_0 = 1$ for convenience. The said difference at $u^i d^{n-i} = u^{2i-n}$ equals

$$A \equiv (u + u^2 + \cdots + u^i + u^{i-1} + \cdots + u^{2i-n}) - (d + d^2 + \cdots + d^{n-i} + d^{n-i-1} + \cdots + d^{n-2i}).$$

Similarly, the said difference at $u^{i-1}d^{n-i+1} = u^{2i-n-2}$ equals

$$B \equiv (u + u^2 + \cdots + u^{i-1} + u^{i-2} + \cdots + u^{2i-n-2}) - (d + d^2 + \cdots + d^{n-i+1} + d^{n-i} + \cdots + d^{n-2i+2}).$$

Now,

$$\begin{aligned} A-B &= u^i + u^{i-1} - u^{2i-n-1} - u^{2i-n-2} + d^{n-i+1} + d^{n-i} - d^{n-2i+1} - d^{n-2i} \\ &= u^i + u^{i-1} - u^{2i-n-1} - u^{2i-n-2} - (u^{2i-n} + u^{2i-n-1} - u^{i-n} - u^{i-n-1}) \\ &= u^i + u^{i-1} - u^{2i-n-1} - u^{2i-n-2} - u^{i-n}(u^i + u^{i-1} - 1 - u^{-1}) \\ &> u^i + u^{i-1} - u^{2i-n-1} - u^{2i-n-2} - (u^i + u^{i-1} - 1 - u^{-1}) > 0, \end{aligned}$$

as claimed.

Exercise 11.7.7: $y = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0) + [\, \frac{y_2 - y_0}{(x_2 - x_0)(x_2 - x_1)} - \frac{y_1 - y_0}{(x_1 - x_0)(x_2 - x_1)} \,](x - x_0)(x - x_1)$.

Exercise 11.7.8: "Bucketing" is no longer necessary because all the running sums at node $N(j, i)$ are integers lying between the integers $(j+1)\, A_{\min}(j, i)$ and $(j+1)\, A_{\max}(j, i)$, hence finite in number. The second error, from interpolation, disappears as a consequence because the running sum – now an integer – that we seek in the next node exists. Specifically, let $V(j, i, k)$ denote the option value at node $N(j, i)$ given that the sum up to that node is k. Use $S(j, i)$ to represent the integral asset price at node $N(j, i)$. By backward induction,

$$V(j, i, k) = [\, pV(j+1, i, k + S(j+1, i)) + (1-p)\, V(j+1, i+1, k + S(j+1, i+1)) \,]\, e^{-r\Delta t},$$

where $0 \le j < n$, $0 \le i \le j$, and $(j+1)\, A_{\min}(j, i) \le k \le (j+1)\, A_{\max}(j, i)$. See [251, 252].

Exercise 11.7.9: Run both algorithms and output their average. This average cannot deviate from the true value by more than the difference of their respective bounds.

Programming Assignment 11.7.11: We follow the terms in Subsection 11.7.1. Consider a lookback call on the minimum and let $S_{\min}$ denote the historical low as it stands now. (1) Figure 11.12 reveals that at node $N(j, i)$, the maximum minimum price between now and $N(j, i)$ is $S_{\max}(j, i) \equiv \min(S_0, S_0 u^{j-i} d^i)$. Similarly, the minimum minimum price is $S_{\min}(j, i) \equiv S_0 u^{-i}$, which is *independent* of j. The states at $N(j, i)$ obviously are $S_{\min}(j, i), uS_{\min}(j, i), u^2 S_{\min}(j, i), \ldots, S_{\max}(j, i)$. The number of states is either (a) $j - i$ when $j \ge 2i$ and $S_{\max}(j, i) = S_0 u^{j-i} d^i$ or (b) i when $j < 2i$ and $S_{\max}(j, i) = S_0$. It is then not hard to show that the total number of states over the whole tree is proportional to n^3, which is also the time bound. Use $N(j, i, k)$ to denote the state at node $N(j, i)$ when $u^k S_{\min}(j, i) = S_0 u^{k-i}$ is the minimum price between now and $N(j, i)$. Note that $k \le i$. Finally, let $C(j, i, k)$ denote the option value at state $N(j, i, k)$.

Backward induction starts with

$$C(n, i, k) = \max(S_0 u^{n-i} d^i - \min(u^k S_{\min}(n, i), S_{\min})),$$

at time n, where $0 \le i \le n$ and $0 \le k \le \ln_u \frac{S_{\max}(n,i)}{S_{\min}(n,i)}$. Inductively, each state at node $N(j, i)$ takes inputs from a state at the up node $N(j+1, i)$ and a state at the down node $N(j+1, i+1)$. Specifically, for $j = n-1, n-2, \ldots, 0$, the algorithm carries out

$$C(j, i, k) = \begin{cases} [\, pC(j+1, i, k) + (1-p)\, C(j+1, i+1, k+1) \,]/R, & \text{if } j > 2i \text{ (i)}, \\ [\, pC(j+1, i, k) + (1-p)\, C(j+1, i+1, k+1) \,]/R, & \text{if } j = 2i\,,\ k < i \text{ (ii)}, \\ [\, pC(j+1, i, k) + (1-p)\, C(j+1, i+1, k) \,]/R, & \text{if } j = 2i\,,\ k = i \text{ (iii)}, \\ [\, pC(j+1, i, k) + (1-p)\, C(j+1, i+1, k+1) \,]/R, & \text{if } j < 2i\,,\ k < j - i \text{ (iv)}, \\ [\, pC(j+1, i, k) + (1-p)\, C(j+1, i+1, k) \,]/R, & \text{if } j < 2i\,,\ k = j - i \text{ (v)}, \end{cases}$$

where $0 \le i \le j$ and $0 \le k \le \ln_u \frac{S_{\max}(j,i)}{S_{\min}(j,i)}$. The observations to follow were utilized in deriving the preceding formula. First, the minimum price in the up-node state always equals $N(j, i, k)$'s minimum price. In case (i), $N(j, i, k)$'s stock price $S_0 u^{j-i} d^i = S_0 u^{j-2i}$ exceeds S_0; thus the minimum price in the down-node state equals $N(j, i, k)$'s minimum price. In both (ii) and (iii), $N(j, i, k)$'s stock price equals S_0. In case (ii), $N(j, i, k)$'s minimum price is less than S_0; thus the minimum price in the down-node state equals $N(j, i, k)$'s minimum price. In case (iii), $N(j, i, k)$'s minimum price equals S_0; thus the minimum price in the down-node state equals $S_0 d$. In both (iv) and (v), $N(j, i, k)$'s stock price $S_0 u^{j-i} d^i = S_0 u^{j-2i}$ is less than S_0. In case (iv), $N(j, i, k)$'s minimum price is less than $S_0 u^{j-2i}$; thus the minimum price in the down-node state equals $N(j, i, k)$'s minimum price. In case (v), $N(j, i, k)$'s minimum price equals $S_0 u^{j-2i}$; thus the minimum price in the down-node state equals $S_0 u^{j-2i-1}$. The returned value is $C(0, 0, 0)$. If the option is American, simply take the greater of $S_0 u^{j-i} d^i - \min(u^k S_{\min}(j, i), S_{\min})$ and the $C(j, i, k)$ above for the final $C(j, i, k)$.

(2) When the option is newly issued, the number of states at each node can be drastically reduced to one. The basic idea is to calculate at each node a newly issued lookback option when the current stock

price is one. See [470, pp. 475–477] (the algorithm in Fig. 29.5 uses a similar idea). Unfortunately, this idea may not work for lookback options with a price history.

CHAPTER 12

Exercise 12.2.1: $4{,}288{,}200 + 400{,}000 \times r$ in U.S. dollars, where r is the spot exchange rate in \$/DEM 3 months from now.

Exercise 12.2.2: (1) Germany has lower interest rates than the U.S.

Exercise 12.2.3: Apply the interest rate parity [514, pp. 328–329].

Exercise 12.2.4: (1) A newly written forward contract is equivalent to a portfolio of one long European call and one short European put on the same underlying asset with a common expiration date equal to the delivery date. This is true because $S_T - X = \max(S_T - X, 0) - \max(X - S_T, 0)$. Hence, if X equals the forward price, the payoff matches exactly that of a forward contract. (2) Because a new forward contract has zero value, the strike price must be such that the put premium is equal to the call premium. Alternatively, we may write put–call parity (8.1) as $C = P + \text{PV}(F - X)$ with the help of Eq. (12.3). The conclusion is immediate as $F = X$. See [159]. (3) With (2), put–call parity (8.1) implies that

$$0 = C - P - S + Fe^{-r\tau} = -S + Fe^{-r\tau},$$

which is exactly Lemma 12.2.1. See [236, pp. 59–61] and [346, p. 244].

Exercise 12.2.5: Consider the case $f > (F - X)\, e^{-r\tau}$. We can create an arbitrage opportunity by buying one forward contract with delivery price F and shorting one forward contract with delivery price X, both maturing τ from now. This generates an initial cash inflow of f because the first contract has zero value by the definition of F. The cash flow at maturity is

$$(S_T - F) + (X - S_T) = -(F - X).$$

Hence a cash flow with a PV of $f - (F - X)\, e^{-r\tau} > 0$ has been ensured. To prove the identity under the remaining case $f < (F - X)\, e^{-r\tau}$, just reverse the above transactions.

Exercise 12.3.1: Repeat the argument leading to Eq. (12.8) but with lending when the cash flow is positive and borrowing when the cash flow is negative. The result is

$$(F_1 - F_0)\, R^{n-1} + (F_2 - F_1)\, R^{n-2} + \cdots + (F_n - F_{n-1}).$$

Clearly the result depends on how $F_n - F_0$ is distributed over the n-day period.

Exercise 12.3.3: Higher rates beget higher futures prices, generating positive cash flows that can be reinvested at higher rates. Similarly, lower rates beget lower futures prices, generating negative cash flows that can be financed at lower rates. Because both are advantageous to holders of futures contracts, their prices must be higher. See [402] or [514, p. 58] for more rigorous arguments.

Exercise 12.3.4: It assumes that the stock index is *not* adjusted for dividend payouts. First, note a crucial element in the proofs of Eqs. (12.4) and (12.6), that is, the dividends of the stock index – in fact, any underlying asset – are predictable. Therefore dividends must have predictable value. Now we come to the problem of adjustments for dividends. If the index were adjusted for dividends, the adjustment would have the effect of making the initial position in the underlying asset, a portfolio of stocks in the case of stock indices, not deliverable as is. This breaks the proofs, in which the only transactions that take place are related to the loan used to take a long position in the underlying asset or the cash outflow incurred when a short position is taken in the underlying asset.

Exercise 12.3.5: (1) The cost of carry is $Se^{r\tau}$ because it costs that much to carry the cash instrument. (2) $F = Se^{r\tau}$ by the condition of full carry. See [88, p. 170].

Exercise 12.3.6: If T-bills are yielding 6% per year and one owns gold outright for 1 year, then the "cost" of ownership is 6% of the cost, representing the interest one would have earned if one had bought a T-bill instead of the gold. At \$350 an ounce, the opportunity cost of owning 100 ounces for 1 year is $100 \times \$350 \times 0.06 = \$2{,}100$. Thus \$2,100 is the cost of carry. See [698, p. 192].

Exercise 12.3.7: Consider a portfolio consisting of borrowing S to buy a unit of the underlying commodity and a short position in a forward contract with the delivery price F (hence zero value). The initial net cash flow is zero. On the delivery date, the cash flow is $F - S - C$. Hence $F - S - C \leq 0$ must hold to prevent arbitrage. See [746, p. 40].

Exercise 12.3.8: Consider the strategy of shorting the commodity, investing the proceeds at the riskless rate, and buying the forward contract with the delivery price F. The initial net cash flow is zero. On the delivery date, the cash flow is $S + I - F - D$. This is because the investor has to pay F to get the commodity and close out the short position. The investor also has to pay any cash flow due the commodity holder. On the other hand, the investor receives $S + I$ for the investment. By Eq. (12.12),

$$S + I - F - D = S - F + C - U.$$

Hence $S - F + C - U \leq 0$ must hold to prevent arbitrage. See [746, p. 41].

Exercise 12.4.1: We have to show that there is no net cash outflow at expiration. Suppose the futures price decreases at expiration. We then exercise the put for a short position in futures, offsetting the long position in futures, and abandon the call. Similar actions can be taken for two other cases. See [95, p. 188] and [328].

Exercise 12.4.2: See [514, p. 204].

Exercise 12.4.3: Consider a portfolio of one long call, one short put, one short futures contract, and a loan of $Fe^{-rt} - X$. The initial portfolio value is $C - P - Fe^{-rt} + X$. At time t, the portfolio value is

$$0 - (X - F_t) - (F_t - F) - (F - Xe^{rt}) = X(e^{rt} - 1) \geq 0 \quad \text{if } F_t \leq X,$$
$$(F_t - X) - 0 - (F_t - F) - (F - Xe^{rt}) = X(e^{rt} - 1) \geq 0 \quad \text{if } F_t > X.$$

Suppose the put is exercised at time $s < t$. The value then is

$$C - (X - F_s) - (F_s - F) - [\, Fe^{-r(t-s)} - Xe^{rs}\,] = C + F[\, 1 - e^{-r(t-s)}\,] + X(e^{rs} - 1) \geq 0.$$

We hence conclude that $C - P - Fe^{-rt} + X \geq 0$.

For the other bound, consider a portfolio of one long put, one futures contract, lending of $F - Xe^{-rt}$, and one short call. The initial portfolio value is $P + F - Xe^{-rt} - C$. At time t, the portfolio value is

$$(X - F_t) + (F_t - F) + (Fe^{rt} - X) - 0 = F(e^{rt} - 1) \geq 0 \quad \text{if } F_t \leq X,$$
$$0 + (F_t - F) + (Fe^{rt} - X) - (F_t - X) = F(e^{rt} - 1) \geq 0 \quad \text{if } F_t > X.$$

Suppose the call option is exercised at time $s < t$. The value then is

$$P + (F_s - F) + [\, Fe^{rs} - Xe^{-r(t-s)}\,] - (F_s - X) = P + F(e^{rs} - 1) + X[\, 1 - e^{-r(t-s)}\,] \geq 0.$$

We hence conclude that $P + F - Xe^{-rt} - C \geq 0$. See [746, pp. 281–282].

Exercise 12.4.4: (1) We shall do it for calls only. Substitute $Se^{(r-q)t}$ for F into Eqs. (12.16) of Black's model to obtain Eq. (9.20). If the underlying asset does not pay dividends, then substitute Se^{rt} for F to obtain the original Black–Scholes formula for European calls. See [346, p. 201] and [470, p. 295]. (2) This holds because the futures price equals the cash asset's price at maturity.

Exercise 12.4.6: (1) From the binomial tree for the underlying stock, we replace the stock price S at each node of the tree with $Se^{r\tau}$, where τ is the time to the futures contract's maturity at that node. The binomial model for the futures price is then $Se^{r\tau} \to Sue^{r(\tau-\Delta t)}$ or $Sde^{r(\tau-\Delta t)}$; in other words, $F \to Fue^{-r\Delta t}$ or $Fde^{-r\Delta t}$. (2) If the underlying stock pays a continuous dividend yield of q, the binomial model for the futures price under the *same* risk-neutral probability is $Se^{(r-q)\tau} \to Sue^{-q\Delta t} e^{(r-q)(\tau-\Delta t)}$ or $Sde^{-q\Delta t} e^{(r-q)(\tau-\Delta t)}$; in other words, $F \to Fue^{-r\Delta t}$ or $Fde^{-r\Delta t}$, identical to (1).

Exercise 12.5.3: The forward dollar/yen exchange rate applicable to time $[\, 0, i\,]$ is $F_i \equiv Se^{(r-q)i}$ by Eq. (12.1). Hence the PV of the forward exchange of cash flow i years from now is $(F_i Y_i - D_i)\, e^{-ri} = SY_i e^{-qi} - D_i e^{-ri}$, in total agreement with Eq. (12.17).

CHAPTER 13

Exercise 13.1.1:

$$\begin{aligned} E[\, X(t+s) - X(0)\,] &= E[\, X(t+s) - X(s)\,] + E[\, X(s) - X(0)\,] \\ &= E[\, X(t) - X(0)\,] + E[\, X(s) - X(0)\,]. \end{aligned}$$

Define $f(t) \equiv E[\, X(t) - X(0)\,]$. The above equality says that $f(t+s) = f(t) + f(s)$. Differentiate it to get $f'(t+s) = f'(s)$. In particular, $f'(t) = f'(t-1+1) = f'(1)$. Hence $f(t) = tf'(1) + a$. However, $a = 0$ because $f(0) = f(0+0) = 2f(0)$; hence $f(t) = tf'(1)$. This implies that $f(1) = f'(1)$ and $f(t) = tf(1)$.

We now proceed to the proof for the variance. Note that

$$\begin{aligned}\mathrm{Var}[\,X(t+s)-X(0)\,] &= \mathrm{Var}[\,X(t+s)-X(s)\,]+\mathrm{Var}[\,X(s)-X(0)\,]\\ &= \mathrm{Var}[\,X(t)-X(0)\,]+\mathrm{Var}[\,X(s)-X(0)\,],\end{aligned}$$

where the first equality is due to independent increments and the second to stationarity. Define $f(t) \equiv \mathrm{Var}[\,X(t)-X(0)\,]$. The above equality says that $f(t+s) = f(t) + f(s)$. Differentiate it to get

$$f'(t+s) = f'(t), \qquad f'(t+s) = f'(s);$$

in particular, $f'(t) = f'(1)$. Hence $f(t) = tf'(1)+a$. However, $a = 0$ because $f(0) = 2f(0)$; hence $f(t) = tf'(1)$. This implies that $f(1) = f'(1)$ and $f(t) = tf(1)$. This proves the claim because

$$\mathrm{Var}[\,X(t)-X(0)\,] = \mathrm{Var}[\,X(t)\,]-\mathrm{Var}[\,X(0)\,].$$

See [566, p. 61].

Exercise 13.1.2: See [280, p. 96].

Exercise 13.1.3: First, $m_X(n) = E[\,X_n\,] = 0$. Now, for $a \neq b$,

$$E[\,X_a X_b\,] = \mathrm{Cov}[\,X_a,\,X_b\,] = 0$$

because X_a and X_b are uncorrelated. Finally,

$$K_X(m+n,m) = E[\,X_{m+n}X_m\,] = \begin{cases}1 & \text{if } n=0\\ 0 & \text{otherwise}\end{cases}.$$

Exercise 13.1.4: (1) In Eq. (13.1), use $\mu = 0$ and

$$\xi_n = \begin{cases}1 & \text{with probability } p\\ -1 & \text{with probability } q \equiv 1-p\end{cases}.$$

(2) Because the mean is zero, the variance equals $n[\,(1/2)\times 1+(1/2)\times(-1)^2\,] = n$.

Exercise 13.1.5: Consider two symmetric random walks whose joint displacement follows this distribution:

$$(X_1,\,X_2) = \begin{cases}(+1,+1) & \text{with probability } (1+\rho)/4\\ (+1,-1) & \text{with probability } (1-\rho)/4\\ (-1,+1) & \text{with probability } (1-\rho)/4\\ (-1,-1) & \text{with probability } (1+\rho)/4\end{cases}.$$

It is straightforward to verify that $\mathrm{Var}[\,X_1\,] = \mathrm{Var}[\,X_2\,] = 1$ and $E[\,X_1X_2\,] = \rho$. See [254].

Exercise 13.2.1: In fact,

$$\begin{aligned}&E[\,X(t_n)\mid X(t_{n-1}),\,X(t_{n-2}),\,\dots,\,X(t_1)\,]\\ &\quad= E[\,X(t_n)-X(t_{n-1})\mid X(t_{n-1}),\,X(t_{n-2}),\,\dots,\,X(t_1)\,]+X(t_{n-1})\\ &\quad= X(t_{n-1}),\end{aligned}$$

where the last equality is true because

$$\begin{aligned}&E[\,X(t_n)-X(t_{n-1})\mid X(t_{n-1}),\,X(t_{n-2}),\,\dots,\,X(t_1)\,]\\ &\quad= E[\,X(t_n)-X(t_{n-1})\mid X(t_{n-1})-X(t_{n-2}),\,\dots,\,X(t_2)-X(t_1),\,X(t_1)-X(0)\,]\\ &\quad= E[\,X(t_n)-X(t_{n-1})\,] = 0.\end{aligned}$$

Exercise 13.2.2: Apply the result of Exercise 6.4.3(2), and the definition of a martingale, Eq. (13.3).

Exercise 13.2.3: See [763, p. 229].

Exercise 13.2.4: By the definition and Eq. (6.5),

$$\mathrm{Var}[\,Z_n\,] = \mathrm{Var}[\,Z_{n-1}\,]+\mathrm{Var}[\,X_n\,]+\mathrm{Cov}[\,Z_{n-1},\,X_n\,].$$

We are done if we can prove that $\mathrm{Cov}[\,Z_{n-1},\,X_n\,] = 0$. Now,

$$\begin{aligned}\mathrm{Cov}[\,Z_{n-1},\,X_n\,] &= E[\,Z_{n-1}X_n\,]-E[\,Z_{n-1}\,]\,E[\,X_n\,]\\ &= E[\,Z_{n-1}X_n\,]\\ &= E[\,Z_{n-1}(Z_n-Z_{n-1})\,]\\ &= E[\,Z_{n-1}Z_n\,]-E\left[\,Z_{n-1}^2\,\right]\\ &= E[\,E[\,Z_{n-1}Z_n\mid Z_{n-1}\,]\,]-E\left[\,Z_{n-1}^2\,\right]\\ &= E[\,Z_{n-1}\,E[\,Z_n\mid Z_{n-1}\,]\,]-E\left[\,Z_{n-1}^2\,\right]\\ &= E[\,Z_{n-1}Z_{n-1}\,]-E\left[\,Z_{n-1}^2\,\right] = 0,\end{aligned}$$

where the first equality is due to Eq. (6.3), the second equality is due to Eq. (13.4) and $E[\, Z_0\,] = 0$, the fifth equality is due to the law of iterated conditional expectations, and the seventh equality is due to the definition of a martingale. See [763, p. 247].

Exercise 13.2.5: Observe that $\{S_n, n \geq 1\}$ is a martingale by Example 13.2.1. Let $Z_n \equiv S_n^2 - n\sigma^2$. Now,

$$\begin{aligned}
&E[\, Z_n \mid Z_1, \ldots, Z_{n-1}\,] \\
&\quad = E\left[\, S_n^2 \,\middle|\, Z_1, \ldots, Z_{n-1}\,\right] - n\sigma^2 \\
&\quad = E\left[\, (S_{n-1} + X_n)^2 \,\middle|\, S_1, \ldots, S_{n-1}\,\right] - n\sigma^2 \\
&\quad = E\left[\, S_{n-1}^2 \,\middle|\, S_1, \ldots, S_{n-1}\,\right] + 2\, E[\, S_{n-1} X_n \mid S_1, \ldots, S_{n-1}\,] \\
&\qquad + E\left[\, X_n^2 \,\middle|\, S_1, \ldots, S_{n-1}\,\right] - n\sigma^2 \\
&\quad = S_{n-1}^2 + 2 S_{n-1}\, E[\, X_n \mid X_1, \ldots, X_{n-1}\,] + E\left[\, X_n^2 \,\middle|\, X_1, \ldots, X_{n-1}\,\right] - n\sigma^2 \\
&\quad = S_{n-1}^2 + 2 S_{n-1} E[\, X_n\,] + E\left[\, X_n^2\,\right] - n\sigma^2 \\
&\quad = S_{n-1}^2 + E\left[\, X_n^2\,\right] - n\sigma^2 \\
&\quad = S_{n-1}^2 - (n-1)\,\sigma^2 \\
&\quad = Z_{n-1}.
\end{aligned}$$

See [763, p. 247].

Exercise 13.2.6: $E[\, Y_n - Y_{n-1} \mid X_1, X_2, \ldots, X_{n-1}\,]$ equals

$$E[\, C_n (X_n - X_{n-1}) \mid X_1, X_2, \ldots, X_{n-1}\,] = C_n E[\, X_n - X_{n-1} \mid X_1, X_2, \ldots, X_{n-1}\,] = 0.$$

See [725, p. 94] or [877, p. 97].

Exercise 13.2.7: By induction. See also [419, p. 227].

Exercise 13.2.8: Let p denote the unknown risk-neutral probability. One unit of foreign currency will be R_{f} units in a period. Translated into domestic currency, the expected value is

$$p R_{\mathrm{f}} Su + (1-p)\, R_{\mathrm{f}} Sd = R_{\mathrm{f}} S[\, pu + (1-p)\, d\,],$$

where S is the current exchange rate. By Eq. (13.7), we must have $S = R_{\mathrm{f}} S[\, pu + (1-p)\, d\,]/R$; so

$$\frac{R}{R_{\mathrm{f}}} = pu + (1-p)\, d = p(u-d) + d.$$

See [514, p. 320].

Exercise 13.2.9: Because $S(i+1) = S(i)\, up + S(i)\, d(1-p)$, where $p = (R-d)/(u-d)$, we have $S(i+1) = S(i)\, R$. The claim is then proved by induction.

Exercise 13.2.10: From relation (13.8), $F_i = E_i^{\pi}[\, F_n\,]$. However, $F_n = S_n$ because the futures price equals the spot price at maturity. See [514, p. 149].

Exercise 13.2.11: It suffices to show that $F_i = E_i^{\pi}[\, F_{i+1}\,]$ as the general case can be derived by recursion. Because futures contracts are marked to market daily, their value (not price) is zero. Hence $0 = E_i^{\pi}[\, (F_{i+1} - F_i)/M(i+1)\,]$ from Eq. (13.7). Because $M(i+1)$ is known at time i, this equality becomes $0 = \frac{E_i^{\pi}[\, F_{i+1}\,] - F_i}{M(i+1)}$, from which the claim easily follows. See [514, p. 171].

Exercise 13.2.12: (1) Assume that the bond's current price is $1/R$ and define $p \equiv (\frac{1}{d} - \frac{1}{R})\, \frac{ud}{u-d}$. The bond's prices relative to the stock price one period from now equal $1/(Su)$ and $1/(Sd)$. Now

$$p\, \frac{1}{Su} + (1-p)\, \frac{1}{Sd} = \left(\frac{1}{d} - \frac{1}{R}\right) \frac{ud}{u-d}\, \frac{1}{Su} + \left(\frac{1}{R} - \frac{1}{u}\right) \frac{ud}{u-d}\, \frac{1}{Sd} = \frac{1}{SR},$$

which is the bond price relative to the stock price today. An alternative is to evaluate Eq. (13.10) by using $S = 1$, $S_1 = S_2 = R$, $P = S$, $P_1 = Su$, and $P_2 = Sd$. See [681, p. 46]. (2) Prob_1 and Prob_2 are the risk-neutral probabilities that use the stock price and the bond price as numeraire, respectively. See [519] or [681, p. 48].

Exercise 13.2.13: Use the zero-coupon bond maturing at time k as numeraire in Eq. (13.9) to obtain

$$\frac{C(i)}{P(i,k)} = E_i^{\pi}\left[\, \frac{C(k)}{P(k,k)}\,\right],$$

where $P(j,k)$ denotes the zero-coupon bond's price at time $j \le k$. We now prove the claim by observing that $P(i,k) = d(k-i)$ and $P(k,k) = 1$. See [731, p. 171].

Exercise 13.3.1: The first two conditions for Brownian motion remain satisfied. Now,

$$E\left[\frac{x(t)-\mu t}{\sigma} - \frac{x(s)-\mu s}{\sigma}\right] = \frac{E[\,x(t)-x(s)\,]-\mu(t-s)}{\sigma} = 0.$$

Furthermore, the variance of $Z \equiv \{[\,x(t)-\mu t\,]-[\,x(s)-\mu s\,]\}/\sigma$ equals

$$\begin{aligned}
E[\,Z^2\,] &= \sigma^{-2}\big(E[\,\{x(t)-x(s)\}^2\,]-2E[\,\{x(t)-x(s)\}(\mu t-\mu s)\,]+\mu^2(t-s)^2\big)\\
&= \sigma^{-2}\big(E[\,\{x(t)-x(s)\}^2\,]-E[\,x(t)-x(s)\,]^2\big)\\
&= \sigma^{-2}\,\mathrm{Var}[\,x(t)-x(s)\,]\\
&= t-s.
\end{aligned}$$

(Note that Z's expected value is zero.)

Exercise 13.3.2: Without loss of generality, assume that $s < t$. Now,

$$\begin{aligned}
K_X(t,s) &\equiv E[\,\{X(t)-\mu t\}\{X(s)-\mu s\}\,]\\
&= E[\,\{X(t)-X(s)-\mu(t-s)+X(s)-\mu s\}\{X(s)-\mu s\}\,]\\
&= E[\,\{X(t)-X(s)-\mu(t-s)\}\{X(s)-\mu s\}\,]+E[\,\{X(s)-\mu s\}^2\,]\\
&= E[\,\{X(t)-X(s)-\mu(t-s)\}X(s)\,]+s\sigma^2\\
&= \{\,E[\,X(t)-X(s)\,]-\mu(t-s)\}\,E[\,X(s)\,]+s\sigma^2 \quad \text{from independence}\\
&= s\sigma^2.
\end{aligned}$$

See [147, p. 345] or [364, p. 36]. Another method is to observe that [763, p. 187]

$$\begin{aligned}
\mathrm{Cov}[\,X(s),X(t)\,] &= \mathrm{Cov}[\,X(s),X(s)+X(t)-X(s)\,]\\
&= \mathrm{Cov}[\,X(s),X(s)\,]+\mathrm{Cov}[\,X(s),X(t)-X(s)\,]\\
&= \mathrm{Var}[\,X(s)\,]\\
&= s\sigma^2.
\end{aligned}$$

Exercise 13.3.3: Let $0 < t_1 < \cdots < t_n$. Then,

$$\begin{aligned}
&E[\,X(t_n)-X(0) \mid X(t_{n-1})-X(0),\ldots,X(t_1)-X(0)\,]\\
&= E[\,X(t_n)-X(t_{n-1})+X(t_{n-1})-X(0) \mid X(t_{n-1})-X(0),\ldots,X(t_1)-X(0)\,]\\
&= E[\,X(t_n)-X(t_{n-1}) \mid X(t_{n-1})-X(0),\ldots,X(t_1)-X(0)\,]+X(t_{n-1})-X(0)\\
&= X(t_{n-1})-X(0).
\end{aligned}$$

The last equality holds because $E[\,X(t_n)-X(t_{n-1})\,]=0$ by the definition of the Wiener process and independent increments. This problem is in fact just a specialization of Exercise 13.2.1. See [280, p. 97] or [541, p. 233].

Exercise 13.3.5: For (2),

$$\begin{aligned}
&E[\,X(t)^2-\sigma^2 t \mid X(u), 0 \le u \le s\,]\\
&= E[\,X(t)^2 \mid X(s)\,]-\sigma^2 t\\
&= E[\,\{X(t)-X(s)\}^2 \mid X(s)\,]+E[\,2X(s)\{X(t)-X(s)\}+X(s)^2 \mid X(s)\,]-\sigma^2 t\\
&= \sigma^2(t-s)+X(s)^2-\sigma^2 t\\
&= X(s)^2-\sigma^2 s.
\end{aligned}$$

For (3),

$$\begin{aligned}
E[\,e^{\alpha X(t)-\alpha^2\sigma^2 t/2} \mid X(u), 0 \le u \le s\,] &= e^{\alpha X(s)-\alpha^2\sigma^2 s/2}\,E[\,e^{\alpha\{X(t)-X(s)\}-\alpha^2(t-s)/2}\,]\\
&= e^{\alpha X(s)-\alpha^2\sigma^2 s/2},
\end{aligned}$$

where the last equality is due to Eq. (6.8). See [419, p. 7] and [543, p. 358].

Exercise 13.3.6: First, $dP = 2^{-n}$ because the random walk is symmetric. On the other hand, $dQ = \prod_{i=1}^{n} [\, 2^{-1} + (p - 2^{-1})\, X_i \,]$. Recall that $p \equiv (1 + \mu\sqrt{\Delta t}\,)/2$. Hence $dQ/dP = \prod_{i=1}^{n} [1 + (2p-1)\, X_i] = \prod_{i=1}^{n} (1 + \mu\sqrt{\Delta t}\, X_i)$.

Exercise 13.3.7: Because $X(t) - X(s) \sim N(\mu(t-s), \sigma^2(t-s))$,

$$E[\, Y(t) \mid Y(s)\,] = E[\, e^{X(t)} \mid e^{X(s)}\,] = e^{X(s)}\, E[\, e^{X(t)-X(s)}\,] = Y(s)\, e^{(t-s)(\mu+\sigma^2/2)},$$

where the last equality is due to Eq. (13.12). See [543, p. 364] or [692, p. 349].

Exercise 13.3.8: (1) From Exercise 13.3.7, the rate of return is $\mu + \sigma^2/2$. Refer to Comment 14.4.1 for a discussion of why it is not μ. (2) It follows easily from the definition of $S(t)$. To establish (3), simply observe that $\ln(\,\cdot\,)$ is a concave function.

Exercise 13.3.9: Recall that $\mathrm{Var}[\, X\,] = E[\, X^2\,] - E[\, X\,]^2$. Hence, for $X \sim N(\mu, \sigma^2)$, we have $E[\, X^2\,] = \sigma^2 + \mu^2$. Returning to our exercise, simply observe that the expression within the brackets is normally distributed with mean $\mu t/2^n$ and variation $\sigma^2 t/2^n$.

Exercise 13.3.10: (1) Recall that $X((k+1)\, t/2^n) - X(kt/2^n) \sim N(0, 2^{-n})$. Note that if $X \sim N(0, \sigma^2)$, then $|X|$ has mean $\sigma\sqrt{2/\pi}$ and variance $\sigma^2(1 - 2/\pi)$. (2) Let $b \equiv 2^{n/2}\sqrt{2/\pi}$ and $c \equiv 1 - 2/\pi$. For any $\alpha > 0$, we can easily find an n_0 such that $\alpha < b - c^{1/2} 2^{n/2}$ for $n > n_0$. Chebyshev's inequality implies that

$$\begin{aligned}
\mathrm{Prob}[\, f_n(X) > \alpha\,] &\geq \mathrm{Prob}[\, f_n(X) \geq b - c^{1/2} 2^{n/2}\,] \\
&\geq \mathrm{Prob}[\, |\, f_n(X) - b\,| \leq c^{1/2} 2^{n/2}\,] \\
&\geq 1 - \frac{c}{(c^{1/2} 2^{n/2})^2} \\
&= 1 - 2^{-n} \to 1.
\end{aligned}$$

See [73, p. 64].

Exercise 13.4.1 For instance,

$$\begin{aligned}
E[\, B(t)^2\,] &= E\left[\, W(t)^2 - \frac{2t}{T}\, W(t)\, W(T) + \frac{t^2}{T^2}\, W(T)^2 \,\right] \\
&= t - \frac{2t}{T}\, E[\, W(t)\,\{\, W(T) - W(t)\,\}\,] - \frac{2t}{T}\, E[\, W(t)^2\,] + \frac{t^2}{T} \\
&= t - 0 - \frac{2t^2}{T} + \frac{t^2}{T} = t - \frac{t^2}{T}.
\end{aligned}$$

See [193, p. 193] or [557, p. 59].

Exercise 13.4.2 It is $x + W(t) - (t/T)[\, W(T) - y + x\,]$, $0 \leq t \leq T$ [557, p. 59].

Chapter 14

Exercise 14.1.1: Consider s and t such that $t_k \leq s < t \leq t_{k+1}$ first. From Eq. (14.1),

$$\begin{aligned}
&E[\, \mathrm{I}_t(X) \mid W(u), 0 \leq u \leq s\,] \\
&\quad = E[\, \mathrm{I}_s(X) + X(t_k)\{\, W(t) - W(s)\,\} \mid W(u), 0 \leq u \leq s\,] \\
&\quad = E[\, \mathrm{I}_s(X) \mid W(u), 0 \leq u \leq s\,] + X(t_k)\, E[\, W(t) - W(s) \mid W(u), 0 \leq u \leq s\,] \\
&\quad = \mathrm{I}_s(X). \qquad (33.8)
\end{aligned}$$

Simple induction can show that $E[\, \mathrm{I}_t(X) \mid W(u), 0 \leq u \leq t_i\,] = \mathrm{I}_{t_i}(X)$ for $t_i < t_k \leq t \leq t_{k+1}$. Hence, for $t_i \leq s < t_{i+1}$, $t_k \leq t < t_{k+1}$, and $i < k$,

$$\begin{aligned}
E[\, \mathrm{I}_t(X) \mid W(u), 0 \leq u \leq s\,] &= E[\, E[\, \mathrm{I}_t(X) \mid W(u), 0 \leq u \leq t_{i+1}\,] \mid W(u), 0 \leq u \leq s\,] \\
&= E[\, \mathrm{I}_{t_{i+1}}(X) \mid W(u), 0 \leq u \leq s\,] \\
&= \mathrm{I}_s(X)
\end{aligned}$$

by Eq. (33.8). See [566, p. 158] and [585, pp. 167–168].

Exercise 14.1.2: The approximating sum is now $\sum_{k=0}^{n-1} W(t_{k+1})[\, W(t_{k+1}) - W(t_k)\,]$ [30, p. 104].

Exercise 14.1.3:

$$\begin{aligned} &E\left[\frac{W(t)^2}{2} \,\middle|\, W(u), 0 \le u \le s \right] \\ &= E\left[\frac{W(s)^2}{2} \,\middle|\, W(u), 0 \le u \le s \right] + E\left[\frac{W(t)^2}{2} - \frac{W(s)^2}{2} \,\middle|\, W(u), 0 \le u \le s \right] \\ &= \frac{W(s)^2}{2} + \frac{t-s}{2}, \end{aligned}$$

which is not a martingale [566, p. 160].

Exercise 14.1.4: Use Eq. (14.3) and the definition $E[\, W(t)^2 \,] = \mathrm{Var}[\, W(t) \,] = t$.

Exercise 14.1.5: This is because $\int_{t_0}^{t} dW(s) = W(t) - W(t_0)$ [30, p. 59].

Exercise 14.2.1:

$$\phi_0 \boldsymbol{S}_0 + G_k = \phi_0 \boldsymbol{S}_0 + \sum_{i=0}^{k-1} \phi_i (\boldsymbol{S}_{i+1} - \boldsymbol{S}_i) = \phi_0 \boldsymbol{S}_0 + \sum_{i=0}^{k-1} (\phi_{i+1} \boldsymbol{S}_{i+1} - \phi_i \boldsymbol{S}_i) = \phi_k \boldsymbol{S}_k.$$

See [70] or [420, p. 226].

Exercise 14.2.2: Applying Ito's formula with $f(x) = x^2$, we have $W(t)^2 = \int_0^t ds + \int_0^t 2W(s)\, dW$. Hence $\int_0^t W(s)\, dW = [\, W(t)^2/2 \,] - (t/2)$. See [585, p. 172].

Exercise 14.3.1: (1) See [746, p. 175]. (2) From Example 14.3.3, we know that $X(t) = e^{Y(t)}$, where Y is a $(\mu - \sigma^2/2, \sigma)$ Brownian motion. Thus $\ln(X(t)/X(0)) \sim N((\mu - \sigma^2/2)\, t, \sigma^2 t)$.

Exercise 14.3.2: Note that $dR = d(\ln X) + (\sigma^2/2)\, dt$. From the solution to Exercise 14.3.1, we arrive at

$$dR = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma\, dW + \frac{\sigma^2}{2}\, dt = \mu\, dt + \sigma\, dW.$$

Exercise 14.3.3: (1) Apply Ito's formula (14.10) to the function $f(x) = x^n$ to obtain

$$dX^n = nX^{n-1}\, dX + \frac{1}{2}\, n(n-1)\, X^{n-2}\, (dX)^2.$$

Substitute W_t for X above to arrive at the stochastic differential equation,

$$dW_t^n = nW_t^{n-1}\, dW_t + \frac{n(n-1)}{2}\, W_t^{n-2}\, dt.$$

See [30, p. 94]. (2) Expand

$$dW(t)^n = [\, W(t) + dW(t) \,]^n - W(t)^n = nW(t)^{n-1}\, dW(t) + \frac{n(n-1)}{2}\, W(t)^{n-2}\, dt$$

with Eq. (13.15) and (13.16). See [373, p. 89].

Exercise 14.3.4: The multidimensional Ito's lemma (Theorem 14.2.2) can be used to show that $dU = (1/2)\, dY + (1/2)\, dZ$, which can be expanded into

$$\frac{dU}{U} = \left(\frac{Y}{Y+Z}\, a + \frac{Z}{Y+Z}\, f \right) dt + \left(\frac{Y}{Y+Z}\, b + \frac{Z}{Y+Z}\, g \right) dW.$$

Exercise 14.3.5: The Ito process $U = YZ$ is defined by Y and Z with differentials, respectively, $dY = a\, dt + b\, dW_{\mathrm{y}}$ and $dZ = f\, dt + g\, dW_{\mathrm{z}}$. Keep in mind that dW_{y} and dW_{z} have correlation ρ. The multidimensional Ito's lemma (Theorem 14.2.3) can be used to show that

$$dU = Z\, dY + Y\, dZ + (a\, dt + b\, dW_{\mathrm{y}})(f\, dt + g\, dW_{\mathrm{z}}) = (Za + Yf + bg\rho)\, dt + Zb\, dW_{\mathrm{y}} + Yg\, dW_{\mathrm{z}}.$$

Exercise 14.3.7: It is $(1/F)\, dF = (1/X)\, dX + (1/Y)\, dY + \sigma^2\, dt$ [746, p. 176].

Exercise 14.3.8: View Y as $Y(X, t) \equiv Xe^{kt}$ and apply Ito's lemma [373, p. 106].

Exercise 14.3.9 Let $f(t) \ge 0$ be any continuous strictly increasing function. Then

$$W(f(t + \Delta t)) - W(f(t)) \sim N(0,\, f(t + \Delta t) - f(t)),$$

which approaches $N(0,\, f'(t)\Delta t) = \sqrt{f'(t)\Delta t}\; N(0, 1)$. Now,

$$dY(t) = \frac{\partial e^{-t}}{\partial t}\, W(e^{2t})\, dt + e^{-t}\, dW(f(t)) = -Y(t)\, dt + e^{-t} \sqrt{2e^{2t}\, dt}\; \xi = -Y(t)\, dt + \sqrt{2}\, dW,$$

where $\xi \sim N(0, 1)$. Hence $dY = -Y\, dt + \sqrt{2}\, dW$. The general formula for $a(t)\, W(f(t))$ appears in [230, p. 229].

Exercise 14.3.10: Ito's lemma (Theorem 14.2.3) says that H satisfies

$$\begin{aligned} dH &= \frac{\partial H}{\partial S}\, dS + \frac{\partial H}{\partial \sigma}\, d\sigma - \frac{\partial H}{\partial \tau}\, d\tau + \frac{1}{2}\,\sigma^2 S^2\, \frac{\partial^2 H}{\partial S^2}\, dt + \sigma\gamma S\rho\, \frac{\partial^2 H}{\partial S \partial \tau}\, dt + \frac{1}{2}\,\gamma^2\, \frac{\partial^2 H}{\partial \tau^2}\, dt \\ &= \left[\mu S\, \frac{\partial H}{\partial S} + \beta(\overline{\sigma} - \sigma)\, \frac{\partial H}{\partial \sigma} - \frac{\partial H}{\partial \tau} + \frac{1}{2}\,\sigma^2 S^2\, \frac{\partial^2 H}{\partial S^2} + \sigma\gamma S\rho\, \frac{\partial^2 H}{\partial S\partial \tau} + \frac{1}{2}\,\gamma^2\, \frac{\partial^2 H}{\partial \tau^2} \right] dt \\ &\quad + \sigma S\, \frac{\partial H}{\partial S}\, dW_1 + \gamma\, \frac{\partial H}{\partial \sigma}\, dW_2. \end{aligned}$$

Note that $d\tau = -dt$. See [790].

Exercise 14.3.11: Without loss of generality, we will verify, instead, that

$$\frac{\partial p}{\partial \tau} = \frac{1}{2}\, \frac{\partial^2 p}{\partial x^2} - \frac{1}{2}\, x\, \frac{\partial p}{\partial x},$$

where $\tau \equiv t - s$. This is admissible because the process is homogeneous, which we can see by observing that both the drift and the diffusion of the Ito process are independent of time t [30, Remark 2.6.8]. From the hint, we know that

$$p(x, s; y, t) = \frac{1}{\sqrt{2\pi\,(1 - e^{-\tau})}}\, \exp\left[-\frac{\left(y - xe^{-\tau/2}\right)^2}{2\,(1 - e^{-\tau})} \right] = B^{-1/2} A,$$

where

$$A \equiv \exp\left[-\frac{(y - xe^{-\tau/2})^2}{2\,(1 - e^{-\tau})} \right], \qquad B \equiv 2\pi(1 - e^{-\tau}).$$

After we verify the following equations, we will be done:

$$\begin{aligned} \frac{\partial p}{\partial \tau} &= B^{-3/2} A\pi \left[-e^{-\tau} + (y - xe^{-\tau/2})^2 \frac{e^{-\tau}}{1 - e^{-\tau}} - (y - xe^{-\tau/2}) xe^{-\tau/2} \right], \\ \frac{1}{2}\, x\, \frac{\partial p}{\partial x} &= B^{-3/2} A\pi\, (y - xe^{-\tau/2}) e^{-\tau/2} x, \\ \frac{1}{2}\, \frac{\partial^2 p}{\partial x^2} &= B^{-3/2} A\pi \left[-e^{-\tau} + \frac{\left(y - xe^{-\tau/2}\right)^2}{1 - e^{-\tau}}\, e^{-\tau} \right]. \end{aligned}$$

Exercise 14.4.1: It follows from Eq. (13.17), the infinite total variation of Brownian motion (with probability one) [723].

Exercise 14.4.2: If rates are negative, then the price exceeds the total sum of future cash flows. To generate arbitrage profits, short the bond and reserve part of the proceeds to service future cash flow obligations.

Exercise 14.4.3: It must be that $g(T) = 0$; hence $\mu = \sigma = 0$. In particular, there are no random changes. See [207].

Exercise 14.4.4: Let the current time be zero and the portfolio contain $b_i > 0$ units of bond i for $i = 1, 2$. (1) Let $\theta(t) \equiv A(t)/L(t)$. To begin with,

$$A(t) = \sum_{i=1}^{2} b_i\, P(r(t), t, t_i), \qquad L(t) = P(r(t), t, s),$$

where $P(r(t), t, s)$ is from Eq. (14.16). Note that $r(t)$ is a random variable. Now,

$$\begin{aligned} \theta(t) &= \sum_{i=1}^{2} b_i\, \frac{P(\,r(t), t, t_i\,)}{P(\,r(t), t, s\,)} \\ &= \sum_{i=1}^{2} b_i \times \exp\left[-r(t)\,(t_i - s) - \mu\, \frac{(t_i - t)^2 - (s - t)^2}{2} + \sigma^2\, \frac{(t_i - t)^3 - (s - t)^3}{6} \right]. \end{aligned}$$

Because $\partial^2\theta(t)/\partial r^2$ equals

$$\sum_{i=1}^{2} b_i (t_i - s)^2 \times \exp\left[-r(t)\,(t_i - s) - \mu\, \frac{(t_i - t)^2 - (s - t)^2}{2} + \sigma^2\, \frac{(t_i - t)^3 - (s - t)^3}{6} \right] > 0,$$

$\theta(t)$ is indeed a convex function of $r(t)$.

(2) Immunization requires that $\partial\theta/\partial r = 0$ at time zero, that is,

$$\sum_{i=1}^{2} -(t_i - s)\, b_i \times \exp\left[-r(0)\,(t_i - s) - \mu\, \frac{t_i^2 - s^2}{2} + \sigma^2\, \frac{t_i^3 - s^3}{6} \right] = 0. \tag{33.9}$$

Incidentally, after some easy manipulations, Eq. (33.9) becomes the standard duration condition, $-[\,\partial A(t)/\partial r\,]/A(t) = s$. Another immunization condition is that the asset and the liability should match in value at time zero, or $\theta(0) = 1$. This leads to

$$\sum_{i=1}^{2} b_i \times \exp\left[-r(0)\,(t_i - s) - \mu\, \frac{t_i^2 - s^2}{2} + \sigma^2\, \frac{t_i^3 - s^3}{6} \right] = 1. \tag{33.10}$$

With the definitions

$$w_i \equiv b_i \times \exp\left[-r(0)\,(t_i - s) - \mu\, \frac{t_i^2 - s^2}{2} + \sigma^2\, \frac{t_i^3 - s^3}{6} \right],$$

we arrive at

$$\begin{aligned}\theta(t) &= \sum_{i=1}^{2} w_i \times \exp\left[-\{\, r(t) - r(0)\,\}\,(t_i - s) - \mu\, \frac{(t_i - t)^2 - (s - t)^2 - t_i^2 + s^2}{2} \right.\\ &\qquad \left. + \sigma^2\, \frac{(t_i - t)^3 - (s - t)^3 - t_i^3 + s^3}{6} \right]\\ &= \sum_{i=1}^{2} w_i \times \exp\left[-\{\, r(t) - r(0)\,\}\,(t_i - s) + \mu t (t_i - s) + \sigma^2\, \frac{t - t_i - s}{2}\, t(t_i - s) \right]\end{aligned} \tag{33.11}$$

The following lemma then implies that $\theta(t) < w_1 + w_2 = 1$, concluding the proof.

LEMMA

$$\exp\left[-\{\, r(t) - r(0)\,\}\,(t_i - s) + \mu t (t_i - s) + \sigma^2\, \frac{t - t_i - s}{2}\, t(t_i - s) \right] < 1$$

for $i = 1, 2$.

Proof of Lemma: Equations (33.9) and (33.10) can be used to obtain $w_1 = [\,(s - t_2)/(t_1 - t_2)\,]$ and $w_2 = [\,(t_1 - s)/(t_1 - t_2)\,]$ because $\sum_{i=1}^{2} -(t_i - s)\, w_i = 0$ and $\sum_{i=1}^{2} w_i = 1$. From Eq. (33.11), the $[\,\partial\theta(t)/\partial r(t)\,] = 0$ condition becomes

$$\sum_{i=1}^{2} (s - t_i)\, w_i \times \exp\left[-\{\, r(t) - r(0)\,\}\,(t_i - s) + \mu t (t_i - s) + \sigma^2\, \frac{t - t_i - s}{2}\, t(t_i - s) \right] = 0.$$

Substitute the solution for (w_1, w_2) into the preceding equation to get

$$\begin{aligned}0 &= (s - t_1)\, \frac{s - t_2}{t_1 - t_2} \exp\left[-\{\, r(t) - r(0)\,\}\,(t_1 - s) + \mu t (t_1 - s) + \sigma^2\, \frac{t - t_1 - s}{2}\, t(t_1 - s) \right]\\ &\quad + (s - t_2)\, \frac{t_1 - s}{t_1 - t_2} \exp\left[-\{\, r(t) - r(0)\,\}\,(t_2 - s) + \mu t (t_2 - s) + \sigma^2\, \frac{t - t_2 - s}{2}\, t(t_2 - s) \right].\end{aligned}$$

Therefore

$$\begin{aligned}&-[\, r(t) - r(0)\,]\,(t_1 - s) + \mu t (t_1 - s) + \sigma^2\, \frac{t - t_1 - s}{2}\, t(t_1 - s)\\ &\quad = -[\, r(t) - r(0)\,]\,(t_2 - s) + \mu t (t_2 - s) + \sigma^2\, \frac{t - t_2 - s}{2}\, t(t_2 - s).\end{aligned} \tag{33.12}$$

Identity (33.12) implies

$$\begin{aligned}r(t) &= \left[r(0)(t_2 - t_1) + \mu t (t_2 - t_1) + \sigma^2\, \frac{t - t_2 - s}{2}\, t(t_2 - s) - \sigma^2\, \frac{t - t_1 - s}{2}\, t(t_1 - s) \right] \Big/ (t_2 - t_1)\\ &= r(0) + \mu t + \frac{\sigma^2 t}{2}\, (t - t_2 - t_1).\end{aligned}$$

Finally, substitute the preceding equation into the $i = 1$ term in Eq. (33.11) to get

$$\begin{aligned}&\exp\left[-\left\{ \mu t + \frac{\sigma^2 t}{2}\, (t - t_2 - t_1) \right\} (t_1 - s) + \mu t (t_1 - s) + \sigma^2\, \frac{t - t_1 - s}{2}\, t(t_1 - s) \right]\\ &\quad = \exp\left[\frac{\sigma^2 t}{2}\, (t_1 - s)(t_2 - s) \right] < 1.\end{aligned}$$

The $i = 2$ case is redundant as its value is identical by identity (33.12).

https://doi.org/10.1017/CBO9780511546839.034 Published online by Cambridge University Press

Exercise 14.4.5: Direct from identity (13.12), it is $Se^{\mu T}$.

Exercise 14.4.6: Although $dS/S = \sigma\, dW$ looks "symmetric" around zero, it is not. We can see this most clearly by looking at its discrete version, which says that the percentage change has zero mean. However, a 2% decrease followed by a 2% increase results in $1.02 \times 0.98 = 0.9996 < 1$. In other words, a 2% decrease has to be followed by a more than 2% increase to be back to the original level.

Exercise 14.4.7: They follow from identity (13.12) and the fact that $\ln(S(t)/S(0))$ is a $(\mu - \sigma^2/2, \sigma)$ Brownian motion.

Exercise 14.4.8: This can be justified heuristically as follows. Partition $[\,0, T\,]$ into n equal periods. Let σ_{i-1} denote the (annualized) volatility during period i and S_i the stock price at time i. Consequently,

$$\ln \frac{S_i}{S_{i-1}} \sim N\left(\mu\,\Delta t - \frac{1}{2}\,\sigma_{i-1}^2\,\Delta t, \sigma_{i-1}^2\,\Delta t\right)$$

by relation (14.17). If S and σ are uncorrelated, preceding relation holds, given σ_{i-1}. The probability distribution of $\ln(S_n/S_0)$, given the path followed by σ (i.e., $\sigma_1, \sigma_2, \dots, \sigma_n$), is thus normal with mean

$$\sum_i \left(\mu\,\Delta t - \frac{1}{2}\,\sigma_{i-1}^2\,\Delta t\right) = \mu T - \frac{1}{2}\sum_i \sigma_{i-1}^2\,\Delta t \to \mu T - \frac{1}{2}\,\widehat{\sigma^2}\,T$$

and variance $\sum_i (\sigma_{i-1}^2\,\Delta t) \to \widehat{\sigma^2}\,T$.

Exercise 14.4.9: This is due to $(\Delta S)^2 \approx \sigma^2 S^2\,\Delta t$ [514, p. 221].

Exercise 14.4.11: Note that

$$\ln u = -\ln 0, \qquad u \approx 1 + \sigma\sqrt{\frac{\tau}{n}} + \frac{\sigma^2\tau}{2n}, \qquad d \approx 1 - \sigma\sqrt{\frac{\tau}{n}} + \frac{\sigma^2\tau}{2n}.$$

Hence, $u + d \approx 2 + \sigma^2(\tau/n)$ and $u - d \approx 2\sigma\sqrt{\tau/n}$. Finally,

$$p \approx \frac{[\,1 + r(\tau/n)\,] - [\,1 - \sigma\sqrt{\tau/n} + (1/2)\,\sigma^2(\tau/n)\,]}{2\sigma\sqrt{\tau/n}} \approx \frac{1}{2} + \frac{r(\tau/n) - (1/2)\,\sigma^2(\tau/n)}{2\sigma\sqrt{\tau/n}}.$$

Now,

$$E[\,X_i\,] = p\ln u + (1-p)\ln d = (2p-1)\ln u \approx r\,\frac{\tau}{n} - \frac{1}{2}\,\sigma^2\frac{\tau}{n}.$$

Hence $E[\sum_{i=1}^n X_i\,] = n\,E[\,X_i\,] \approx r\tau - \sigma^2\tau/2$, as desired.

We proceed to calculate the variance:

$$\begin{aligned}
\mathrm{Var}[\,X_i\,] &= p(\ln u - E[\,X_i\,])^2 + (1-p)(\ln d - E[\,X_i\,])^2\\
&= p[\,\ln u - (2p-1)\ln u\,]^2 + (1-p)[\,\ln d - (2p-1)\ln u\,]^2\\
&= p[\,\ln u - (2p-1)\ln u\,]^2 + (1-p)[\,-\ln u - (2p-1)\ln u\,]^2\\
&= 4p(1-p)\ln^2 u\\
&\approx 4\left[\frac{1}{2} + \frac{r(\tau/n) - (1/2)\,\sigma^2(\tau/n)}{2\sigma\sqrt{\tau/n}}\right]\left[\frac{1}{2} - \frac{r(\tau/n) - (1/2)\,\sigma^2(\tau/n)}{2\sigma\sqrt{\tau/n}}\right]\sigma^2\frac{\tau}{n}\\
&\approx \sigma^2\frac{\tau}{n}.
\end{aligned}$$

Hence $\mathrm{Var}[\sum_{i=1}^n X_i\,] = n\,\mathrm{Var}[\,X_i\,] \approx \sigma^2\tau$, as desired.

Exercise 14.4.12: The troubling step is

$$X_{i+1} = \ln\left(1 + \frac{S_{i+1} - S_i}{S_i}\right) \approx \frac{S_{i+1} - S_i}{S_i} \equiv \frac{\Delta S_i}{S_i}.$$

It should have been

$$X_{i+1} = \ln\left(1 + \frac{S_{i+1} - S_i}{S_i}\right) \approx \frac{\Delta S_i}{S_i} - \frac{1}{2}\left(\frac{\Delta S_i}{S_i}\right)^2.$$

Now, because $(dS)^2 = \sigma^2 S^2\,dt$, the preceding approximation becomes $X_{i+1} \approx (\Delta S_i/S_i) - (1/2)\,\sigma^2\,\Delta t$. This corrects the problem.

Exercise 14.4.13: Apply Eq. (14.7) to $dX = (r - \sigma^2/2)\,dt + \sigma\,dW$ [229].

https://doi.org/10.1017/CBO9780511546839.034 Published online by Cambridge University Press

CHAPTER 15

Exercise 15.2.2: Different specifications of the underlying process imply different ways of estimating the diffusion coefficient from the data, and they may not give the same value. Hence the σ that gets plugged into the Black–Scholes formula could be different for different models. Misspecification of the drift therefore may lead to misestimation of the diffusion of the Ito process [819]. See [613] for the complete argument.

Exercise 15.2.3: Write valuation formula (9.4) as

$$C(S,t) = \frac{pC(S^+, t+\Delta t) + (1-p)\, C(S^-, t+\Delta t)}{e^{r\Delta t}},$$

where $p \equiv (Se^{r\Delta t} - S^-)/(S^+ - S^-)$. Express $C(S^+, t+\Delta t)$ and $C(S^-, t+\Delta t)$ in Taylor expansion around (S,t) and substitute them into the preceding formula. See [696, p. 416].

Exercise 15.2.4: This is the risk-neutral valuation with the density function of the lognormal distribution [492, p. 312].

Exercise 15.2.5: Follow the same steps with slightly different boundary conditions: $B(0,\tau) = 1$ for $\tau > 0$ and $B(S,0) = \max(1 - S/X, 0)$. In the end, we get $\Theta(z,0) = \max(e^{-z} - 1, 0)$ instead. Now,

$$\Theta(z,u) = \frac{1}{\sqrt{2\pi u}} \int_{-\infty}^{0} (e^{-y} - 1)\, e^{-(z-y)^2/(2u)}\, dy = \cdots = -N\left(-\frac{z}{\sqrt{u}}\right) + \frac{1}{x}\, N\left(-\frac{z-u}{\sqrt{u}}\right).$$

The rest is straightforward.

Exercise 15.2.6:

$$\frac{\partial P}{\partial t} + rS\,\frac{\partial P}{\partial S} + \frac{1}{2}\sigma^2 S^2\,\frac{\partial^2 P}{\partial S^2} - rP = rX < 0$$

when $X - S$ is substituted into the equation [879, p. 112].

Exercise 15.3.1: It is

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 F^2\,\frac{\partial^2 C}{\partial F^2} = rC.$$

This can be derived by the hedging argument with the observation that it costs nothing to enter into a futures contract. An alternative replaces $\partial C/\partial t$ with $\partial C/\partial t - rF(\partial C/\partial F)$, $\partial C/\partial S$ with $e^{r(T-t)}(\partial C/\partial F)$, and $\partial^2 C/\partial S^2$ with $e^{2r(T-t)}(\partial^2 C/\partial F^2)$ in the original Black–Scholes differential equation. See [125], [575, Subsection 2.3.3], and [879, p. 100].

Exercise 15.3.3: (1) It is straightforward [492, p. 380]. (2)

$$\frac{\partial C}{\partial t} + \sum_{i=1}^{n} rS_i\,\frac{\partial C}{\partial S_i} + \sum_{i=1}^{n} \frac{\sigma_i^2 S_i^2}{2}\,\frac{\partial^2 C}{\partial S_i^2} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij}\sigma_i\sigma_j S_i S_j\,\frac{\partial^2 C}{\partial S_i \partial S_j} = rC.$$

Exercise 15.3.4: $\max(S_1, S_2) = S_1 + \max(S_2 - S_1, 0)$.

Exercise 15.3.5: Its terminal value can be written as

$$\max(S_1(\tau) - X, 0) + \max(S_2(\tau) - X, 0) - \max(\max(S_1(\tau), S_2(\tau)) - X, 0).$$

The last term is the terminal value of a call on the maximum of two assets with strike price X. See [746, p. 376].

Exercise 15.3.6: See [492, p. 4].

Exercise 15.3.8: Ito's lemma applied to $f(S_1, S_2) = S_2/S_1$ gives

$$df = (\cdots)\, dt - \frac{S_2}{S_1}\,\sigma_1\, dW_1 + \frac{S_2}{S_1}\,\sigma_2\, dW_2.$$

From variance (6.9), we know that the variance for $df/f = d(S_2/S_1)/(S_2/S_1)$ is the σ^2 in Eq. (15.5).

Exercise 15.3.10: It is due to $dV - V_1\, dS_1 - V_2\, dS_2 = 0$.

Exercise 15.3.11: Just plug in the interpretation into formulas (9.20) for a stock paying a continuous dividend yield.

Exercise 15.3.12: (1) Let F denote the forward exchange rate τ years from now. Start with S units of domestic currency and convert it into one unit of foreign currency. Also sell forward the foreign currency. From then on, apply the standard hedging argument by using only foreign assets to construct replicating portfolios, earning the riskless rate r_{f} in foreign currency. To verify that the strategy is self-financing in terms of domestic value, observe that the forward price is $Se^{(r-r_{\mathrm{f}})\tau'}$ at τ' years remaining to expiration. The final wealth is $e^{r_{\mathrm{f}}\tau}$, identical to a call option in a foreign risk-neutral economy. This can be converted for $e^{r_{\mathrm{f}}\tau}F = Se^{r\tau}$ dollars in domestic currency because $F = Se^{(r-r_{\mathrm{f}})\tau}$. Correlation between the exchange rate and the foreign asset price is eliminated by the forward contract. (2) The results in (1) show that G_{f} grows at a rate of $r_{\mathrm{f}} - q_{\mathrm{f}} + r - r_{\mathrm{f}} = r - q_{\mathrm{f}}$ in a domestic risk-neutral economy, equivalent to a payout rate of q_{f}. Furthermore, in this economy, X_{f} becomes $X_{\mathrm{f}}F/S = X_{\mathrm{f}}e^{(r-r_{\mathrm{f}})\tau}$ at expiration.

Exercise 15.3.13: Let $S_{\mathrm{f}}(t) \equiv S(t)\,B_{\mathrm{f}}(t)$, which follows geometric Brownian motion. What we have is an exchange option that buys $S(T)\,B_{\mathrm{f}}(T)$ with $XB(T)$ at T. Note that $B(T) = B_{\mathrm{f}}(T) = 1$. Hence the desired formula is $SB_{\mathrm{f}}N(x) - XBN(x - \sigma\sqrt{\tau})$, where $x \equiv \frac{\ln(SB_{\mathrm{f}}/(XB)) + (\sigma^2/2)\tau}{\sigma\sqrt{\tau}}$, $\sigma^2 \equiv \sigma_{S_{\mathrm{f}}}^2 - 2\rho\sigma_{S_{\mathrm{f}}}\sigma_B + \sigma_B^2$, and ρ is the correlation between S_{f} and B. See [575, p. 110].

Exercise 15.3.15: Identity the U.S. dollar with the currency C in the case of foreign domestic options. The case of forex options is trivial.

Exercise 15.3.16: The portfolio has the following terminal payoff in U.S. dollars:

$$\max(S_{\mathrm{A}} - X_{\mathrm{A}}, 0) + X \times \max(X_{\mathrm{C}} - S_{\mathrm{C}}, 0).$$

Now consider all six possible relations between the spot prices (S_{A}, S_{C}, and S) and the strike prices. See [775].

Exercise 15.3.17: For simplicity, let $\widehat{S} = 1$. The exchange rate and the foreign asset's price follow

$$dS = \mu_{\mathrm{s}} S\,dt + \sigma_{\mathrm{s}} S\,dW_{\mathrm{s}}, \qquad dG_{\mathrm{f}} = \mu_{\mathrm{f}} G_{\mathrm{f}}\,dt + \sigma_{\mathrm{f}} G_{\mathrm{f}}\,dW_{\mathrm{f}},$$

respectively. The foreign asset pays a continuous dividend yield of q_{f}. Let C be the price of the quanto option. From Ito's lemma (Theorem 14.2.2),

$$\begin{aligned} dC &= \left(\mu_{\mathrm{s}} S\,\frac{\partial C}{\partial S} + \mu_{\mathrm{f}} G_{\mathrm{f}}\,\frac{\partial C}{\partial G_{\mathrm{f}}} + \frac{\partial C}{\partial t} + \frac{1}{2}\,\sigma_{\mathrm{s}}^2 S^2\,\frac{\partial^2 C}{\partial S^2} + \frac{1}{2}\,\sigma_{\mathrm{f}}^2 G_{\mathrm{f}}^2\,\frac{\partial^2 C}{\partial G_{\mathrm{f}}^2} + \rho\sigma_{\mathrm{s}}\sigma_{\mathrm{f}} S G_{\mathrm{f}}\,\frac{\partial^2 C}{\partial S\,\partial G_{\mathrm{f}}}\right) dt \\ &\quad + \sigma_{\mathrm{s}} S\,\frac{\partial C}{\partial S}\,dW_{\mathrm{s}} + \sigma_{\mathrm{f}} G_{\mathrm{f}}\,\frac{\partial C}{\partial G_{\mathrm{f}}}\,dW_{\mathrm{f}}. \end{aligned}$$

Set up a portfolio that is long one quanto option, short δ_{s} units of foreign currency, and short δ_{f} units of the foreign asset. Its value is $\Pi = C - \delta_{\mathrm{s}} S - \delta_{\mathrm{f}} G_{\mathrm{f}} S$. The total wealth change of the portfolio at time dt is given by

$$d\Pi = dC - \delta_{\mathrm{s}}\,dS - \delta_{\mathrm{f}}\,d(G_{\mathrm{f}} S) - \delta_{\mathrm{s}} S r_{\mathrm{f}}\,dt - \delta_{\mathrm{f}} G_{\mathrm{f}} S q_{\mathrm{f}}\,dt.$$

The last two terms are due to foreign interest and dividends. From Example 14.3.5,

$$d(G_{\mathrm{f}} S) = G_{\mathrm{f}} S(\mu_{\mathrm{s}} + \mu_{\mathrm{f}} + \rho\sigma_{\mathrm{s}}\sigma_{\mathrm{f}})\,dt + \sigma_{\mathrm{s}} G_{\mathrm{f}} S\,dW_{\mathrm{s}} + \sigma_{\mathrm{f}} G_{\mathrm{f}} S\,dW_{\mathrm{f}}.$$

Substitute the formulas for dC, dS, and $d(G_{\mathrm{f}}S)$ into $d\Pi$ to yield

$$\begin{aligned} d\Pi &= \left[\mu_{\mathrm{s}} S\,\frac{\partial C}{\partial S} + \mu_{\mathrm{f}} G_{\mathrm{f}}\,\frac{\partial C}{\partial G_{\mathrm{f}}} + \frac{\partial C}{\partial t} + \frac{1}{2}\,\sigma_{\mathrm{s}}^2 S^2\,\frac{\partial^2 C}{\partial S^2} + \frac{1}{2}\,\sigma_{\mathrm{f}}^2 G_{\mathrm{f}}^2\,\frac{\partial^2 C}{\partial G_{\mathrm{f}}^2}\right. \\ &\quad \left. +\ \rho\sigma_{\mathrm{s}}\sigma_{\mathrm{f}} S G_{\mathrm{f}}\,\frac{\partial^2 C}{\partial S\,\partial G_{\mathrm{f}}} - \delta_{\mathrm{s}}\mu_{\mathrm{s}} S - \delta_{\mathrm{f}} G_{\mathrm{f}} S(\mu_{\mathrm{s}} + \mu_{\mathrm{f}} + \rho\sigma_{\mathrm{s}}\sigma_{\mathrm{f}}) - \delta_{\mathrm{s}} S r_{\mathrm{f}} - \delta_{\mathrm{f}} G_{\mathrm{f}} S q_{\mathrm{f}}\right] dt. \\ &\quad + \left(\sigma_{\mathrm{s}} S\,\frac{\partial C}{\partial S} - \delta_{\mathrm{s}}\sigma_{\mathrm{s}} S - \delta_{\mathrm{f}}\sigma_{\mathrm{s}} G_{\mathrm{f}} S\right) dW_{\mathrm{s}} + \left(\sigma_{\mathrm{f}} G_{\mathrm{f}}\,\frac{\partial C}{\partial G_{\mathrm{f}}} - \delta_{\mathrm{f}}\sigma_{\mathrm{f}} G_{\mathrm{f}} S\right) dW_{\mathrm{f}}. \end{aligned}$$

Clearly, we have to pick $\delta_{\mathrm{f}} = (\partial C/\partial G_{\mathrm{f}})\,S^{-1}$ and $\delta_{\mathrm{s}} = (\partial C/\partial S) - (\partial C/\partial G_{\mathrm{f}})(G_{\mathrm{f}}/S)$ to remove the randomness. Under these choices,

$$\begin{aligned} d\Pi &= \left[\mu_{\mathrm{s}} S\,\frac{\partial C}{\partial S} + \mu_{\mathrm{f}} G_{\mathrm{f}}\,\frac{\partial C}{\partial G_{\mathrm{f}}} + \frac{\partial C}{\partial t} + \frac{1}{2}\,\sigma_{\mathrm{s}}^2 S^2\,\frac{\partial^2 C}{\partial S^2} + \frac{1}{2}\,\sigma_{\mathrm{f}}^2 G_{\mathrm{f}}^2\,\frac{\partial^2 C}{\partial G_{\mathrm{f}}^2} + \rho\sigma_{\mathrm{s}}\sigma_{\mathrm{f}} S G_{\mathrm{f}}\,\frac{\partial^2 C}{\partial S\,\partial G_{\mathrm{f}}}\right. \\ &\quad \left. - \left(\frac{\partial C}{\partial S} - \frac{\partial C}{\partial G_{\mathrm{f}}}\,\frac{G_{\mathrm{f}}}{S}\right)\mu_{\mathrm{s}} S - \frac{\partial C}{\partial G_{\mathrm{f}}}\,G_{\mathrm{f}}(\mu_{\mathrm{s}} + \mu_{\mathrm{f}} + \rho\sigma_{\mathrm{s}}\sigma_{\mathrm{f}}) - \left(\frac{\partial C}{\partial S} - \frac{\partial C}{\partial G_{\mathrm{f}}}\,\frac{G_{\mathrm{f}}}{S}\right) S r_{\mathrm{f}} - \frac{\partial C}{\partial G_{\mathrm{f}}}\,G_{\mathrm{f}} q_{\mathrm{f}}\right] dt \\ &= r\,\Pi\,dt = r\,(C - \delta_{\mathrm{s}} S - \delta_{\mathrm{f}} G_{\mathrm{f}} S)\,dt = r\left[C - \left(\frac{\partial C}{\partial S} - \frac{\partial C}{\partial G_{\mathrm{f}}}\,\frac{G_{\mathrm{f}}}{S}\right) S - \frac{\partial C}{\partial G_{\mathrm{f}}}\,G_{\mathrm{f}}\right] dt. \end{aligned}$$

https://doi.org/10.1017/CBO9780511546839.034 Published online by Cambridge University Press

After simplification,

$$rC = (r - r_{\mathrm{f}})\, S\, \frac{\partial C}{\partial S} + \frac{\partial C}{\partial t} + \frac{\sigma_{\mathrm{s}}^2 S^2}{2}\, \frac{\partial^2 C}{\partial S^2} + \frac{\sigma_{\mathrm{f}}^2 G_{\mathrm{f}}^2}{2}\, \frac{\partial^2 C}{\partial G_{\mathrm{f}}^2}$$
$$+ \rho\sigma_{\mathrm{s}}\sigma_{\mathrm{f}} S G_{\mathrm{f}}\, \frac{\partial^2 C}{\partial S\, \partial G_{\mathrm{f}}} + \frac{\partial C}{\partial G_{\mathrm{f}}}\, G_{\mathrm{f}}\,(r_{\mathrm{f}} - \rho\sigma_{\mathrm{s}}\sigma_{\mathrm{f}} - q_{\mathrm{f}}).$$

Finally, C does not directly depend on the exchange rate S [734]. Hence the equation becomes

$$rC = \frac{\partial C}{\partial t} + \frac{\sigma_{\mathrm{f}}^2 G_{\mathrm{f}}^2}{2}\, \frac{\partial^2 C}{\partial G_{\mathrm{f}}^2} + \frac{\partial C}{\partial G_{\mathrm{f}}}\, G_{\mathrm{f}}\,(r_{\mathrm{f}} - \rho\sigma_{\mathrm{s}}\sigma_{\mathrm{f}} - q_{\mathrm{f}}).$$

See [878, pp. 156–157].

Exercise 15.3.18: This is because the bond exists beyond time t^* only if $C(V, t^*) < P(t^*)$ *and* the market value if not called is exceeded by $P(t)$ at $t = t^*$ for, otherwise, it would be called at $t = t^*$. From that moment onward, the bond will be called the moment its market value if not called rises to the call price. Because the market value if not called cannot be exceeded by the conversion value, $C(V, t) \leq P(t)$. As a consequence, on a coupon date and when the bond is callable (i.e., $t > t^*$),

$$W(V, t^-) = \min(W(V - mc, t^+) + c,\ P(t))$$

because, as argued in the text, the bond should be called only when its value if not called equals the call price, which equals the value if called under $C(V, t) \leq P(t)$.

Exercise 15.4.1: Let $f_1, f_2, \ldots, f_{n+1}$ denote the prices of securities whose value depends on $S_1, S_2, \ldots, S_n$, and t. By Ito's lemma (Theorem 14.2.2),

$$\begin{aligned} df_j &= \left(\frac{\partial f_j}{\partial t} + \sum_i \boldsymbol{\mu}_i S_i\, \frac{\partial f_j}{\partial S_i} + \sum_{i,k} \frac{1}{2}\, \boldsymbol{\rho}_{ik} \boldsymbol{\sigma}_i \boldsymbol{\sigma}_k S_i S_k\, \frac{\partial^2 f_j}{\partial S_i \partial S_k} \right) dt + \sum_i \boldsymbol{\sigma}_i S_i\, \frac{\partial f_j}{\partial S_i}\, dW_i \\ &\equiv \mu_j f_j\, dt + \sum_i \sigma_{ij} f_j\, dW_i \end{aligned} \tag{33.13}$$

Maintain a portfolio of k_j units of f_j such that

$$\sum_j k_j \sigma_{ij} f_j = 0 \quad \text{for } i = 1, 2, \ldots, n. \tag{33.14}$$

Because

$$\begin{aligned} \sum_j k_j\, df_j &= \sum_j k_j \mu_j f_j\, dt + \sum_j k_j \sum_i \sigma_{ij} f_j\, dW_i \\ &= \sum_j k_j \mu_j f_j\, dt + \sum_i \sum_j k_j \sigma_{ij} f_j\, dW_i = \sum_j k_j \mu_j f_j\, dt, \end{aligned}$$

the portfolio is instantaneously riskless. Its return therefore equals the short rate, or $\sum_j k_j \mu_j f_j = r \sum_j k_j f_j$. After rearrangements,

$$\sum_j k_j f_j (\mu_j - r) = 0. \tag{33.15}$$

Equations (33.14) and (33.15) under the condition that not all k_js are zeros imply that $\mu_j - r = \sum_i \lambda_i \sigma_{ij}$ for some $\lambda_1, \lambda_2, \ldots, \lambda_n$, which depend on only $S_1, S_2, \ldots, S_n$, and t.[1] Therefore any derivative whose value depends only on $S_1, S_2, \ldots, S_n$, and t and that follows

$$\frac{df}{f} = \mu\, dt + \sum_i \sigma_i\, dW_i \tag{33.16}$$

must satisfy

$$\mu - r = \sum_i \lambda_i \sigma_i, \tag{33.17}$$

where λ_i is the market price of risk for S_i. Equation (33.17) links the excess expected return and risk. The term $\lambda_i \sigma_i$ measures the extent to which the required return on a security is affected by its dependence on S_i. From Eq. (33.13), μ and σ_i in Eq. (33.16) are

$$\begin{aligned} \mu &= \frac{\partial f}{\partial t} + \sum_i \boldsymbol{\mu}_i S_i\, \frac{\partial f}{\partial S_i} + \frac{1}{2} \sum_{i,k} \boldsymbol{\rho}_{ik} \boldsymbol{\sigma}_i \boldsymbol{\sigma}_k S_i S_k\, \frac{\partial^2 f}{\partial S_i \partial S_k}, \\ \sigma_i &= \sum_i \boldsymbol{\sigma}_i S_i\, \frac{\partial f}{\partial S_i}. \end{aligned}$$

Plugging them into Eq. (33.17) and rearranging the terms, we obtain Eq. (15.14). Risk-neutral valuation discounts the expected payoff of f at the riskless interest rate assuming that $dS_i/S_i = (\mu_i - \lambda_i \sigma_i)\, dt + \sigma_i\, dW_i$. The correlation between the dW_is is unchanged.

Exercise 15.4.2: (1) Just note that $S = X$. (2) Use the risk-neutral argument. See [894, p. 186].

Chapter 16

Exercise 16.2.1: It reduces risk (standard deviation of returns) without paying a risk premium [3, p. 38].

Exercise 16.2.2: Because the formula for $\widehat{\beta}_1$ in formula (6.14) is exactly the hedge ratio $h = \rho\delta_S/\delta_F$, resulting from plugging in Eq. (6.2) for δ_S and δ_F and plugging in Eq. (6.18) for ρ.

Exercise 16.2.3: By formula (16.2) we short $(1.25 - 2.0) \times [\,2{,}400{,}000/(600 \times 500)\,] = -6$ futures contracts, or, equivalently, long six futures contracts.

Exercise 16.3.1: (1) In the current setup, the stock price forms a continuum instead of only two states as in the binomial model. (2) This happens, for example, when their respective values are related linearly. See [119].

Exercise 16.3.3: Because of the convexity of option values, the value of the hedge, being a linear function of the stock price, loses money whichever way the price moves; one is buying stock when the stock price rises and selling it when its price falls. So the premium can be seen as the cash reserve to offset future hedging losses in order to maintain a self-financing strategy. See [593].

Alternatively, we can check that the equivalent portfolio at any time, Δ shares of stock plus B dollars of bonds, is *not* self-financing unless B exceeds the stock value by exactly the option premium. This means that we borrow money to buy the stock but have to initially invest our own money in the amount equal to the option premium.

Exercise 16.3.5: From Eq. (15.3), it is clear that $\Theta = 0$ when $\Delta = \Gamma = f = 0$.

Exercise 16.3.8: If the two strike prices used in the bull call spread are relatively close to each other, the payoff of the position will be approximately that of a binary option [514, p. 604].

CHAPTER 17

Exercise 17.1.1: Use the reflection principle [725, p. 106].

Exercise 17.1.3: Just note that for each terminal price, the number of paths that have the minimum price level M equals the number of paths that hit M minus the number of paths that hit $M - 1$. Both numbers can be obtained by the reflection principle.

Exercise 17.1.4: Set

$$\kappa \equiv \left\lfloor \frac{\ln\,(K/\,(Sd^n))}{\ln(u/d)} \right\rfloor = \left\lfloor \frac{\ln(K/S)}{2\sigma\sqrt{\Delta t}} + \frac{n}{2} \right\rfloor.$$

It is easy to see that $\tilde{K} \equiv Su^{\kappa} d^{n-\kappa}$ is the price among $Su^j d^{n-j}$ $(0 \leq j \leq n)$ closest to, but not exceeding, K. The role of the trigger price is played by the effective trigger price $\tilde{K}$ in the binomial model. Assume that the trigger price exceeds the current stock price, i.e., $2\kappa \geq n$.

Take any node A that is reachable from the root in l moves and with a price level equal to the trigger price, i.e., $Su^j d^{l-j} = \tilde{K}$ for some $0 \leq j \leq l$. The PV of the payoff is $R^{-l}(\tilde{K} - X)$. What we need to calculate is the probability that the price barrier $\tilde{K}$ has never been touched until now. Consider the node B that reaches A by way of two up moves. B's price level is $Su^{j-2} d^{l-j}$. Observe that the probability a path of length $l - 2$ that reaches node B without touching the barrier, $\tilde{K} = Su^{j-1} d^{l-2-(j-1)}$, is precisely what we are after. However, this equivalent problem is the *complement* of the ballot problem! Write B's price level as $Su^{j-2} d^{(l-2)-(j-2)}$. The desired probability can be obtained from formula (17.5) as

$$\binom{l-2}{(l-2) - 2(j-1) + (j-2)} p^{j-2}(1-p)^{(l-2)-(j-2)} = \binom{l-2}{l-j-2} p^{j-2}(1-p)^{l-j}.$$

The nature of the process dictates that $n - l$ be even. By Eq. (17.4), we have $l \geq 2\kappa - n$. Furthermore, because $Su^j d^{l-j} = Su^{\kappa} d^{n-\kappa}$, we have $j = (l - n)/2 + \kappa$. The preceding probability therefore becomes

$$\binom{l-2}{(l+n)/2-\kappa-2}\, p^{(l-n)/2+\kappa-2}(1-p)^{(l+n)/2-\kappa}.$$

We conclude that the payoff from early exercise is

$$\sum_{\substack{2\kappa-n\leq l\leq n\\ n-j \text{ is even}}} R^{-l}\binom{l-2}{(l+n)/2-\kappa-2}\, p^{(l-n)/2+\kappa-2}(1-p)^{(l+n)/2-\kappa}(\tilde{K}-X).$$

The remaining part is to add the payoff from exercise at maturity, which is just the down-and-out option with the barrier set to the trigger price.

Exercise 17.1.6: First verify that

$$\left[\frac{up}{d(1-p)}\right]^{2h-n}\to\left(\frac{H}{S}\right)^{2\lambda}, \tag{33.18}$$

where $\lambda\equiv(r+\sigma^2/2)/\sigma^2$, by plugging in the formulas for u, d, and p from Eq. (17.1) and using $\ln(1+x)\approx x$. Now,

$$\begin{aligned}
(17.6) &= S\sum_{j=a}^{2h}\binom{n}{n-2h+j}\left(\frac{up}{R}\right)^j\left[\frac{d(1-p)}{R}\right]^{n-j} - XR^{-n}\sum_{j=a}^{2h}\binom{n}{n-2h+j}p^j(1-p)^{n-j}\\
&= S\left[\frac{up}{d(1-p)}\right]^{2h-n}\sum_{j=a}^{2h}\binom{n}{n-2h+j}\left(\frac{up}{R}\right)^{n-2h+j}\left[\frac{d(1-p)}{R}\right]^{2h-j} - X\times(\cdots).
\end{aligned}$$

We analyze the first term; the second term can be handled analogously. Note that $h\leq n/2$ because $H<S$. Thus,

$$\begin{aligned}
\sum_{j=a}^{2h}\binom{n}{n-2h+j}\left(\frac{up}{R}\right)^{n-2h+j}\left[\frac{d(1-p)}{R}\right]^{2h-j} &= \sum_{j=n-2h+a}^{n}\binom{n}{j}\left(\frac{up}{R}\right)^j\left[\frac{d(1-p)}{R}\right]^{n-j}\\
&= \Phi\,(n-2h+a; n,\, pu/R)\,.
\end{aligned}$$

It is straightforward to check that $n-2h+a\approx\ln(\frac{SX}{H^2d^n})/\ln(u/d)$. Finally, apply convergence (9.18) to obtain $\Phi(n-2h+a;n,\,pu/R)\to N(x)$, where $x\equiv\frac{\ln(H^2/(SX))+(r+\sigma^2/2)\,\tau}{\sigma\sqrt{\tau}}$. With convergence (33.18), we have proved the validity of Eq. (11.4) for the no-dividend-yield case.

Exercise 17.1.7: From Lemma 9.3.3, $\ln S_\tau\sim N(\ln S+(r-q-\sigma^2/2)\,\tau,\sigma^2\tau)$ in a risk-neutral economy, where q is the continuous dividend yield. Hence, $\ln S_\tau^2\sim N(2\ln S+(2r-2q-\sigma^2)\,\tau, 2\sigma^2\tau)$. The desired formula thus equals the Black–Scholes formula after the following substitutions: $S\to S^2$, $r\to 2r+\sigma^2$, $q\to 2q$, and $\sigma\to 2\sigma$. See [845].

Programming Assignment 17.1.8: See Programming Assignment 17.3 and [624].

Programming Assignment 17.1.9: We use the current stock price and the geometric average of past stock prices as the state information each node keeps. Based on Exercise 11.7.2, the number of states at time i is approximately $i^3/2$, and backward induction over n periods gives us a running time of $O(n^4)$. Consult [249] for the subtle issue of data structures.

To reduce the running time to $O(n^3)$, we prove that the number of paths of length n having the same geometric average is precisely the number of unordered partitions of some integer into unequal parts, none of which exceeds n. Let $q(m)$ denote the number of such partitions of integer $m>0$. Any legitimate partition of m, say $\lambda\equiv(x_1,x_2,\ldots,x_k)$, satisfies $\sum_i x_i=m$, where $n\geq x_1>x_2>\cdots>x_k>0$. Interpret λ as the path of length n that makes the first up move at step $n-x_1$ (i.e., during period $n-x_1+1$), the second up move at step $n-x_2$, and so on. Each up move at step $n-x_i$ contributes x_i to the sum m. This path has a terminal geometric average of $SM^{1/(n+1)}$, where $M\equiv u^m d^{n(n+1)/2-m}$, in which the ith up move contributes u^{x_i} to the u^m term. For completeness, let $q(0)=1$. (See [26] for more information on integer partitions.) The crucial step is to observe that

$$(1+x)\,(1+x^2)\,(1+x^3)\cdots(1+x^n)=\sum_{m=0}^{n(n+1)/2}q(m)\,x^m.$$

The above polynomial clearly can be expanded in time $O(n^3)$. Thus the $q(m)$s can be computed in cubic time as well. With the parameters from Exercise 9.3.1, the option value is

$$R^{-n}\sum_{m=0}^{n(n+1)/2} 2^{-n}q(m)\times\max\left(S[\,u^m d^{n(n+1)/2-m}\,]^{1/(n+1)}-X,0\right).$$

Note that this particular parameterization makes calculating the probability of reaching any given terminal geometric average easy: $2^{-n}q(m)$. The standard parameters $u=e^{\sigma\sqrt{\tau/n}}$ and $d=1/u$ would have led to complications because of the inequality between the probabilities of up and down moves.

Programming Assignment 17.1.10: (1) Let $H_i \equiv Sd^{n-i}$, $i=0,1,\ldots,n$, and $H_{-1}\equiv-\infty$. A lookback option on the minimum on an n-period binomial tree is equivalent to a portfolio of $n+1$ barrier-type options, the ith of which is an option that knocks in at H_i but knocks out at H_{i-1}. Because this option is basically long a knock-in option with barrier H_i and short a knock-out option with barrier H_{i-1}, it can be priced in $O(n)$ time. With $n+1$ such options the total running time is $O(n^2)$. See Exercise 17.1.3 and [285]. (2) For the barrier options mentioned in (1), successive pairs are related: Each can be priced from the prior one in $O(1)$ time.

Exercise 17.1.12: This holds because h is the result of taking the floor operation; the opposite would hold if h were the result of taking the ceiling. The influence of a is negligible in comparison.

Exercise 17.1.13: Consider the j that makes $H'=Su^jd^{n-1-j}$ the largest such number not exceeding H. Call this number h'. Use $2h+1$ in place of $2h$ if $Su^{h'}d^{n-1-h'}>Su^hd^{n-h}$, as this effective barrier H' is "tighter" (closer to H) than Su^hd^{n-h}. An equivalent procedure is to test if $Su^hd^{n-1-h}\leq H$ and, if true, replace $2h$ with $2h+1$. The effect on the choice of n is that of simplification:

$$n=\left\lfloor\frac{\tau}{[\,\ln(S/H)/(j\sigma)\,]^2}\right\rfloor,\quad j=1,2,3,\ldots.$$

Exercise 17.1.15: We prove formula (17.8) with reference to Fig. 33.3. Consider any legitimate path from $(0,-\boldsymbol{a})$ to $(n,-\boldsymbol{b})$ that hits H. Let J denote the first position at which this happens. (The path may have hit the L line earlier.) By reflecting the portion of the path from $(0,-\boldsymbol{a})$ to J, a path from $(0,\boldsymbol{a})$ to $(n,-\boldsymbol{b})$ is constructed. Note that this path hits H at J. The number of paths from $(0,-\boldsymbol{a})$ to $(n,-\boldsymbol{b})$ in which an L hit is preceded by an H hit is exactly the number of paths from $(0,\boldsymbol{a})$ to $(n,-\boldsymbol{b})$ that hits L. The desired number is as claimed by application of the reflection principle.

Exercise 17.1.16: See [675, p. 7]. Note that $|B_i|$ can be derived from $|A_i|$ if $\boldsymbol{a}$ is replaced with $s-\boldsymbol{a}$ and $\boldsymbol{b}$ with $s-\boldsymbol{b}$.

Exercise 17.1.17: Apply the inclusion–exclusion principle [604].

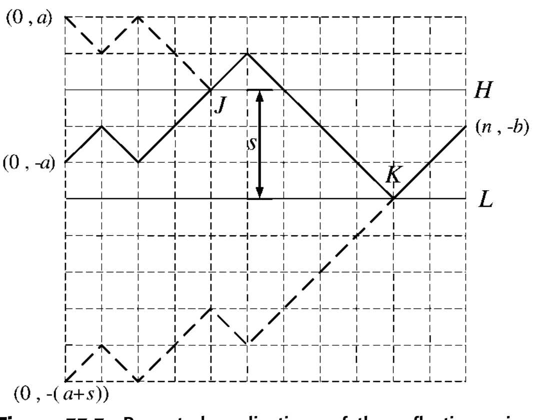


Figure 33.3: Repeated applications of the reflection principle. The random walk from $(0,-\boldsymbol{a})$ to $(n,-\boldsymbol{b})$ must hit either barrier, and there must exist an L hit preceded by an H hit. In counting the number of such walks, reflect the path first at J and then at K.

Trinomial backward induction with the diagonal method:

```
input:    S, σ, τ, X, r, n;
real      p_u, p_m, p_d, u, d, C[n+1][2n+1];
integer   i, j;
Calculate u and d;
Calculate the branching probabilities p_u, p_m, and p_d;
for (i = 2n down to 0)
        for (j = n down to ⌈i/2⌉)
            if [j = n] // Expiration date.
                Calculate C[j][i] based on Su^{j-i};
            else { // Backward induction.
                Calculate C[j][i] based on C[j+1][i],
                    C[j+1][i+1], C[j+1][i+2],
                    p_u, p_m, p_d, and Su^{j-i} (with
                    discounting factor e^{-rτ/n});
            }
return C[0][0];
```

Figure 33.4: Trinomial backward induction with the diagonal method. $C[j][i]$ represents the derivative value at time j if the stock price makes $j-i$ more up moves than down moves; it corresponds to the node with the $(i+1)$th largest underlying asset price. Recall that $ud = 1$. The space requirement can be further reduced by using a 1-dimensional array for C.

Exercise 17.1.18: (1) A standard European call is equivalent to a double-barrier option and a double-barrier option that knocks out if neither barrier is hit. (2) Construct a portfolio of long a down-and-in option, long an up-and-in option, and short the double-barrier option. (3) It equals $\sum_{i=2}^{n}(-1)^i$ $(|A_i|+|B_i|)$.

Exercise 17.1.19: Consider the process $Y(t) \equiv [X(t) - f_\ell(t)]/[f_\mathrm{h}(t) - f_\ell(t)]$, where $X(t) \equiv \ln S(t)$. The payoff of the option at maturity becomes $\max(e^{[f_\mathrm{h}(T)-f_\ell(T)]Y(T)+f_\ell(T)} - X, 0)$ with barriers at 0 and 1. See [757].

Programming Assignment 17.1.20: Calculate $\binom{n}{0}, \binom{n}{1}, \binom{n}{2}, \ldots, \binom{n}{n}$ and store them for random access when needed by the $|A_i|$s and the $|B_i|$s. Note that the calculation of $\boldsymbol{N}(\boldsymbol{a}, \boldsymbol{b}, s)$ takes time $O(n/s)$ because $|A_i| = |B_i| = 0$ for $i > n/s$. The implication is that expression (17.10) only has $(h-a)\,\frac{n}{2(h-l)} < n/2$ nonzero terms because $l < a$.

Exerciset 17.2.2: Consider one period to maturity with $S = X$.

Programming Assignment 17.2.3: See the algorithm in Fig. 33.4.

Exercise 17.2.4: See [745, p. 21].

Exercise 17.2.5: We focus on the down-and-in call with barrier $H < X$. Assume without loss of generality that $H < S$ for, otherwise, the option is identical to a standard call. Under the trinomial model, there are $2n+1$ stock prices Su^j $(-n \le j \le n)$. Let

$$a \equiv \left\lceil \frac{\ln(X/S)}{\lambda\sigma\sqrt{\Delta t}} \right\rceil, \qquad h \equiv \frac{\ln(S/H)}{\lambda\sigma\sqrt{\Delta t}} > 0.$$

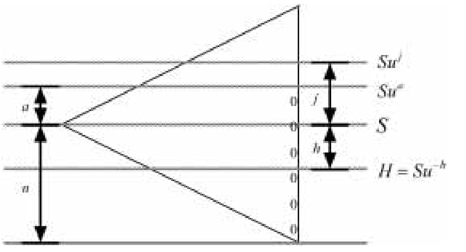
A process with n moves ends up at a price at or above X if and only if the number of up moves exceeds that of down moves by at least a because $Su^{a-1} < X \le Su^a$. The starting price is separated from the barrier by h down moves because $Su^{-h} = H$ (see Fig. 33.5). The following formula is the pricing formula for European down-and-in calls:

$$R^{-n} \sum_{m=0}^{n-2h-a} \sum_{j \ge \max(a, m-n)}^{n-m-2h} \frac{n!}{[(n-m+j+2h)/2]!\, m!\, [(n-m-j-2h)/2]!}$$
$$p_u^{(n-m+j)/2} p_m^m p_d^{(n-m-j)/2} (Su^j - X). \tag{33.19}$$

This is an alternative characterization of the trinomial tree algorithm for down-and-in calls.

https://doi.org/10.1017/CBO9780511546839.034 Published online by Cambridge University Press

Figure 33.5: Down-and-in calls under trinomial tree. Note that the interpretations of a, j, and h differ from those in Fig. 17.2.



We use the reflection principle for trinomial random walks to prove the correctness of pricing formula (33.19). Consider paths that hit the barrier and end up at Su^j with m level moves. Any such path must contain, cumulatively, $(n-m+j)/2$ up moves, m level moves, and $(n-m-j)/2$ down moves with probability $p_u^{(n-m+j)/2} p_m^m p_d^{(n-m-j)/2}$. The reflection principle tells us that their number equals the number of paths from $(0, -h)$ to $(n, j+h)$, which is

$$\frac{n!}{[\,(n-m+j+2h)/2\,]!\, m!\, [\,(n-m-j-2h)/2\,]!}$$

because the number of up moves is $(n-m+j+2h)/2$, the number of down moves is $(n-m-j-2h)/2$, and the number of level moves is m.

The correctness of the bounds on m (the number of level moves) and j (the number of up moves minus the number of down moves) in pricing formula (33.19) can be verified as follows. The former cannot exceed $n-2h-a$ as it takes $2h+a$ nonlevel moves to hit the barrier and end up at a price of at least X. That $(n-m+j+2h)/2$, $(n-m-j-2h)/2$, $(n-m+j)/2$, and $(n-m-j)/2$ are integers means that $n-m+j$ must be even. That j must be at least a is obvious. The bound $j \geq m-n$ is necessitated by $(n-m+j)/2 \geq 0$. The bound $j \leq n-m-2h$ is needed because $(n-m-j-2h)/2 \geq 0$. Finally, we can easily check that all the terms are now nonnegative integers within their respective bounds.

Exercise 17.3.1: The number of nodes at time n is $1+4+8+12+16+\cdots+4n$.

Exercise 17.3.2: The covariance between R_1 and R_2, or $E[\,R_1 R_2\,] - \mu_1\mu_2$, equals

$$\sigma_1\sigma_2(p_1 - p_2 - p_3 + p_4) = \sigma_1\sigma_2[p_1 - (1 - p_1 - p_4) + p_4] = \sigma_1\sigma_2(2p_1 + 2p_4 - 1).$$

Hence their correlation is $2p_1 + 2p_4 - 1$, as claimed.

Exercise 17.3.5: We use asset 1 to denote the stock S and asset 2 to denote the futures F. Let h denote the hedge ratio and the stock price has a correlation of ρ with the derivative's price. The variance at time Δt is $\delta_S^2 + h^2\delta_F^2 - 2h\rho\delta_S\delta_F$, where δ_S denotes the standard deviation of the stock price and δ_F the standard deviation of the futures price at time Δt. The optimal hedge ratio is $h = \rho(\delta_S/\delta_F) = E[\,(S - S\mu_1)(F - F\mu_2)\,]/\delta_F^2$ as shown in Eq. (16.1), where μ_1 and μ_2 are the expected gross returns of the respective assets at Δt. The formula for h is

$$\frac{(Fu_2 - F\mu_2)[\,p_1(Su_1 - S\mu_1) + p_3(Sd_1 - S\mu_1)\,] + (Fd_2 - F\mu_2)[\,p_2(Su_1 - S\mu_1) + p_4(Sd_1 - S\mu_1)\,]}{(p_1 + p_3)(Fu_2 - F\mu_2)^2 + (p_2 + p_4)(Fd_2 - F\mu_2)^2}.$$

Of course, there is no stopping at stocks and futures. For instance, we can use index futures to hedge equity options dynamically, and the same formula applies!

Chapter 18

Exercise 18.1.1: See [706, p. 135].

Exercise 18.1.2: $U_{i,T} = \max(0,\, X - e^{V_{\min} + i\times\Delta V})$.

Exercise 18.1.4: After discretization, $U_{i,j-1} = aU_{i+1,j} + bU_{i,j} + cU_{i-1,j}$, where

$$a \equiv \left[\left(\frac{\sigma S_i}{\Delta S}\right)^2 + \frac{rS_i}{\Delta S}\right]\frac{1}{2\Delta t}, \qquad b \equiv \left[-r + \frac{1}{\Delta t} - \left(\frac{\sigma S_i}{\Delta S}\right)^2\right]\Delta t, \qquad c \equiv \left[\left(\frac{\sigma S_i}{\Delta S}\right)^2 - \frac{rS_i}{\Delta S}\right]\frac{1}{2\Delta t},$$

and S_i is the stock price for $U_{i,j}$. Two conditions must be met: (1) $1/\Delta t > (\sigma S_i/\Delta S)^2 + r$ to ensure that $b > 0$ and (2) $\sigma^2 S_i/r > \Delta S$ to ensure that $a > 0$ and $c > 0$. See [897].

Exercise 18.2.2: The node with the maximum discounted intrinsic value may not be exercised [129].

Exercise 18.2.4: From Exercise 14.3.1 we know that $d(\ln X) = (\mu - \sigma^2/2)\,dt + \sigma\,dW$, a Brownian motion process. Because Brownian motion's sample paths can be generated without loss of accuracy once we have access to a perfect random-number generator, a better algorithm generates sample paths for $\ln X$ and then turns them into ones for X by taking exponentiation. Specifically, $X_i = X_{i-1} e^{(\mu-\sigma^2/2)\,\Delta t + \sigma\sqrt{\Delta t}\,\xi}$, where $\xi \sim N(0, 1)$.

Exercise 18.2.7: Observe that $\text{Prob}[\,X < -y\,] = 1 - \text{Prob}[\,X < y\,]$ if X is normally distributed with mean zero. Hence the method alluded to does indeed fall into the category of the antithetic-variates method by the inverse-transform method. It also works for any random variable with a density function that is symmetric around zero.

Programming Assignment 18.2.9: Generalize the model for stochastic volatility in Section 15.5 as follows. In a risk-neutral economy, $dS = r\,S\,dt + \sigma\,S(\rho\,dW_v + \sqrt{1-\rho^2}\,dW_s)$ and $d\sigma^2 = \mu\,dt + \sigma_v\,dW_v$, where $\mu(t)$ and $\sigma_v(t)$ may depend on the history of $\sigma^2(t)$, and $W_v(t)$ and $W_s(t)$ are uncorrelated. Now,

$$
\begin{aligned}
S_T &= S_0 \times \exp\left[\, r\,T - \frac{1}{2}\int_0^T \sigma^2\,du + \rho\int_0^T \sigma\,dW_v + \sqrt{1-\rho^2}\int_0^T \sigma\,dW_s \,\right] \\
&= S_0 \times \exp\left[\, -\frac{\rho^2}{2}\int_0^T \sigma^2\,du + \rho\int_0^T \sigma\,dW_v \,\right] \\
&\quad \times \exp\left[\, r\,T - \frac{1-\rho^2}{2}\int_0^T \sigma^2\,du + \sqrt{1-\rho^2}\int_0^T \sigma\,dW_s \,\right].
\end{aligned}
$$

If the path of W_v is known, then $\xi \equiv \exp[\,-\frac{\rho^2}{2}\int_0^T \sigma^2\,du + \rho\int_0^T \sigma\,dW_v\,]$ is a constant and $\ln S_T$ is normally distributed with mean $\ln(S_0\xi) + r\,T - \frac{1-\rho^2}{2}\int_0^T \sigma^2\,du$ and variance $(1-\rho^2)\int_0^T \sigma^2\,du$, in which case the Black–Scholes formula applies. The algorithmic idea is now clear. Simulate W_v to obtain a path of σ values. Use the analytic solution to obtain the option value conditional on that path. Repeat a few times and average. There is no need to simulate W_s. See [876].

Exercise 18.2.10: They are $\text{Var}[\,Y\,] < 2\,\text{Cov}[\,X, Y\,]$ for $\beta = -1$ and $\text{Var}[\,Y\,] < -2\,\text{Cov}[\,X, Y\,]$ for $\beta = 1$. Both are more stringent than relation (18.6). See [584].

Exercise 18.2.11: This scheme will make $\text{Cov}[\,X, Y\,] = 0$ and thus $\text{Var}[\,W\,] > \text{Var}[\,X\,]$.

Programming Assignment 18.2.12: See the algorithm in Fig. 33.6.

Control variates for pricing average-rate calls on a non-dividend-paying stock:

```
input:    S, X, n, r, σ, τ, m;
real      P, C, M₁, M₂;
real      ξ();   //ξ() ~ N(0, 1).
integer   i, j;
C := 0;
for (i = 1 to m) {
          P := S; M₁ := S; M₂ := S^{1/(n+1)};
          for (j = 1 to n) {
                P := Pe^{(r−σ²/2)(τ/n)+σ√(τ/n) ξ()};
                M₁ := M₁ + P;
                M₂ := M₂ × P^{1/(n+1)};
          }
          C := C + e^{−rτ} × max(M₁/(n + 1) − X, 0)−
                (e^{−rτ} × max(M₂ − X, 0) − analytic value);
}
return C/m;
```

Figure 33.6: Control-variates method for arithmetic average-rate calls. m is the number of replications, n is the number of periods, M_1 is the arithmetic average, and M_2 is the geometric average. The `analytic value` is computed by Eq. (11.8) for the geometric average-rate option. In practice, the option value under the binomial model may be preferred as the `analytic value` (but it takes more time, however, to calculate). Note carefully that the expected value of the control variate is not exactly M_2.

Exercise 18.2.13: Let $\epsilon \equiv |B|/|A|$. The Monte Carlo approach's probability of failure is clearly $(1-\epsilon)^N$. The refined search scheme, on the other hand, has a probability of failure equal to $\prod_i (1-\epsilon_i)^{N/m}$, where ϵ_i is the proportion of A_i that intersects B, in other words, $\epsilon_i \equiv |B \cap A_i|/|A_i|$. Clearly, $\sum_i \epsilon_i = m\epsilon$. Now, $[\prod_i (1-\epsilon_i)]^{N/m} \leq (\sum_i \frac{1-\epsilon_i}{m})^N = (1-\epsilon)^N$, where the inequality is by the relation between the arithmetic average and the geometric average.

Exercise 18.3.1: They are $\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{4}{9}, \frac{7}{9}, \frac{2}{9}, \frac{5}{9}, \frac{8}{9}, \frac{1}{27}, \frac{10}{27}$.

Chapter 19

Exercise 19.1.1: This follows from the definition and $c_{ii} = \sigma_i^2$, the variance of x_i.

Exercise 19.1.2: Consider any normalized linear combination $\sum_i b_i x_i = b^{\mathrm{T}} \boldsymbol{x}$, where $\sum_i b_i^2 = 1$. Let $b = Be$ for some unit-length vector $e \equiv [e_1, e_2, \dots, e_n]^{\mathrm{T}}$. Then $b^{\mathrm{T}}\boldsymbol{x} = e^{\mathrm{T}} B^{\mathrm{T}} \boldsymbol{x} = e^{\mathrm{T}} P$. Finally,

$$\begin{aligned} \mathrm{Var}[e^{\mathrm{T}} P] &= E[(e^{\mathrm{T}} P)^2] = \sum_{i=1}^{n} e_i^2 \, E[p_i^2] \quad \text{from the uncorrelatedness of the } p_i\text{s} \\ &= \lambda_1 + \sum_{i=2}^{n} e_i^2 (\lambda_i - \lambda_1) \quad \text{because } E[p_i^2] = \lambda_i \text{ and } e_1^2 = 1 - \textstyle\sum_{i=2}^{n} e_i^2. \end{aligned}$$

The preceding equation achieves the maximum at λ_1 when $e_2 = e_3 = \cdots = e_n = 0$, that is, when $e^{\mathrm{T}} P = p_1$. In general, p_j is the normalized linear combination of the x_is, which is uncorrelated with $p_1, p_2, \dots, p_{j-1}$ and has the maximum variance. This can be verified as follows. Uncorrelatedness implies that $e_1 = e_2 = \cdots = e_{j-1} = 0$. Hence

$$\mathrm{Var}[e^{\mathrm{T}} P] = \sum_{i=j}^{n} e_i^2 \, E[p_i^2] = \lambda_j + \sum_{i=j+1}^{n} e_i^2 (\lambda_i - \lambda_j).$$

The preceding equation achieves the maximum at λ_j when $e_{j+1} = e_{j+2} = \cdots = e_n = 0$, that is, when $e^{\mathrm{T}} P = p_j$.

Exercise 19.1.3: Note that

$$\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} 1 & \rho \\ 0 & \sqrt{1-\rho^2} \end{bmatrix}.$$

Exercise 19.1.4: See Exercise 6.1.3.

Exercise 19.1.5: (1) Let $B \equiv \sum_{i=1}^{p} \sigma_i u_i v_i^{\mathrm{T}}$. It is sufficient to prove that $Av_i = Bv_i$ for $i = 1, 2, \dots, n$ because $\{v_1, v_2, \dots, v_n\}$ forms a basis of $\boldsymbol{R}^n$. Now $Av_i = \sigma_i u_i$ and $Bv_i = \sum_j \sigma_j u_j (v_j^{\mathrm{T}} v_i) = \sigma_i u_i$. See [870, p. 395]. (2) It is because $Av_i = \sigma_i u_i$ and $A^{\mathrm{T}} u_i = \sigma_i v_i$. See [392, p. 258], or [870, p. 392].

Exercise 19.2.1: From Eq. (19.5), we have the following LS problem for linear regression:

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

The normal equations are then

$$\begin{bmatrix} m & \sum_i x_i \\ \sum_i x_i & \sum_i x_i^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \sum_i y_i \\ \sum_i x_i y_i \end{bmatrix}.$$

See [381, p. 263].

Exercise 19.2.2: (1) Let the data be

$$\{(x_{1,1}, x_{1,2}, \dots, x_{1,n}, y_1), (x_{2,1}, x_{2,2}, \dots, x_{2,n}, y_2), \dots, (x_{m,1}, x_{m,2}, \dots, x_{m,n}, y_m)\}.$$

We desire to fit the data to the linear model $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$. The LS formulation is

$$\begin{bmatrix} 1 & x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ 1 & x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{m,1} & x_{m,2} & \cdots & x_{m,n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

(2) It is

$$\begin{bmatrix} m & \sum_{i=1}^m x_{i,1} & \sum_{i=1}^m x_{i,2} & \cdots & \sum_{i=1}^m x_{i,n} \\ \sum_{i=1}^m x_{i,1} & \sum_{i=1}^m x_{i,1}x_{i,1} & \sum_{i=1}^m x_{i,1}x_{i,2} & \cdots & \sum_{i=1}^m x_{i,1}x_{i,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^m x_{i,n} & \sum_{i=1}^m x_{i,n}x_{i,1} & \sum_{i=1}^m x_{i,n}x_{i,2} & \cdots & \sum_{i=1}^m x_{i,n}x_{i,n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_{i,1}y_i \\ \vdots \\ \sum_{i=1}^m x_{i,n}y_i \end{bmatrix}.$$

Exercise 19.2.3: See [392, p. 223].

Exercise 19.2.4: Let $A \equiv [a_{ij}] \in \boldsymbol{R}^{m\times n}$. Then $a_{ij}a_{ik} \neq 0$ only if $|j-k| < \omega$. Hence, the (j,k)th entry of $A^{\mathrm{T}}A$, which is $\sum_{i=1}^m a_{ij}a_{ik}$, equals 0 for $|j-k| \geq \omega$. See [75, p. 217].

Exercise 19.2.5: Because $A^{\mathrm{T}}(Ax_{\mathrm{LS}} - b) = 0$ [415, p. 201].

Exercise 19.2.6: Let $C = PP^{\mathrm{T}}$ be the Cholesky decomposition, where P is nonsingular. With $Py = b$ and $P\epsilon' = \epsilon$, we have transformed the problem into $Py = Ax + P\epsilon'$ or, equivalently, $y = P^{-1}Ax + \epsilon'$. We claim that $y = P^{-1}Ax$ is the LS problem to solve; in other words, we regress $P^{-1}b$ on $P^{-1}A$. This can be verified by

$$\mathrm{Cov}[\epsilon'] = \mathrm{Cov}[P^{-1}\epsilon] = P^{-1}\mathrm{Cov}[\epsilon]\left(P^{\mathrm{T}}\right)^{-1} = P^{-1}\sigma^2 C\left(P^{\mathrm{T}}\right)^{-1} = \sigma^2 C.$$

See Exercise 6.1.3 and [802, p. 54].

Exercise 19.2.7: The covariance matrix of x_{LS} is

$$\begin{aligned}(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}E[bb^{\mathrm{T}}]A[(A^{\mathrm{T}}A)^{-1}]^{\mathrm{T}} &= \sigma^2(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}A[(A^{\mathrm{T}}A)^{-1}]^{\mathrm{T}} \\ &= \sigma^2(A^{\mathrm{T}}A)^{-1}[A^{\mathrm{T}}A(A^{\mathrm{T}}A)^{-1}]^{\mathrm{T}} \\ &= \sigma^2(A^{\mathrm{T}}A)^{-1}.\end{aligned}$$

Exercise 19.2.8: Because both U and V are orthogonal, we have $U^{-1} = U^{\mathrm{T}}$ and $V^{-1} = V^{\mathrm{T}}$. Expand $(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ and use the property $\Sigma^+ = \Sigma^{-1}$ to obtain the desired result.

Exercise 19.2.9: Just do the transformation $A = U\Sigma V^{\mathrm{T}} \to A^+ = V\Sigma^+U^{\mathrm{T}}$ twice.

Exercise 19.2.10: This can be verified informally as follows. To begin with,

$$A\widehat{x} = AA^+b + AV_2y = U\Sigma V^{\mathrm{T}}V\Sigma^+U^{\mathrm{T}}b + U\Sigma V^{\mathrm{T}}V_2y = b + U\mathbf{0}y = b.$$

We complete the proof by noting that our solution set has dimension $n - m$. See [586].

Exercise 19.2.11: By Exercise 19.2.6 we should solve $y = P^{-1}Ax$ for x, where $C = PP^{\mathrm{T}}$ and $Py = b$. By normal equations (19.6), the solution must satisfy

$$(P^{-1}A)^{\mathrm{T}}(P^{-1}A)x = (P^{-1}A)^{\mathrm{T}}y = (P^{-1}A)^{\mathrm{T}}P^{-1}b.$$

After expansion, the preceding identity becomes $A^{\mathrm{T}}(P^{-1})^{\mathrm{T}}P^{-1}Ax = A^{\mathrm{T}}(P^{-1})^{\mathrm{T}}P^{-1}b$, which implies that $A^{\mathrm{T}}C^{-1}Ax = A^{\mathrm{T}}C^{-1}b$, as claimed.

Exercise 19.2.12: Let $B = U\Sigma V^{\mathrm{T}}$ be the SVD of B and $V \equiv [\underbrace{V_1}_{p}, \underbrace{V_2}_{n-p}]\}n$. The solution to the constrained problem is [586]

$$B^+d + V_2(AV_2)^+(b - AB^+d). \tag{33.20}$$

Solution (33.20) clearly satisfies the constraint $Bx = d$ because all solutions to $Bx = d$ are of the form $B^+d + V_2y$ for arbitrary $y \in \boldsymbol{R}^{n-p}$ by virtue of Eq. (19.10). An algorithm implementing solution (33.20) is given in Fig. 33.7. We verify its correctness as follows. After step 1, $U^{\mathrm{T}}BV = \Sigma$. Observe that $B^+BV_1 = V_1$ even if $B^+B \neq I$. This holds because, by pseudoinverse (19.9),

$$\begin{aligned}B^+BV_1 &= V\Sigma^+U^{\mathrm{T}}U\Sigma V^{\mathrm{T}}V_1 \\ &= V\Sigma^+\Sigma V^{\mathrm{T}}V_1 \\ &= V\Sigma^+ \times \mathrm{diag}_{p\times p}[\sigma_1, \sigma_2, \ldots, \sigma_p] \\ &= V \times \mathrm{diag}_{n\times p}[\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_p^{-1}] \times \mathrm{diag}_{p\times p}[\sigma_1, \sigma_2, \ldots, \sigma_p] \\ &= V \times \mathrm{diag}_{n\times p}[1, 1, \ldots, 1] \\ &= V_1.\end{aligned}$$

Therefore, after step 3, $B^+d = B^+BV_1x_1 = V_1x_1$. Step 4 makes $\widehat{b} = b - AV_1x_1 = b - AB^+d$. After step 5, $x_2 = (AV_2)^+\widehat{b} = (AV_2)^+(b - AB^+d)$. Finally the returned value is $V_1x_1 + V_2x_2 = B^+d + V_2(AV_2)^+(b - AB^+d)$, matching solution (33.20).

Exercise 19.2.13: See [586, p. 139].

Algorithm for the LS problem with linear equality constraints:

```
input:  A[m][n], b[m], B[p][n], d[p], m, n, p (p ≤ n);
real    x1[p], x2[n−p], b̂[m], U[p][p], V[n][n], Σ[p][n];
1.      Compute the SVD of B, B = UΣV^T;
2.      Partition V ≡ [ V1 (p), V2 (n−p) ]} n;
3.      Solve BV1x1 = d for x1;
4.      b̂ := b − AV1x1;
5.      Solve the LS problem AV2x2 = b̂ for x2;
6.      return V1x1 + V2x2;
```

Figure 33.7: Algorithm for the LS problem with linear equality constraints. See the text for the meanings of the input variables. Because only the diagonal elements of Σ are needed, a one-dimensional array suffices (the needed change to the algorithm is straightforward).

Exercise 19.2.14: Define $W \equiv \text{diag}[\,1/\sqrt{\Psi_1}, 1/\sqrt{\Psi_2}, \dots, 1/\sqrt{\Psi_p}\,]$. It plays the role of W in weighted LS problem (19.11). In other words, we claim that our problem is equivalent to

$$\min_{x\in R^n} \| WLf_t - W(y_t - \mu) \| . \tag{33.21}$$

This granted, because $\Psi^{-1} = W^{\mathrm{T}}W$, the solution for f_t is $(L^{\mathrm{T}}\Psi^{-1}L)^{-1}L^{\mathrm{T}}\Psi^{-1}(y_t - \mu)$ by formula (19.13), as desired. Hence we need to verify only equivalence (33.21). Because the covariance matrix of ϵ_t in Eq. (19.15) is Ψ, the discussion leading to Eq. (19.14) says that the inverse of $W^{\mathrm{T}}W$ should equal that matrix, which is true.

Exercise 19.2.15: $\text{Cov}[\,y_t, f_t^{\mathrm{T}}\,] = E[\,(y_t - \mu) f_t^{\mathrm{T}}\,] = E[\,(Lf_t + \epsilon_t) f_t^{\mathrm{T}}\,] = E[\,Lf_t f_t^{\mathrm{T}}\,] = L$ [523, p. 399].

Exercise 19.2.16: See [239, 390].

Exercise 19.3.1: First we derive

$$\begin{aligned} f''(x) &= f'_{i-1}\,\frac{6x - 4x_i - 2x_{i-1}}{h_i^2} - f'_i\,\frac{-6x + 2x_i + 4x_{i-1}}{h_i^2} \\ &\quad + y_{i-1}\,\frac{12x - 6x_i - 6x_{i-1}}{h_i^3} + y_i\,\frac{-12x + 6x_i + 6x_{i-1}}{h_i^3} \end{aligned} \tag{33.22}$$

for $x \in [\,x_{i-1}, x_i\,]$. For the $f''(x_i-)$ case, we simply evaluate $f''(x_i)$ as x_i is the right end point of Eq. (33.22) for $f''(x)$. As for the $f''(x_i+)$ case, we first derive an analogous formula for $f''(x)$, $x \in [\,x_i, x_{i+1}\,]$. It is obvious that the desired formula is merely Eq. (33.22) but with i replaced with $i+1$. Finally, we evaluate $f''(x_i)$ for the answer as x_i is the left end point of the formula in question. See [447, p. 479].

Exercise 19.3.2: With the help of Eq. (19.16), $f''(x_0) = f''(x_n) = 0$ implies that

$$\begin{aligned} f''(x_0) &= -\frac{2}{h_1}\,(2f'_0 + f'_1) + 6\,\frac{y_1 - y_0}{h_1^2}, \\ f''(x_n) &= \frac{2}{h_n}\,(f'_{n-1} + 2f'_n) - 6\,\frac{y_n - y_{n-1}}{h_n^2}. \end{aligned}$$

Hence

$$\begin{bmatrix} \frac{2}{h_1} & \frac{1}{h_1} & 0 & 0 & 0 & \cdots & 0 \\ \frac{1}{h_1} & \frac{2}{h_1} + \frac{2}{h_2} & \frac{1}{h_2} & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{h_2} & \frac{2}{h_2} + \frac{2}{h_3} & \frac{1}{h_3} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \frac{1}{h_{n-1}} & \frac{2}{h_{n-1}} + \frac{2}{h_n} & \frac{1}{h_n} \\ 0 & \cdots & \cdots & 0 & 0 & \frac{1}{h_n} & \frac{2}{h_n} \end{bmatrix} \begin{bmatrix} f'_0 \\ f'_1 \\ \vdots \\ f'_n \end{bmatrix} = \begin{bmatrix} 3\,\frac{y_1 - y_0}{h_1^2} \\ 3\,\frac{y_1 - y_0}{h_1^2} + 3\,\frac{y_2 - y_1}{h_2^2} \\ 3\,\frac{y_2 - y_1}{h_2^2} + 3\,\frac{y_3 - y_2}{h_3^2} \\ \vdots \\ 3\,\frac{y_{n-1} - y_{n-2}}{h_{n-1}^2} + 3\,\frac{y_n - y_{n-1}}{h_n^2} \\ 3\,\frac{y_n - y_{n-1}}{h_n^2} \end{bmatrix}.$$

https://doi.org/10.1017/CBO9780511546839.034 Published online by Cambridge University Press

Exercise 19.3.4: 19.3.4: This time we solve

$$\begin{bmatrix} 4 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} f_1'' \\ f_2'' \\ f_3'' \end{bmatrix} = \begin{bmatrix} -16.2 \\ 26.4 \\ -28.2 \end{bmatrix}.$$

The solutions are $f_1'' = -6.72857$, $f_2'' = 10.7143$, and $f_3'' = -9.72857$. Thus

$$\begin{aligned}
p_1(x) &= 2x + \frac{6.72857}{6}(x - x^3), \\
p_2(x) &= 2(2-x) + 1.3\,(x-1) + \frac{6.72857}{6}[\,(2-x)-(2-x)^3\,] - \frac{10.7143}{6}[\,(x-1)-(x-1)^3\,], \\
p_3(x) &= 1.3\,(3-x) + 5(x-2) - \frac{10.7143}{6}[\,(3-x)-(3-x)^3\,] + \frac{9.72857}{6}[\,(x-2)-(x-2)^3\,], \\
p_4(x) &= 5(4-x) + 4(x-3) + \frac{9.72857}{6}[\,(4-x)-(4-x)^3\,].
\end{aligned}$$

Exercise 19.3.5: B is

$$\begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 & -1 & -x_1 & -x_1^2 & -x_1^3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & x_2 & x_2^2 & x_2^3 & -1 & -x_2 & -x_2^2 & -x_2^3 & 0 & \cdots \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix},$$

d is the zero vector, the jth row of $A \in \boldsymbol{R}^{m\times 4n}$ is

$$[\,\overbrace{0,\ldots,0}^{4(i-1)}, 1, \tilde{x}_j, \tilde{x}_j^2, \tilde{x}_j^3, \overbrace{0,\ldots,0}^{4(m-i)}\,]^{\mathrm{T}}$$

when $\tilde{x}_j$ falls within $[\,x_{i-1}, x_i\,]$, and $b \equiv [\,\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_m\,]^{\mathrm{T}}$.

Chapter 20

Exercise 20.1.1: As $\Delta S/S = \mu\Delta t + \sigma\sqrt{\Delta t}\,\xi$, we have $\Delta S_i - S_i\mu\Delta t \sim N(0, S^2\sigma^2\Delta t)$. The log-likelihood function is

$$\sum_{i=1}^{n} -\ln\left(2\pi S^2\sigma^2\Delta t\right) - \sum_{i=1}^{n} \frac{(\Delta S_i - S_i\mu\Delta t)^2}{2S^2\sigma^2\Delta t}.$$

After differentiation with respect to μ and σ^2, the estimators are found to be

$$\widehat{\mu} = \frac{\sum_{i=1}^{n}(\Delta S_i/S_i)}{n\Delta t}, \qquad \widehat{\sigma^2} = \frac{\sum_{i=1}^{n}[\,(\Delta S_i/S_i) - \widehat{\mu}\Delta t\,]^2}{n\Delta t}.$$

They are essentially the estimators of Eqs. (20.2) and (20.3) with R_i replaced with $(S_{i+1} - S_i)/S_i$ and α replaced with μ. Estimating μ remains difficult.

Exercise 20.1.2: Exercise 14.4.7 shows that the simple rate of return has mean $\mu_s = e^{t\mu} - 1$ and variance $\sigma_s^2 = (e^{t\sigma^2} - 1)\,e^{2t\mu}$. Hence

$$\sigma = \sqrt{\frac{1}{t}\ln\left(1 + \frac{\sigma_s^2}{(1+\mu_s)^2}\right)}.$$

Published statistics for stock returns are mostly based on simple rates of return [147, p. 362]. Simple rates of return also should be used when calculating the VaR in Section 31.4.

Exercise 20.1.3: Equation (20.5) can be simplified to

$$0 = \sum_i (\Delta r_i - \beta(\mu - r_i)\,\Delta t)\, r_i^{-1} \tag{33.23}$$

with the help of Eq. (20.6) and the assumption that $\mu \neq 0$. Multiply out Eqs. (33.23) and (20.6) to obtain

$$\begin{aligned}
0 &= \sum_i \Delta r_i r_i^{-1} - \beta\mu\Delta t \sum_i r_i^{-1} + \beta n\Delta t, \\
0 &= \sum_i \Delta r_i r_i^{-2} - \beta\mu\Delta t \sum_i r_i^{-2} + \beta\Delta t \sum_i r_i^{-1}.
\end{aligned}$$

Eliminate μ to gain a formula for β. Then plug the formula into one of the preceding equations to obtain the formula for μ. Finally, μ and β can be used in Eq. (20.7) to calculate σ^2.

Exercise 20.1.4: From Eq. (20.4) and Ito's lemma, $d\ln r = [\frac{\beta(\mu-r)}{r} - \frac{1}{2}\sigma^2]\,dt + \sigma\,dW$. The discrete approximation thus satisfies

$$\ln\left(1+\frac{\Delta r}{r}\right) - \left(\frac{\beta(\mu-r)}{r} - \frac{1}{2}\sigma^2\right)\Delta t \sim N\left(0, \sigma^2\Delta t\right).$$

The log-likelihood function of n observations $\Delta r_1, \Delta r_2, \dots, \Delta r_n$ after the removal of the constant terms and simplification is

$$-n\ln\sigma - \left(2\sigma^2\Delta t\right)^{-1}\sum_{i=1}^{n}\left\{\ln\left(1+\frac{\Delta r_i}{r_i}\right) - \left[\frac{\beta(\mu-r_i)}{r_i} - \frac{1}{2}\sigma^2\right]\Delta t\right\}^2.$$

We differentiate the log-likelihood function with respect to β, μ, and σ and equate them to zero. After simplification, we arrive at

$$0 = \sum_{i=1}^{n}\left\{\ln\left(1+\frac{\Delta r_i}{r_i}\right) - \left[\frac{\beta(\mu-r_i)}{r_i} - \frac{1}{2}\sigma^2\right]\Delta t\right\} r_i^{-1}, \tag{33.24}$$

$$0 = \sum_{i=1}^{n}\left\{\ln\left(1+\frac{\Delta r_i}{r_i}\right) - \left[\frac{\beta(\mu-r_i)}{r_i} - \frac{1}{2}\sigma^2\right]\Delta t\right\}, \tag{33.25}$$

$$\sigma^2 = \frac{1}{n\Delta t}\sum_{i=1}^{n}\left\{\ln\left(1+\frac{\Delta r_i}{r_i}\right) - \left[\frac{\beta(\mu-r_i)}{r_i} - \frac{1}{2}\sigma^2\right]\Delta t\right\}\ln\left(1+\frac{\Delta r_i}{r_i}\right).$$

This set of three equations can be solved numerically. (The preceding formulas can be obtained as follows. We derive Eq. (33.24) first by differentiating the log-likelihood function with respect to μ and setting it to zero. Then we differentiate the function with respect to β and set it to zero. The resulting equation can be simplified with the help of Eq. (33.24) to obtain Eqs. (33.25). Finally, we differentiate the function with respect to σ and set it to zero. With the help of Eq. (33.24),

$$\frac{1}{n\Delta t}\sum_{i=1}^{n}\left\{\ln\left(1+\frac{\Delta r_i}{r_i}\right) - \left[\frac{\beta(\mu-r_i)}{r_i} - \frac{1}{2}\sigma^2\right]\Delta t\right\}^2 = \sigma^2.$$

Multiplying out the summand and using Eqs. (33.24) and (33.25) again leads to the last formula.)

Exercise 20.1.5: The log-likelihood function for a sample of size n is

$$-\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\lambda^2) - \theta\sum_{i=1}^{n}\ln S_i - \frac{1}{2\lambda^2}\sum_{i=1}^{n}\left(\frac{\Delta S_i}{S_i}\right)^2 S^{-2\theta}.$$

Straightforward maximization of the function with respect to λ^2 and θ results in

$$\lambda^2 = \frac{1}{n}\sum_i\left(\frac{\Delta S_i}{S_i}\right)^2 S_i^{-2\theta}, \qquad \lambda^2 = \sum_i\frac{\ln S_i}{\sum_i \ln S_i}\left(\frac{\Delta S_i}{S_i}\right)^2 S_i^{-2\theta}.$$

There are two nonlinear equations in two unknowns λ^2 and θ, that can be estimated by the Newton–Raphson method. See [208] and also [113] for a trinomial model for the CEV process.

Exercise 20.1.8: From Exercise 6.4.3, (2), the desired prediction of X_{t+1} given $X_1, X_2, \dots, X_t$ should be $E[\,X_{t+1} \mid X_1, X_2, \dots, X_t\,]$, which equals $E[\,X_{t+1}\,] = \mu$ by stationarity.

Exercise 20.1.9: Let $\boldsymbol{X} \equiv [\,X_t, X_{t-1}, \dots, X_1\,]^{\mathrm{T}}$ and $\boldsymbol{a} \equiv [\,a_1, a_2, \dots, a_t\,]^{\mathrm{T}}$. The linear prediction is then $a_0 + \boldsymbol{a}^{\mathrm{T}}\boldsymbol{X}$. The matrix version of Exercise 6.4.1 says that we should pick $a_0 = \mu$ and $\boldsymbol{a}^{\mathrm{T}} = \mathrm{Cov}[\,X_{t+1}, \boldsymbol{X}\,]\,\mathrm{Cov}[\,\boldsymbol{X}\,]^{-1}$ [416, p. 75]. (See Exercise 19.2.2(2), for the finite-sample version and Exercise 6.4.1, (1), for the one-dimensional version called beta.) By definition $\mathrm{Cov}[\,\boldsymbol{X}\,] = [\,\lambda_{|i-j|}\,]_{1\le i,j\le t}$. As the covariance between X_{t+1} and X_{t-i} is λ_{i+1}, $\mathrm{Cov}[\,X_{t+1}, \boldsymbol{X}\,] = [\,\lambda_1, \lambda_2, \dots, \lambda_t\,]$. See [416, p. 86].

Exercise 20.1.10: Let S_t denote the price at time t and $S_t = S_{t-1} + \epsilon_t$. By the assumption, $\mathrm{Var}[\,S_t\,] = \mathrm{Var}[\,S_{t-1}\,] + \mathrm{Var}[\,\epsilon_t\,]$ based on Eq. (6.5).

Exercise 20.1.12: Take a positive autocorrelation for example. It implies that a higher-than-average (lower-than-average) return today is likely to be followed by higher-than-average (lower-than-average, respectively) returns in the future. In other words, today's returns can be used to predict future returns [424, 767].

Exercise 20.1.13: Note that $E[\,Y_t\,] = \sum_{k=0}^{l} E[\,a_k X_{t-k}\,] = 0$ and

$$\lambda_\tau = E\left[\left(\sum_{j=0}^{l} a_j X_{t+\tau-j}\right)\left(\sum_{k=0}^{l} a_k X_{t-k}\right)\right] = \begin{cases} a_l a_{l-\tau} + \cdots + a_\tau a_0, & \text{if } \tau \le l \\ 0, & \text{otherwise} \end{cases}.$$

See [541, p. 167].

Exercise 20.1.14: $\lambda_\tau \equiv E[\,(X_{t-\tau} - b)(X_t - b)\,]$ equals

$$E\left[\left(\sum_{j=0}^{\infty} c_j \epsilon_{t-j-\tau}\right)\left(\sum_{j=0}^{\infty} c_j \epsilon_{t-j}\right)\right] = \sigma^2(c_0 c_\tau + c_1 c_{\tau+1} + c_2 c_{\tau+2} + \cdots) = \sigma^2 \sum_{j=0}^{\infty} c_j c_{j+\tau}.$$

See [667, p. 68].

Exercise 20.1.15: Similar to Exercise 19.2.2(1), the LS problem is

$$\begin{bmatrix} 1 & X_p & X_{p-1} & \cdots & X_1 \\ 1 & X_{p+1} & X_p & \cdots & X_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n-1} & X_{n-2} & \cdots & X_{n-p} \end{bmatrix} \begin{bmatrix} c \\ a_1 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} X_{p+1} \\ X_{p+2} \\ \vdots \\ X_n \end{bmatrix}.$$

Exercise 20.2.1: (1) Let k_X denote the kurtosis of X_t and k_U the kurtosis of U_t. First, $k_U = E[\,U_t^4\,]$ because $E[\,U_t\,] = 0$ and $E[\,U_t^2\,] = 1$. Now,

$$k_X \equiv \frac{E[\,(X_t - \mu)^4\,]}{E[\,(X_t - \mu)^2\,]^2} = \frac{E[\,V_t^4 U_t^4\,]}{E[\,V_t^2\,]^2\,E[\,U_t^2\,]^2} = \frac{k_U\,E[\,V_t^4\,]}{E[\,V_t^2\,]^2} > k_U. \tag{33.26}$$

The inequality is due to Jensen's inequality. Note that we need to require only that V_t and U_t be independent. See [839, p. 72].

Exercise 20.2.2: Keep $E[\,V_t^n\,] = e^{na + n^2 b^2/2}$ (see p. 68) and $E[\,V_t^i U_t^i\,] = E[\,V_t^i\,]\,E[\,U_t^i\,]$ for any i in mind in the following. (1) By Eq. (33.26) and the fact that the kurtosis of the standard normal distribution is three. (3) $|X_t - \mu| = V_t\,|U_t|$ and the mean of $|U_t|$ is $\sqrt{2/\pi}$. (4) $(X_t - \mu)^2 = V_t^2 U_t^2$.

Exercise 20.2.3: (1) By repeated substitutions. (2) $V_{t+1}^2 = (1 - a_1 - a_2)\,V + a_1(X_t - \mu)^2 + a_2 V_t^2$ given V_t. Take expectations of both sides and rearrange to obtain $E[\,V_{t+1}^2 \mid V_t\,] = V + (a_1 + a_2)(V_t^2 - V)$. Repeat the above steps to arrive at $E[\,V_{t+k}^2 \mid V_t\,] = V + (a_1 + a_2)^k (V_t^2 - V)$. Finally, apply Eq. (20.11), which holds as long as U_t and V_t are uncorrelated. See [470, p. 379].

CHAPTER 21

Exercise 21.1.1: A change of 0.01 translates into a change in the yield on a bank discount basis of one basis point, or 0.0001, which will change the dollar discount, and therefore the invoice price, by

$$0.0001 \times \$1{,}000{,}000 \times \frac{t}{360} = \$0.2778 \times t,$$

where t is the number of days to maturity. As a result, the tick value is $\$0.2778 \times 91 = \$25.28$. See [88, p. 163] or [325, p. 393].

Exercise 21.1.2: The implied rate over the 9-month period is

$$1.01656 \times 1.01694 \times [\,1 + 0.0689 \times (91/360)\,] - 1 = 5.179\%.$$

The annualized rate is therefore $5.179\% \times (360/274) = 6.805\%$.

Exercise 21.1.4: In calculating the conversion factor, the bond is assumed to have exactly 19 years and 9 months to maturity. Therefore there are 40 coupon payments, starting 3 months from now. The value of the bond is

$$\sum_{i=0}^{39} \frac{6.5}{(1.04)^{i+.5}} + \frac{100}{(1.04)^{38.5}} - 3.25 = 149.19.$$

The conversion factor is therefore 1.4919.

Exercise 21.1.5: Buy the September T-note contracts and sell them in August [95, p. 107].

Exercise 21.1.6: (1) The forward rate for the Tth period is $f(T, T+1)$. Now sell one unit of the T-period zero-coupon bond and buy $d(T)/d(T+1)$ units of the $(T+1)$-period zero-coupon bond. The net cash flow today is zero. This portfolio generates a cash outflow of \$1 at time T and a cash inflow of $d(T)/d(T+1)$ dollars at time $T+1$. As the forward rate can be locked in today, it must satisfy $1+f(T,T+1)=d(T)/d(T+1)$ to prevent arbitrage profits. (2) It is the cost to replicate the FRA now.

Exercise 21.2.2: See Exercise 12.2.4 or [346, p. 244].

Exercise 21.4.1: From A's point of view, by issuing a fixed-rate loan and entering into a swap with the bank, it effectively pays a floating rate of (LIBOR $-S'_{\text{A}}-F'_{\text{A}}$)%. It is better off if $S_{\text{A}}>-S'_{\text{A}}-F'_{\text{A}}$. Similarly, from B's point of view, by issuing a floating-rate loan and entering into a swap, it effectively pays a fixed rate of $(F_{\text{B}}+F'_{\text{B}}+S_{\text{B}}+S'_{\text{B}})$%. It is better off if $F_{\text{B}}>F_{\text{B}}+F'_{\text{B}}+S_{\text{B}}+S'_{\text{B}}$. Finally, the bank is better off by entering into both swaps if $F_{\text{B}}+F'_{\text{B}}-S'_{\text{A}}-F_{\text{A}}-F'_{\text{A}}+S'_{\text{B}}>0$. Put these inequalities together and rearrange terms to get the desired result. See [849].

Exercise 21.4.2: The cash flow is identical to that of the fixed-rate payer/floating-rate receiver. To start with, there is no initial cash flow. Furthermore, on all six payment dates, the net position is a cash inflow of LIBOR plus 0.5% and a cash outflow of $5 million.

Exercise 21.4.3: Because the swap value is

$$(0.5-x)\,e^{-0.101\times 0.3}+\left(\frac{0.106962}{2}\times 10-x\right)e^{-0.103\times 0.8}$$

million, the x that makes it zero is 0.51695. The desired fixed rate is 10.339%.

Exercise 21.4.4: To start with,

$$\begin{aligned}P_2-P_1-(21.4) &= (\mathcal{N}+C^*)\,e^{-r_1t_1}-\sum_{i=1}^{n}Ce^{-r_it_i}-\mathcal{N}e^{-r_nt_n}\\&\qquad -(C^*-C)\,e^{-r_1t_1}-\sum_{i=2}^{n}\left(\frac{f_i}{k}\mathcal{N}-C\right)e^{-r_it_i}\\&= \mathcal{N}e^{-r_1t_1}-\mathcal{N}e^{-r_nt_n}-\sum_{i=2}^{n}\frac{f_i}{k}\mathcal{N}e^{-r_it_i}.\end{aligned}$$

So we need to prove only that

$$e^{-r_1t_1}-e^{-r_nt_n}-\sum_{i=2}^{n}\frac{f_i}{k}\,e^{-r_it_i}=0. \tag{33.27}$$

The annualized, continuously compounded forward rate t_{i-1} years from now is

$$c_i\equiv\frac{r_it_i-r_{i-1}t_{i-1}}{t_i-t_{i-1}}=\frac{r_it_i-r_{i-1}t_{i-1}}{(1/k)}$$

from Eq. (5.9). Hence, $f_i=k(e^{c_i/k}-1)=k(e^{r_it_i-r_{i-1}t_{i-1}}-1)$ according to Eq. (3.3), as the desired rate needs to be one that is compounded k times per annum. Now, plug the formula for f_i into Eq. (33.27) to get $e^{-r_1t_1}-e^{-r_nt_n}-\sum_{i=2}^{n}(e^{r_it_i-r_{i-1}t_{i-1}}-1)\,e^{-r_it_i}=0$, as claimed.

Exercise 21.4.5: It is $(C-\widehat{C})\,e^{-r_1t_1}+\sum_{i=2}^{n}(C-\widehat{C})\,e^{-r_it_i}$.

Exercise 21.4.7: Callable bonds are called away when rates decline. This leaves the institution with long interest rate swap positions that have a high fixed rate. Callable swaps can alleviate such a situation. See [821, p. 602].

Exercise 21.4.8: A cap gives the holder an option at each reset date to borrow at a capped rate. An option on a swap, in contrast, offers the holder the one-time option to borrow at a fixed rate over the remaining lifetime of the swap. Clearly, swaptions are less flexible. More rigorously, a swaption can be viewed as an option on a portfolio and a cap as a portfolio of options. Theorem 8.6.1 says that a portfolio of options is more valuable than an option on a portfolio. See [346, p. 248] or [746, p. 513].

Chapter 22

Exercise 22.1.1: See Exercise 5.6.3 and Eq. (5.11).

Exercise 22.2.1: As $d(t)=e^{-ts(t)}$, Eq. (22.1) becomes $e^{-ts(t)}=e^{-ts(t_1)\frac{t_2-t}{t_2-t_1}}\,e^{-ts(t_2)\frac{t-t_1}{t_2-t_1}}$, which can be simplified to $s(t)=s(t_1)\,\frac{t_2-t}{t_2-t_1}+s(t_2)\,\frac{t-t_1}{t_2-t_1}$.

Exercise 22.3.1: Consistent with the numbers in Eq. (22.2), the LS problem is

$$\begin{aligned} P_i &= \sum_{j=1}^{n_i} C_i(a_0 + a_1 j + a_2 j^2) + (a_0 + a_1 n_i + a_2 n_i^2) \\ &= a_0(n_i C_i + 1) + a_1 \left(n_i + C_i \sum_{j=1}^{n_i} j \right) + a_2 \left(n_i^2 + C_i \sum_{j=1}^{n_i} j^2 \right), \quad 1 \le i \le m, \end{aligned}$$

which can be solved by multiple regression.

Exercise 22.3.2: Suppose that $y \approx ae^{bx}$ for each piece of data. Then $\ln y \approx bx + \ln a$. Transform the data into pairs $(x_i, \ln y_i)$ and perform linear regression to obtain a line $z = a' + b'x$. Finally, set $a = e^{a'}$ and $b = b'$ so that $y = e^z = e^{a'+b'x} = ae^{bx}$. See [846, p. 545].

Exercise 22.4.2: Note that

$$f(T) = \frac{\partial[-\ln d(T)]}{\partial T} = -\frac{1}{d(T)} \frac{\partial d(T)}{\partial T}$$

by Eq. (5.7). For the forward rate curve to be continuous, the discount function must have at least one continuous derivative, hence claim (1). The preceding identities also show that if the forward rate curve should be continuously differentiable, then a cubic spline is needed. See [147, p. 411].

Chapter 23

Exercise 23.1.1: No [492, p. 388].

Exercise 23.1.2: The accumulated value of a \$1 investment is $a(n) \equiv \prod_{t=1}^{n}(1 + i_t)$. As $\ln(1 + i_t) \sim N(\mu, \sigma^2)$, the mean and the variance of this lognormal variable are given by $e^{\mu+\sigma^2/2}$ and $e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$, respectively, according to Eq. (6.11). Now $\ln a(n) = \sum_{t=1}^{n} \ln(1 + i_t)$ is normal with mean $n\mu$ and variance $n\sigma^2$. Equivalently, $a(n)$ is lognormal with mean $e^{n\mu+n\sigma^2/2}$ and variance $e^{2n\mu+n\sigma^2}(e^{n\sigma^2} - 1)$. See [547, p. 368].

Exercise 23.2.1: The logarithms of the short rates are $\ln r_k, \ln r_k + \ln v_k, \ln r_k + 2\ln v_k, \ldots, r_k + (k-1)\ln v_k$. As far as variance is concerned, the term $\ln r_k$ is irrelevant; hence it is deleted from the list. The mean of the logarithms of the short rates is

$$\sum_{i=0}^{k-1} (i \ln v_k) \frac{\binom{k-1}{i}}{2^{k-1}} = \frac{k-1}{2} \ln v_k.$$

Hence the variance equals

$$\begin{aligned} \sum_{i=0}^{k-1} (i \ln v_k)^2 \frac{\binom{k-1}{i}}{2^{k-1}} - \left(\frac{k-1}{2} \ln v_k \right)^2 &= \frac{(\ln v_k)^2}{2^{k-1}} \sum_{i=0}^{k-1} i^2 \binom{k-1}{i} - \left(\frac{k-1}{2} \ln v_k \right)^2 \\ &= \frac{(\ln v_k)^2}{2^{k-1}} (k-1)[2^{k-2} + (k-2)\,2^{k-3}] - \left(\frac{k-1}{2} \ln v_k \right)^2 \\ &= (k-1) \left(\frac{\ln v_k}{2} \right)^2, \end{aligned}$$

which equals $\sigma_k^2(k-1)\Delta t$ based on Eq. (23.2). See [731, p. 440].

Exercise 23.2.2: Consider rate $r \equiv r_j v_j^i$ for period j and its two subsequent rates $r_{j+1} v_{j+1}^i$ and $r_{j+1} v_{j+1}^{i+1}$. By Eq. (23.2) and the constant-volatility assumption, we have $v_j = v_{j+1} = e^{2\sigma\sqrt{\Delta t}}$. Hence the dynamics becomes $r \to r(r_{j+1}/r_j)\, e^{2\sigma\sqrt{\Delta t}\,\xi}$, where $\xi = 0, 1$, each with a probability of one-half. We establish the claim by defining μ such that $r_{j+1}/r_j = e^{\mu-\sigma\sqrt{\Delta t}}$.

Exercise 23.2.3: (1) It is equal to

$$q(\ln r_\ell)^2 + (1-q)(\ln r_{\mathrm{h}})^2 - [\, q \ln r_\ell + (1-q) \ln r_{\mathrm{h}} \,]^2 = q(1-q)(\ln r_{\mathrm{h}} - \ln r_\ell)^2.$$

(2) Use $r(\Delta t)$ to denote the r after a time period of Δt. From (1),

$$\sigma^2 \Delta t = \mathrm{Var}[\, \ln r(\Delta t) \,] = \mathrm{Var}[\, \ln(r + \Delta r) \,] = \mathrm{Var}\left[\ln \left(1 + \frac{\Delta r}{r} \right) \right] \approx \mathrm{Var}\left[\frac{\Delta r}{r} \right] = \frac{\mathrm{Var}[\, \Delta r \,]}{r^2}.$$

So $\mathrm{Var}[\, r + \Delta r \,] = \mathrm{Var}[\, \Delta r \,] = r^2 \sigma^2 \Delta t$.

Exercise 23.2.4: (1) Let the rate for the first period be r and the forward rate one period from now be f. Suppose the binomial interest rate tree gives r_ℓ and r_h for the forward rates applicable in the second period. By construction, $r_h/r_\ell = v$ and $(r_h + r_\ell)/2 = f$. The price of a zero-coupon bond two periods from now is priced by the tree as

$$\frac{1}{1+r}\,\frac{1}{2}\left(\frac{1}{1+r_h} + \frac{1}{1+r_\ell}\right).$$

It is not difficult to see that the above number exceeds $1/[\,(1+r)(1+f)\,]$ unless $v = 1$.

(2) By (1) we know that the claim holds for trees with two periods. Assume that the claim holds for trees with $n-1$ periods and proceed to prove its validity for trees with n periods. Suppose the tree has baseline rates $r_1, r_2, \ldots, r_n$, where r_i is the baseline rate for the ith period. Denote this tree by $T(r_1, \ldots, r_n)$. Split the tree into two $(n-1)$-period subtrees by taking out the root. The first tree can be denoted by $T(r_2, \ldots, r_n)$ and the second by $T(r_2 v_2, \ldots, r_n v_n)$. If $V(T)$ is used to signify the value of a security as evaluated by the binomial interest rate tree T, clearly

$$V(T(r_1, \ldots, r_n)) = \frac{V(T(r_2, \ldots, r_n)) + V(T(r_2 v_2, \ldots, r_n v_n))}{2(1+r_1)}. \tag{33.28}$$

It suffices to prove the claim for zero-coupon bonds by the additivity of the valuation process. Hence the problem is reduced to proving

$$V(T(r_1, \ldots, r_n)) > [\,(1+f_1)(1+f_2)\cdots(1+f_n)\,]^{-1},$$

where f_i is the one-period forward rate for period i.

By Eq. (23.4), $f_j = r_j(1+v_j/2)^{j-1}$. So $T(r_2, \ldots, r_n)$ implies that its ith-period forward rate (counting from its root) is

$$f_i' = r_{i+1}\left(\frac{1+v_{i+1}}{2}\right)^{i-1} = f_{i+1}\,\frac{2}{1+v_{i+1}}.$$

Similarly, $T(r_2 v_2, \ldots, r_n v_n)$ implies that its ith-period forward rate is

$$f_i'' = r_{i+1} v_{i+1}\left(\frac{1+v_{i+1}}{2}\right)^{i-1} = f_{i+1}\,\frac{2v_{i+1}}{1+v_{i+1}}.$$

Apply the induction hypothesis to each of the subtrees to obtain

$$V(T(r_2, \ldots, r_n)) > [\,(1+f_1')(1+f_2')\cdots(1+f_{n-1}')\,]^{-1}$$
$$V(T(r_2 v_2, \ldots, r_n v_n)) > [\,(1+f_1'')(1+f_2'')\cdots(1+f_{n-1}'')\,]^{-1}$$

Add them up:

$$\begin{aligned}&V(T(r_2, \ldots, r_n)) + V(T(r_2 v_2, \ldots, r_n v_n))\\ &\quad > [\,(1+f_1')(1+f_2')\cdots(1+f_{n-1}')\,]^{-1} + [\,(1+f_1'')(1+f_2'')\cdots(1+f_{n-1}'')\,]^{-1}.\end{aligned}$$

By Eq. (33.28) we are done if we can show that

$$[\,(1+f_1')(1+f_2')\cdots(1+f_{n-1}')\,]^{-1} + [\,(1+f_1'')(1+f_2'')\cdots(1+f_{n-1}'')\,]^{-1} > 2[(1+f_2)\cdots(1+f_n)]^{-1}.$$

Recall that $r_1 = f_1$ by definition. Now,

$$\begin{aligned}&[\,(1+f_1')(1+f_2')\cdots(1+f_{n-1}')\,]^{-1} + [\,(1+f_1'')(1+f_2'')\cdots(1+f_{n-1}'')\,]^{-1}\\ &\quad \geq 2\sqrt{[\,(1+f_1')(1+f_2')\cdots(1+f_{n-1}')\,]^{-1}[\,(1+f_1'')(1+f_2'')\cdots(1+f_{n-1}'')\,]^{-1}}\end{aligned}$$

because of the relation between arithmetic and geometric means. It is sufficient to prove that

$$(1+f_1')(1+f_2')\cdots(1+f_{n-1}')(1+f_1'')(1+f_2'')\cdots(1+f_{n-1}'') < [\,(1+f_2)\cdots(1+f_n)\,]^2.$$

We now prove the validity of the preceding inequality by showing that $(1+f_i')(1+f_i'') < (1+f_{i+1})^2$ for each $1 \leq i \leq n-1$.

Because

$$(1+f_i')(1+f_i'') = \left(1+f_{i+1}\,\frac{2}{1+v_{i+1}}\right)\left(1+f_{i+1}\,\frac{2v_{i+1}}{1+v_{i+1}}\right) = \frac{4v_{i+1}}{(1+v_{i+1})^2}\,f_{i+1}^2 + 2f_{i+1} + 1,$$

we have

$$(1+f_i')(1+f_i'') - (1+f_{i+1})^2 = -\left(\frac{1-v_{i+1}}{1+v_{i+1}}\right)^2 f_{i+1}^2 < 0,$$

as desired.

Exercise 23.2.5: 23.2.5: Fix a period $[k-1, k]$. (1) Let f be the forward rate for that period. Initiate an FRA at time zero for that period. Its payoff at time k equals $s - f$, where s is the future spot rate for the period. The forward rate f is the fixed contract rate that makes the FRA zero valued. Adopt the forward-neutral probability measure π_k with the zero-coupon bond maturing at time k as numeraire. By Exercise 13.2.13, $0 = d(k)\, E_0^{\pi_k}[\, s - f\,]$, which implies that $f = E_0^{\pi_k}[\, s\,]$. (2) Let f' denote the same forward rate one period from now. Again, $f' = E_1^{\pi_k}[\, s\,]$. By applying $E_0^{\pi_k}$ to both sides of the equation and then Eq. (6.6), the law of iterated conditional expectations, we obtain $E_0^{\pi_k}[\, f'\,] = E_0^{\pi_k}[\, E_1^{\pi_k}[\, s\,]\,] = E_0^{\pi_k}[\, s\,]$, which equals f. See [691, p. 398] and [731, p. 177].

Exercise 23.2.6: Solve for

$$\frac{0.112832}{1+r} + \frac{0.333501}{1+1.5\times r} + \frac{0.327842}{1+(1.5)^2\times r} + \frac{0.107173}{1+(1.5)^3\times r} = \frac{1}{(1+0.044)^4}.$$

The result is $r = 0.024329$. The forward rate for the fourth period is $(1+0.044)^4/(1+0.043)^3 - 1 = 1.047006$. So the baseline rate would have been $2^3\times 0.047006/(2.5)^3 = 0.024067$ had we used Eq. (23.4). It is lower than 0.024329.

Exercise 23.2.7: (1) Because $d(j) = \sum_i P_i$, where P_i is the state price in state i at time j, simply define the risk-neutral probabilities as $q_i \equiv P_i/d(j)$. Clearly q_i sum to one. These probabilities are called forward-neutral probabilities in Exercise 13.2.13. (2) They are 0.232197/0.92101, 0.460505/0.92101, and 0.228308/0.92101. See [40].

Exercise 23.2.8: Calculate all the state prices in $O(n^2)$ time by using forward induction. Then sum the state prices of each column.

Exercise 23.2.9: (1) We do it inductively. Suppose we are at time j and there are $j+1$ nodes with the state prices $P'_1, P'_2, \ldots, P'_{j+1}$. Let the baseline rate for period j be $r \equiv r_j$, the multiplicative ratio be $v \equiv v_j$, and $P_1, P_2, \ldots, P_j$ be the state prices a period prior, corresponding to rates $r, rv, \ldots, rv^{j-1}$. Each P_i has branching probabilities p_i for the up move and $1-p_i$ for the down move. Add $P_0 \equiv 0$ for convenience. Clearly, $P'_i = (1-p_{i-1})P_{i-1}/(1+rv^{i-2}) + p_i P_i/(1+rv^{i-1})$, $i = 1, 2, \ldots, j+1$ (see Fig. 23.7). We are done because there are $j+1$ equations (one of which is redundant) and j unknowns $p_1, p_2, \ldots, p_j$. (2) Although the j state prices $P_1, P_2, \ldots, P_j$ are now unknown, $p_1 = p_2 = \cdots = p_j$ from the assumption. Hence $j+1$ unknowns remain. We use the $j+1$ equations, none of which is redundant, to solve for the unknowns.

Exercise 23.3.1: No [427].

Exercise 23.3.4: (1) It is identical to the calculation for the futures price except that (a) discounting should be used during backward induction and (b) the final price, the PV of a bond delivered at the delivery date T, should be divided by the discount factor $d(T)$ to obtain its future value [623, p. 390]. (2) The forward price is

$$\frac{97.186\times 0.232197 + 95.838\times 0.460505 + 93.884\times 0.228308}{0.92101} = 95.693.$$

Note that the forward price exceeds the futures price 95.687 as Exercise 12.3.3 predicts.

Exercise 23.3.6: Consider the strategy of buying a coupon bond, selling a call on this bond struck at X, and buying a put on this bond also struck at X. The options have the same expiration date t. The cash flow is $C, C, \ldots, C, X$, where C is the periodic coupon payment of the bond. Note that there is no cash flow beyond time t. Hence $\text{PV}(I) + \text{PV}(X) = B + P - C$, where $\text{PV}(I)$ denotes the present value of the bond's cash flow on and before time t. See [731, p. 269].

Exercise 23.4.1: At any node on the correlated binomial tree with the stock price–short rate pair (S, r), it must hold that $d < 1 + r < u$ by Exercise 9.2.1. Unless there are prior restrictions on r, this implies that u and d must be state, time, and path dependent. See [779].

Chapter 24

Exercise 24.2.1: (1) If the forward price satisfies it, the following two strategies yield the same dollar amount after M time for the same cost: (1) Buying an M-time zero-coupon bond for $P(t, M)$ dollars and (2) spending $P(t, M)$ dollars to buy $P(t, M)/P(t, T)$ units of the T-time zero-coupon bond and entering into a T-time forward contract for the $(M-T)$-time zero-coupon bond. Observe that we have $F(t, T, M)$ dollars at the end of time T in strategy (2), exactly what is needed to take delivery of one $(M-T)$-time zero-coupon bond. If $P(t, M) > P(t, T)\, F(t, T, M)$, arbitrage opportunities exist by taking short positions in the M-time zero-coupon bonds and long positions in both the T-time zero-coupon bonds and the T-time forward contracts for $(M-T)$-time zero-coupon bonds. Reverse the positions if $P(t, M) < P(t, T)\, F(t, T, M)$. Note the affinity between Eq. (24.1) and Eq. (5.2).

(2) Suppose the cash flows before the forward contract's delivery date are $C_1, C_2, \ldots,$ at times $T_1, T_2, \ldots$. Let F denote the forward price. Our replicating strategy is this: Borrow $FP(t, T)$ dollars for T years, borrow $C_i P(t, T_i)$ dollars for T_i years for $i = 1, 2, \ldots$, and buy one unit of the underlying bond for $B(t, M)$ dollars. Each of the loans' obligations before T can be paid off by the bond's coupon. On the delivery date T, the position is worth the bond's FV minus the forward price, $B(T, M) - F$, replicating the forward contract's payoff. Because the forward contract has zero value at time t, we must have $FP(t, T) + \sum_i C_i P(t, T_i) - B(t, M) = 0$. Thus $F = \frac{B(t,M) - \sum_i C_i P(t,T_i)}{P(t,T)}$. In words, the forward price equals the FV of the underlying bond's current invoice price minus the PV of the bond's cash flow until the delivery date. The forward price includes the accrued interest of the underlying bond payable at time T. See [848, p. 169].

Exercise 24.2.2: There is no net investment at time t, and we pay \$1 at time T and receive $P(t, T)/P(t, M)$ dollars at time M. In other words, the investment at time T results in a certain gross return of $P(t, T)/P(t, M)$ at time M. So the said forward price must be $P(t, M)/P(t, T)$. See [76, p. 22].

Exercise 24.2.3: Note that Eq. (5.11) is equivalent to $f(T) = \partial(TS(T))/\partial T$ and $TS(T) = -\ln d(T)$ by Eq. (5.7). The rest is simple translation between the two systems of notations.

Exercise 24.2.4: Equation (24.2) is

$$1 + f(t, s, 1) = 1 + f(t, s) = \frac{P(t, s)}{P(t, s+1)}.$$

With the preceding identity applied recursively, we have

$$P(t, T) = \frac{P(t, t+1)}{[\,1 + f(t, t+1)\,] \cdots [\, f(t, T-1)\,]}$$

Now,

$$P(t, t+1) = [\,1 + r(t, t+1)\,]^{-1} = [\,1 + r(t)\,]^{-1}.$$

Combine the preceding two equations above to obtain the desired result. See [492, p. 388].

Exercise 24.2.5: We know that the spot rate s satisfies $P = e^{-s\tau}$. On the other hand, Eq. (24.6)'s continuous compounding analog, Eq. (5.8), says that $P = e^{-\sum_{i=0}^{\tau-1} f(t, t+i)}$. See [746, p. 422].

Exercise 24.2.6: Straightforward from Eqs. (5.12) and (5.13). An alternative starts from Eq. (24.1) with $F(t, T, T+L) = \frac{P(t, T+L)}{P(t,T)}$. Then observe that

$$f(t, T, L) = \frac{1}{L}\left[\frac{1}{F(t, T, T+L)} - 1\right] = \frac{1}{L}\left[\frac{P(t, T)}{P(t, T+L)} - 1\right].$$

Exercise 24.2.8: Apply Eqs. (14.16) and (24.4) to obtain $f(t, T) = r(t) + \mu(T-t) - \sigma^2(T-t)^2/2$. From the definition, $E_t[\,r(T) \mid r(t)\,] = r(t) + \mu(T-t)$. So the premium is

$$f(t, T) - E_t[\,r(T) \mid r(t)\,] = -\frac{\sigma^2(T-t)^2}{2} < 0.$$

As for the forward rate, use Eqs. (14.16) and (24.4) to obtain $r + \mu(T-t) - \sigma^2(T-t)^2/2$.

Exercise 24.2.9: Exercise 5.6.6 and Eq. (24.5) imply that $P(t, T) = e^{-\int_t^T r(s)\, ds}$ for a certain economy [746, p. 527].

Exercise 24.2.10: From $r_c \equiv \ln(1 + r_e)$, we have $r_e/(1 + r_e) = 1 - e^{-r_c}$. Hence,

$$\begin{aligned} dr_c &= (1 + r_e)^{-1}\, dr_e - \frac{1}{2}\,(1 + r_e)^{-2}\,(dr_e)^2 \\ &= (1 + r_e)^{-1} r_e(\mu\, dt + \sigma\, dW) - \frac{1}{2}\,(1 + r_e)^{-2} r_e^2 \sigma^2\, dt \\ &= (1 - e^{-r_c})(\mu\, dt + \sigma\, dW) - \frac{1}{2}(1 - e^{-r_c})^{-2} \sigma^2\, dt, \end{aligned}$$

from which the equation follows. See [781, Theorem 3].

Exercise 24.3.1: From $P(t,T) = e^{-r(t,T)(T-t)}$ and Eq. (24.9), $r(t,T) = -\frac{\ln E_t[e^{-\int_t^T r(s)\,ds}]}{T-t}$. Jensen's inequality says it is less than $\frac{E_t[\int_t^T r(s)\,ds]}{T-t}$ unless there is no uncertainty. See [302, Theorem 1].

Exercise 24.3.2: From Eqs. (24.4) and (24.9) and Leibniz's rule,

$$f(t,T) = -\frac{\partial P(t,T)/\partial T}{P(t,T)} = \frac{E_t[\,r(T)\,e^{-\int_t^T r(s)\,ds}\,]}{E_t[\,e^{-\int_t^T r(s)\,ds}\,]}.$$

So the forward rate is a weighted average of future spot rates. Finally, the assumption says that the above average is less than the simple average $E_t[\,r(T)\,]$. See [302, Theorem 2].

Exercise 24.3.3: Yes for the calibrated tree and independent of the term structure of volatilities. Consider an n-period bond with $n > 1$. One period from now (time t), the bond will have two prices, P_u and P_d, such that $(P_u + P_d)/\{2[1 + r(t)]\} = P(t,n)$, the market price of n-period zero-coupon bonds. The expected one-period return for this bond is, by construction,

$$\frac{(P_u + P_d)/2}{P(t,n)} = 1 + r(t),$$

exactly as demanded by the local expectations theory. As for the uncalibrated tree, it does not satisfy the theory. The reason is simple:

$$\frac{P_u + P_d}{2(1 + r(t))} > P(t,n),$$

as the tree is known to overestimate the discount factor by Theorem 23.2.2.

Exercise 24.3.4: From Eqs. (24.5) and (24.6) and the theory, which says that $f(s,t) = E_s[\,r(t)\,]$.

Exercise 24.3.5: Every cash flow has to be discounted by the appropriate discount function [292].

Exercise 24.3.7: See [725, p. 231].

Exercise 24.3.8: An interest rate cap, we recall, is a contract in which the seller promises to pay a certain amount of cash to the holder if the interest rate exceeds a certain predetermined level (the cap rate) at certain future dates. In the same way, the seller of a floor contract promises to pay cash when future interest rates fall below a certain level. Technically, a cap contract is a sum of caplets. We now give a precise description of the caplet.

Let t stand for the time at which the contract is written and $[\,t_0, t_1\,]$ be the period for which the caplet is in effect with $\Delta t \equiv t_1 - t_0$. Denote the cap rate by x. For simplicity, assume that the notional principal is \$1.

The interest rate that in real life determines the payments of the cap is some market rate such as LIBOR. The rate is quoted as a simple rate over the period $[\,t_0, t_1\,]$. This simple rate, which we denote by f, is determined at t_0 and defined by the relation $P(t_0,t_1) = 1/(1 + f\Delta t)$. Finally a caplet is a contingent t_1 claim that at time t_1 will pay $\max(f - x, 0) \times \Delta t$ to the holder of the contract. The payment is in arrears. Specifically, the payoff at time t_1 is

$$\begin{aligned}\max\left(\frac{1 - P(t_0,t_1)}{P(t_0,t_1)\,\Delta t} - x, 0\right) \times \Delta t &= \max\left(\frac{1}{P(t_0,t_1)} - (1 + x\Delta t), 0\right)\\ &= \max\left(\frac{1}{P(t_0,t_1)} - \alpha, 0\right),\end{aligned}$$

where $\alpha \equiv 1 + x\Delta t$. The price of the caplet at time t can be easily proved to be

$$E_t^{\pi}\left[e^{-\int_t^{t_1} r(s)\,ds} \max\left(\frac{1}{P(t_0,t_1)} - \alpha, 0\right)\right] = \alpha\, E_t^{\pi}\left[e^{-\int_t^{t_0} r(s)\,ds} \max\left(\frac{1}{\alpha} - P(t_0,t_1), 0\right)\right].$$

Thus a caplet is equivalent to α put options on a t_1 bond with delivery date t_0 and strike price $1/\alpha$ (see also Subsection 21.2.4).

Exercise 24.4.1: From Eq. (24.13) we have $\Theta_1 - \Theta_2 = \sigma(r,t)^2(C_2 - C_1)/2$ [207].

Exercise 24.4.3: Rearrange the equation as

$$\frac{1}{2}\sigma(r)^2\frac{\partial^2 P}{\partial r^2}+\mu(r)\frac{\partial P}{\partial r}-rP-\frac{\partial P}{\partial T}=0$$

and assume that μ and σ are independent of time for simplicity. The partial differential equation becomes the following $N-1$ difference equations:

$$\frac{1}{2}\sigma_i^2\frac{P_{i+1,j}-2P_{i,j}+P_{i-1,j}}{(\Delta r)^2}+\mu_i\frac{P_{i+1,j}-P_{i,j}}{\Delta r}-rP_{i,j}-\frac{P_{i,j}-P_{i,j-1}}{\Delta t}=0$$

for $1\le i\le N-1$, where $\mu_i\equiv\mu(i\Delta r)$ and $\sigma_i\equiv\sigma(i\Delta r)$. (In the preceding equation, we could have used $(P_{i+1,j}-P_{i-1,j})/(2\Delta r)$ for $\partial P/\partial r$.) Regroup the terms by $P_{i,j-1}$, $P_{i-1,j}$, $P_{i,j}$, and $P_{i+1,j}$ to obtain the following system of equations at every time step:

$$a_iP_{i-1,j}+b_iP_{i,j}+c_iP_{i+1,j}=P_{i,j-1},$$

where

$$a_i\equiv-\left(\frac{\sigma_i}{\Delta r}\right)^2\frac{\Delta t}{2},\qquad b_i\equiv 1+i\Delta r\,\Delta t+\left(\frac{\sigma_i}{\Delta r}\right)^2\Delta t+\frac{\mu_i\Delta t}{\Delta r},\qquad c_i\equiv-\left(\frac{\sigma_i}{\Delta r}\right)^2\frac{\Delta t}{2}-\frac{\mu_i\Delta t}{\Delta r}.$$

Initially, the terminal conditions $P_{i,0}$ are given. The condition $\sigma(0)=0$ implies that $a_0=0$. The other condition, $\lim_{r\to\infty}P(r,T)=0$, leads to $P_{N,j}=0$. The system of equations can be written as

$$\begin{bmatrix} b_0 & c_0 & 0 & \cdots & \cdots & \cdots & 0\\ a_1 & b_1 & c_1 & 0 & \cdots & \cdots & 0\\ 0 & a_2 & b_2 & c_2 & 0 & \cdots & 0\\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots\\ 0 & \cdots & \cdots & 0 & a_{N-2} & b_{N-2} & c_{N-2}\\ 0 & \cdots & \cdots & \cdots & 0 & a_{N-1} & b_{N-1}\end{bmatrix}\begin{bmatrix}P_{0,j}\\ P_{1,j}\\ P_{2,j}\\ \vdots\\ P_{N-1,j}\end{bmatrix}=\begin{bmatrix}P_{0,j-1}\\ P_{1,j-1}\\ P_{2,j-1}\\ \vdots\\ \vdots\\ P_{N-2,j-1}\\ P_{N-1,j-1}\end{bmatrix}.$$

We can obtain the $P_{i,j}$s by inverting the tridiagonal matrix given the $P_{i,j-1}$s. See Section 18.1 or [38, pp. 81–84]. Note that the time partial derivative of the Black–Scholes differential equation in Section 18.1 is with respect to t, not T here. This accounts for the slight notational differences from the implicit equations in Subsection 18.1.3.

In practice, a finite interval in the r axis such as $[\,0\%, 30\%\,]$ suffices to obtain good approximations. A better approach is to consider a transform like $x\equiv 1/(1+r)$ and solve the differential equation in terms of x, whose domain $(\,0,1\,]$ is finite [38]. See also [859].

Exercise 24.4.4: The argument is the same as that leading to the term structure equation except that the futures contract's return is zero, not r [746, p. 565].

Exercise 24.5.1: In a risk-neutral probability measure, $\mu_p(t,T)=r(t)$ and

$$d\ln P(t,T)=\left[r(t)-\frac{1}{2}\sigma_p(t,T,P(t,T))^2\right]dt+\sigma_p(t,T,P(t,T))\,dW_t.$$

Recall from Eq. (24.3) that

$$f(t,T,\Delta t)=\frac{\ln P(t,T)-\ln P(t,T+\Delta t)}{\Delta t}.$$

Hence $df(t,T,\Delta t)$ equals

$$\frac{\sigma_p(t,T+\Delta t,P(t,T+\Delta t))^2-\sigma_p(t,T,P(t,T))^2}{2\Delta t}\,dt$$
$$+\frac{\sigma_p(t,T,P(t,T))-\sigma_p(t,T+\Delta t,P(t,T+\Delta t))}{\Delta t}\,dW_t.$$

As $\Delta t\to 0$ the preceding equation becomes

$$d(f,T)=\sigma_p(t,T,P(t,T))\frac{\partial\sigma_p(t,T,P(t,T))}{\partial T}\,dt-\frac{\partial\sigma_p(t,T,P(t,T))}{\partial T}\,dW_t.$$

See [692, Subsection 19.3.2].

Exercise 24.5.2: Let $\sigma_p(t, T, P) = \psi(t) \ln P(t, T)$. The desired diffusion term, $-\partial\sigma_p(t, T, P)/\partial T$, equals $-\psi(t)(\partial \ln P(t, T)/\partial T) = \psi(t)\, f(t, T)$ because $f(t, T) = -\partial \ln P(t, T)/\partial T$. See [731, p. 378].

Exercise 24.6.2: In a period, bond one's price can go to $P_1 u_1$ or $P_1 d_1$, and bond two's price can go to $P_2 u_2$ or $P_2 d_2$. A portfolio consisting of one short unit of bond one and $\frac{P_1 u_1 - P_1 d_1}{P_2 u_2 - P_2 d_2}$ long units of bond two is riskless because its value equals $\frac{P_1(u_1 d_2 - u_2 d_1)}{u_2 - d_2}$ in either state. Hence

$$\left[-B_1 + \frac{P_1(u_1 - d_1)}{P_2(u_2 - d_2)}\, B_2\right](1 + r) = \frac{P_1(u_1 d_2 - u_2 d_1)}{u_2 - d_2},$$

in which B_i is bond i's current price. Use the preceding equation and Eqs. (24.16) and (24.17) in the formula for λ to make sure that the final result is indeed independent of the bond. See [38, §3.3] and [781, Footnote 3].

Exercise 24.6.3: (1) In a period the portfolio's value is either $P_u + V_u(P_d - P_u)/(V_u - V_d)$ or $P_d + V_d(P_d - P_u)/(V_u - V_d)$. Because both equal $\frac{-P_u V_d + P_d V_u}{V_u - V_d}$, the return must be riskless. (2) Solve $(1 + x\, \frac{P_d - P_u}{V_u - V_d}) R = \frac{-P_u V_d + P_d V_u}{V_u - V_d}$ for x. See [739].

Exercise 24.6.4: (1) To match the payoff one period from now, we need

$$\begin{aligned} u P(t, t+2)\, \Delta + e^r B &= V_u, \\ d P(t, t+2)\, \Delta + e^r B &= V_d. \end{aligned}$$

It is easy to deduce from the preceding equations that the values are correct as listed. (2) Plug the values in (1) into $V = \Delta \times P(t, t+2) + B$. See [441].

Exercise 24.6.5: Simple algebraic manipulations [38, p. 44].

Exercise 24.6.6: Without the current term structure, the step $100/(1.05)^2 = 90.703$ would not have been taken, and everything that followed would have broken down. See [725, Example 6.1] for more information.

Exercise 24.6.7: $\frac{(90.703 \times 1.04) - 92.593}{98.039 - 92.593}$.

Exercise 24.6.8: When a security's value is matched in every state by the bond portfolio which is set up dynamically, it will be immunized. This notion admits of no arbitrage profits.

Exercise 24.7.1: (1) Because the bond can never be worth more than \$100 if interest rates are nonnegative. (2) This is due to the lognormal assumption for bond prices and the result that there is some probability that the bond will reach *any* given positive price. However, in fact, bond prices must lie between zero and the sum of the remaining cash flows if interest rates are nonnegative. See [304].

Exercise 24.7.2: No. The value of this option equals the maximum of zero and the discounted value of par minus the strike price, i.e., $\max(0, P(t, T)(100 - X))$, where T is the expiration date. See [304].

Chapter 25

Exercise 25.1.1: Rearrange the AR(1) process $X_t - b = a(X_{t-1} - b) + \epsilon_t$ to yield $X_t - X_{t-1} = (1 - a)(b - X_{t-1}) + \epsilon_t$, which is the discrete-time analog of the Vasicek model [41, p. 94]. Recall that the autocorrelation of X at lag one is a, which translates into $1 - \beta\, dt$ in the Vasicek model.

Exercise 25.1.2: (2) Observe that $\partial B(t, T)/\partial T = e^{-\beta(T-t)}$. From Eq. (24.4) we have

$$f(t, T) = -\frac{\sigma^2 \left(e^{-\beta(T-t)} - 1\right)^2}{2\beta^2} - \mu(e^{-\beta(T-t)} - 1) + r(t)\, e^{-\beta(T-t)}.$$

As $E_t[\, r(T) \mid r(t)\,] = \mu + (r(t) - \mu)\, e^{-\beta(T-t)}$, the liquidity premium is $-\sigma^2[\, e^{-\beta(T-t)} - 1\,]^2/(2\beta^2)$. Note that this premium is zero for $T = t$, as it should be. It is negative otherwise and converges to $-\sigma^2/(2\beta^2)$ as $T \to \infty$, which is the difference between the long rate and the long-term mean of the short rate, μ. See [855].

Exercise 25.1.3: Just plug in the following equations into Eq. (24.12):

$$\begin{aligned}
\lambda(t, r) &= 0,\\
\mu(r, t) &= \beta(\mu - r),\\
\sigma(r, t) &= \sigma,\\
\frac{\partial P/\partial T}{P} &= -re^{-\beta(T-t)} + \mu e^{-\beta(T-t)} - \frac{\sigma^2}{\beta^2}\, e^{-\beta(T-t)} - \mu + \frac{\sigma^2}{2\beta^2} + \frac{\sigma^2}{2\beta^2}\, e^{-2\beta(T-t)},\\
\frac{\partial P/\partial r}{P} &= -B(t, T) = \frac{e^{-\beta(T-t)} - 1}{\beta},\\
\frac{\partial^2 P/\partial r^2}{P} &= B(t, T)^2 = \left[\frac{1 - e^{-\beta(T-t)}}{\beta}\right]^2.
\end{aligned}$$

The $\beta = 0$ case is simpler.

Exercise 25.1.4: Obtain the formula for $\partial P/\partial r$ from Exercise 25.1.3.

Exercise 25.1.5: From Exercise 25.1.4,

$$d \ln P = \left[r - \frac{1}{2}\, B(t, T)^2 \sigma^2\right] dt - B(t, T)\, \sigma\, dW.$$

Because $f(t, T) = -\partial \ln P(t, T)/\partial T$,

$$df = -\left[-B(t, T)\, \sigma^2\, \frac{\partial B(t, T)}{\partial T}\right] dt + \sigma\, \frac{\partial B(t, T)}{\partial T}\, dW.$$

We can simplify the preceding equation by using $\partial B(t, T)/\partial T = e^{-\beta(T-t)}$ to

$$df = \left[\frac{1 - e^{-\beta(T-t)}}{\beta}\, e^{-\beta(T-t)} \sigma^2\right] dt + e^{-\beta(T-t)} \sigma\, dW.$$

See [16].

Exercise 25.1.6: Use $\ln P(t, T) = \ln A(t, T) - B(t, T)\, r(t)$ and Eq. (14.14) to prove that

$$\mathrm{Var}[\,\ln P\,] = B(t, T)^2\, \mathrm{Var}[\,r(t)\,] = \frac{\sigma^2}{2\beta}\left[1 - e^{-2\beta(T-t)}\right].$$

Exercise 25.2.1: It approaches $2\beta\mu/(\beta + \gamma)$ by letting $T \to \infty$.

Exercise 25.2.4: The spot rate curve is $r(t, T) = -\frac{\ln P(t,T)}{T-t} = -\frac{\ln A(t,T)}{T-t} + \frac{B(t,T)\, r(t,T)}{T-t}$. So the spot rate volatility structure becomes $\frac{\partial r(t,T)}{\partial r}\, \sigma(r, t) = \sigma(r, t)\, B(t, T)/(T - t)$. See [731, p. 415].

Exercise 25.2.5: (2) It is $\sqrt{2(\phi_1\phi_2 - \phi_2^2)}$ [138].

Exercise 25.2.6: Equations (25.1) and (25.3) say that the spot rate is a linear function of the short rate under these two models, say $r(t, T) = a(T - t) + b(T - t)\, r(t)$. In particular, $r(T, T_1) = a' + b' r(T, T_2)$ for some a' and b'. Now

$$\begin{aligned}
\max(0, r(T, T_1) - r(T, T_2)) &= \max(0, a' + (b' - 1)\, r(T, T_2))\\
&= (b' - 1) \times \max\left(0, r(T, T_2) - \frac{a'}{1 - b'}\right),
\end{aligned}$$

which means a portfolio of caplets. The case of floorlets is symmetric. See [616].

Exercise 25.2.9: Substitute the definitions of r^+, r^-, and $r = f(x)$ into Eq. (25.4) [268].

Exercise 25.2.11: See [868].

Exercise 25.2.12: Observe that all the rates on the same horizontal row are identical. Nodes with identical short rates generate identical term structures because the term structure depends solely on the prevailing short rate. As a result, we need to store only a vector of rates and a vector of probabilities. To answer (1), for example, we slide the two vectors backward in time as we perform the necessary computation on a third vector. See [405].

Exercise 25.2.13: Let $x \equiv f(y, t)$. By Ito's lemma,

$$dx = \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} \alpha(y,t) + \frac{1}{2} \frac{\partial^2 f}{\partial y^2} \sigma(y,t)^2 \right] dt + \frac{\partial f}{\partial y} \sigma(y,t)\, dW.$$

The said transformation makes the diffusion term above equal one. See [696, 841].

Exercise 25.2.14: In asking if

$$\sigma(r,t)\sqrt{\Delta t} - \sigma(r_{\mathrm{u}},t)\sqrt{\Delta t} = -\sigma(r,t)\sqrt{\Delta t} + \sigma(r_{\mathrm{d}},t)\sqrt{\Delta t}, \tag{33.29}$$

we have $\sigma(r,t) = \sigma\sqrt{r}$ for the CIR model and $\sigma(r,t) = r\sigma$ for the geometric Brownian motion. Hence, for the CIR model,

$$\begin{aligned} r_{\mathrm{u}} &= r + \beta(\mu - r)\,\Delta t + \sigma\sqrt{r}\sqrt{\Delta t}, \\ r_{\mathrm{d}} &= r + \beta(\mu - r)\,\Delta t - \sigma\sqrt{r}\sqrt{\Delta t}, \end{aligned}$$

whereas for the geometric Brownian motion,

$$\begin{aligned} r_{\mathrm{u}} &= r + r\mu\Delta t + r\sigma\sqrt{\Delta t}, \\ r_{\mathrm{d}} &= r + r\mu\Delta t - r\sigma\sqrt{\Delta t}, \end{aligned}$$

It is easy to verify that Eq. (33.29) holds for the geometric Brownian motion but not for the CIR model.

Exercise 25.3.1: The tree should model $x(r) \equiv r^{1-\gamma}$ [475].

Exercise 25.5.1: Changes in slope are due to changes in a, changes in curvature are due to changes in b, and parallel moves are due to changes in r [239, 564].

CHAPTER 26

Exercise 26.1.1: The equilibrium model, as the no-arbitrage model is calibrated to the government bonds.

Exercise 26.2.3: Consider an n-period bond. Let $P_{\mathrm{ud}}(t+2, t+n)$ denote its price when the short rate first rises and then declines. Similarly, let $P_{\mathrm{du}}(t+2, t+n)$ denote its price when the short rate first declines and then rises. From Eq. (26.3),

$$\begin{aligned} P_{\mathrm{ud}}(t+2, t+n) &= \frac{P_{\mathrm{u}}(t+1, t+n)}{P_{\mathrm{u}}(t+1, t+2)} \frac{2e^{v_3 + \cdots + v_n}}{1 + e^{v_3 + \cdots + v_n}} \\ &= \frac{P(t, t+n)}{P(t, t+1)} \frac{2}{1 + e^{v_2 + \cdots + v_n}} e^{r_2 + v_2} \frac{2e^{v_3 + \cdots + v_n}}{1 + e^{v_3 + \cdots + v_n}}, \\ P_{\mathrm{du}}(t+2, t+n) &= \frac{P_{\mathrm{d}}(t+1, t+n)}{P_{\mathrm{d}}(t+1, t+2)} \frac{2}{1 + e^{v_3 + \cdots + v_n}} \\ &= \frac{P(t, t+n)}{P(t, t+1)} \frac{2e^{v_2 + \cdots + v_n}}{1 + e^{v_2 + \cdots + v_n}} e^{r_2} \frac{2}{1 + e^{v_3 + \cdots + v_n}}. \end{aligned}$$

These two formulas are indeed equal.

Exercise 26.2.5: The portfolio value is $P(t, t+T_1) + \beta P(t, t+T_2)$. It becomes $V_{\mathrm{u}} = P_{\mathrm{u}}(t+1, t+T_1) + \beta P_{\mathrm{u}}(t+1, t+T_2)$ if the short rate rises and $V_{\mathrm{d}} = P_{\mathrm{d}}(t+1, t+T_1) + \beta P_{\mathrm{d}}(t+1, t+T_2)$ if the short rate declines. To make it riskless, there must be no uncertainty, that is $V_{\mathrm{u}} = V_{\mathrm{d}}$. The implication is, by Eq. (26.3),

$$\begin{aligned} \beta &= -\frac{P_{\mathrm{u}}(t+1, t+T_1) - P_{\mathrm{d}}(t+1, t+T_1)}{P_{\mathrm{u}}(t+1, t+T_2) - P_{\mathrm{d}}(t+1, t+T_2)} \\ &= \frac{2\, \frac{P(t,t+T_1)}{P(t,t+1)} \left(\frac{1 - \exp[\, v_2 + \cdots + v_{T_1}\,]}{1 + \exp[\, v_2 + \cdots + v_{T_1}\,]} \right)}{2\, \frac{P(t,t+T_2)}{P(t,t+1)} \left(\frac{1 - \exp[\, v_2 + \cdots + v_{T_2}\,]}{1 + \exp[\, v_2 + \cdots + v_{T_2}\,]} \right)} \\ &= \frac{P(t, t+T_1)(1 - \exp[\, v_2 + \cdots + v_{T_1}\,])(1 + \exp[\, v_2 + \cdots + v_{T_2}\,])}{P(t, t+T_2)(1 - \exp[\, v_2 + \cdots + v_{T_2}\,])(1 + \exp[\, v_2 + \cdots + v_{T_1}\,])}. \end{aligned}$$

See [458].

Exercise 26.2.7: The return $r(t, t+n)$ is either $\ln P_{\rm d}(t+1, t+n) - \ln P(t, t+n) = (n-1)\, v + C$ with probability $1-p$ or $\ln P_{\rm u}(t+1, t+n) - \ln P(t, t+n) = C$ with probability p for some constant C. The variance of $r(t, t+n)$ is hence $p(1-p)[\,(n-1)\, v\,]^2$. Equation (26.2) says that $\sigma^2 = p(1-p)\, v^2$. Hence the variance of $r(t, t+n)$ is $(n-1)^2\sigma^2$. As for the covariance, use Eq. (6.3). See [72].

Exercise 26.2.8: Observe that $\text{Var}[\,\xi_s\,] = (1/4)\ln\delta = v^2/4$ and $v = 2\sigma$.

Exercise 26.2.9:

$$\begin{aligned}
-\ln\left(\frac{1+\delta^n}{1+\delta^{n+1}}\right) - \frac{1}{2}\ln\delta &= \ln\left(\frac{\delta^{-1/2}+\delta^{n+1/2}}{1+\delta^n}\right)\\
&= \ln\left(\frac{e^{-\sigma(\Delta t)^{1.5}} + e^{2\sigma\sqrt{\Delta t}\,[\,(T-\boldsymbol{t})+\Delta t/2\,]}}{1+e^{2\sigma\sqrt{\Delta t}\,(T-\boldsymbol{t})}}\right)\\
&\approx \ln\left(\frac{1-\sigma(\Delta t)^{1.5} + e^{2\sigma\sqrt{\Delta t}\,[\,(T-\boldsymbol{t})+\Delta t/2\,]}}{1+e^{2\sigma\sqrt{\Delta t}\,(T-\boldsymbol{t})}}\right)\\
&\approx \frac{-\sigma(\Delta t)^{1.5} + e^{2\sigma\sqrt{\Delta t}\,[\,(T-\boldsymbol{t})+\Delta t/2\,]} - e^{2\sigma\sqrt{\Delta t}\,(T-\boldsymbol{t})}}{1+e^{2\sigma\sqrt{\Delta t}\,(T-\boldsymbol{t})}}\\
&\approx \frac{(1/2)(2\sigma\sqrt{\Delta t}\,[\,(T-\boldsymbol{t})+\Delta t/2)\,]^2 - (1/2)[\,2\sigma\sqrt{\Delta t}\,(T-\boldsymbol{t})\,]^2}{2}\\
&\approx \sigma^2(T-\boldsymbol{t})(\Delta t)^2.
\end{aligned}$$

Exercise 26.2.10: Following the proof of Exercise 26.2.9, we have

$$-\ln\left(\frac{1+\delta^{t-1}}{1+\delta^t}\right) - \frac{1}{2}\ln\delta = \ln\left(\frac{\delta^{-1/2}+\delta^{t-1/2}}{1+\delta^{t-1}}\right) \approx \sigma^2\boldsymbol{t}\,(\Delta t)^2.$$

Exercise 26.3.1: Rearrange Eqs. (26.7) and (26.8) as simultaneous equations:

$$\begin{aligned}
f(P_{\rm u}, P_{\rm d}) &\equiv P_{\rm u} + P_{\rm d} - \frac{2(1+r_1)}{(1+y)^i} = 0,\\
g(P_{\rm u}, P_{\rm d}) &\equiv P_{\rm u}{}^{-1/(i-1)} - 1 - e^{2\kappa}[\,P_{\rm d}{}^{-1/(i-1)} - 1\,] = 0.
\end{aligned}$$

Because $P_{\rm u}$ and $P_{\rm d}$ are functions of r_i and v_i, $f(P_{\rm u}, P_{\rm d})$ and $g(P_{\rm u}, P_{\rm d})$ are also functions of r_i and v_i. Denote them by $F(r_i, v_i)$ and $G(r_i, v_i)$, respectively. For brevity, we use $f(r, v)$ instead of $f(r_i, v_i)$, $g(r, v)$ instead of $g(r_i, v_i)$, and so on. By the Newton–Raphson method, the $(k+1)$th approximation to (r_i, v_i) – denoted as $(r(k+1), v(k+1))$ – satisfies

$$\begin{bmatrix} \dfrac{\partial F(r(k), v(k))}{\partial r} & \dfrac{\partial F(r(k), v(k))}{\partial v}\\[2ex] \dfrac{\partial G(r(k), v(k))}{\partial r} & \dfrac{\partial G(r(k), v(k))}{\partial v}\end{bmatrix}\begin{bmatrix}\Delta r(k+1)\\ \Delta v(k+1)\end{bmatrix} = -\begin{bmatrix} F(r(k), v(k))\\ G(r(k), v(k))\end{bmatrix},$$

where $\Delta r(k+1) \equiv r(k+1) - r(k)$ and $\Delta v(k+1) \equiv v(k+1) - v(k)$. We need $\partial F/\partial r$, $\partial F/\partial v$, $\partial G/\partial r$, and $\partial G/\partial v$ to solve the preceding matrix for $(r(k+1), v(k+1))$. Obviously,

$$\frac{\partial F}{\partial r} = \frac{\partial P_{\rm u}}{\partial r} + \frac{\partial P_{\rm d}}{\partial r}, \qquad \frac{\partial F}{\partial v} = \frac{\partial P_{\rm u}}{\partial v} + \frac{\partial P_{\rm d}}{\partial v}.$$

By the chain rule,

$$\frac{\partial G}{\partial r} = \frac{\partial g}{\partial P_{\rm u}}\frac{\partial P_{\rm u}}{\partial r} + \frac{\partial g}{\partial P_{\rm d}}\frac{\partial P_{\rm d}}{\partial r}, \qquad \frac{\partial G}{\partial v} = \frac{\partial g}{\partial P_{\rm u}}\frac{\partial P_{\rm u}}{\partial v} + \frac{\partial g}{\partial P_{\rm d}}\frac{\partial P_{\rm d}}{\partial v}.$$

In the preceding four equations, the items we need to evaluate are

$$\frac{\partial P_{\rm u}}{\partial r},\quad \frac{\partial P_{\rm d}}{\partial r},\quad \frac{\partial P_{\rm u}}{\partial v},\quad \frac{\partial P_{\rm d}}{\partial v},\quad \frac{\partial g}{\partial P_{\rm u}},\quad \frac{\partial g}{\partial P_{\rm d}}.$$

The differential tree method can compute them as follows. Working backward, the method ends up with $P_{\rm u}$, $P_{\rm d}$, $\partial P_{\rm u}/\partial r$, $\partial P_{\rm d}/\partial r$, $\partial P_{\rm u}/\partial v$, and $\partial P_{\rm d}/\partial v$. The remaining $\partial g/\partial P_{\rm u}$ and $\partial g/\partial P_{\rm d}$ can be computed directly from the definition of g:

$$\frac{\partial g}{\partial P_{\rm u}} = -\frac{1}{i-1}\, P_{\rm u}^{-i/(i-1)}, \qquad \frac{\partial g}{\partial P_{\rm d}} = \frac{e^{2\kappa}}{i-1}\, P_{\rm d}^{-i/(i-1)}.$$

Thus $\partial F/\partial r$, $\partial F/\partial v$, $\partial G/\partial r$, and $\partial G/\partial v$ can be computed.

This backward-induction algorithm runs in cubic time because the Newton–Raphson method takes only a few iterations to get to the desired accuracy. See [625].

Exercise 26.3.4: The proof is similar to Exercise 23.2.3. From the process,

$$\begin{aligned}\sigma(t)^2\Delta t &= \mathrm{Var}[\,\ln r(t+\Delta t) - \ln r(t)\,]\\ &= \mathrm{Var}[\,\ln r(t)\{\,1+\Delta r(t)/r(t)\,\} - \ln r(t)\,]\\ &= \mathrm{Var}[\,\ln(1+\Delta r(t)/r(t))\,]\\ &\approx \mathrm{Var}[\,\Delta r(t)/r(t)\,]\\ &= \mathrm{Var}[\,\Delta r(t)\,]/r(t)^2.\end{aligned}$$

So $\mathrm{Var}[\,\Delta r(t)\,] \approx r(t)^2\sigma(t)^2\Delta t$. See [514, p. 482].

Exercise 26.4.1: The normal model does, for the following reason. It has fatter left tails and thinner right tails for the probability distribution of interest rates in the future (see Fig. 6.1). It therefore gives thinner left tails and fatter right tails for the probability distribution of bond prices. See [476].

Exercise 26.4.2: Apply the moment generating function of the normal distribution with mean $r_j + \mu_{i,j}\Delta t$ and variance $\sigma^2\Delta t$. In other words, use formula (6.8) with $t = -\Delta t$, $\mu = r_j + \mu_{i,j}\Delta t$, and $\sigma = \sigma\sqrt{\Delta t}$. The result is

$$e^{-r_j\Delta t - \mu_{i,j}\,(\Delta t)^2 + \sigma^2(\Delta t)^3/2} \approx e^{-r_j\Delta t}\left[1 - \mu_{i,j}(\Delta t)^2 + \frac{1}{2}\,\sigma^2(\Delta t)^3\right].$$

See [477].

Programming Assignment 26.4.3: Because the unreachable nodes' corresponding branch entries are useless, we can improve the algorithm by keeping track of the upper and the lower bounds of the reachable nodes in each column of branch and limiting the calculation to those nodes.

Programming Assignment 26.4.5: Apply the trinomial tree algorithm for the Hull–White model to $x \equiv \ln r$. See also [215, 848].

Exercise 26.5.1: It is $(\sigma^2/\kappa)(e^{-\kappa(T-t)} - e^{-2\kappa(T-t)})$ [746, p. 581].

Exercise 26.5.2: From Eq. (26.14), the process for $df(t, T)$ in Eq. (26.15) can be written as

$$df(t, T) = \mu(t, T)\,dt + \sigma(t, T)\,dW_t.$$

Now $r(t) = f(t, t) = f(0, t) + \int_0^t \mu(s, t)\,ds + \int_0^t \sigma(s, t)\,dW_s$. Differentiate it with respect to t to obtain

$$\begin{aligned}dr(t) &= \frac{\partial f(0, t)}{\partial t}\,dt + \left[\int_0^t \frac{\partial\mu(s, t)}{\partial t}\,ds\right]dt + \left[\int_0^t \frac{\partial\sigma(s, t)}{\partial t}\,dW_s\right]dt + \mu(t, t)\,dt + \sigma(t, t)\,dW_t\\ &= \left\{\frac{\partial}{\partial T}\left[f(0, T) + \int_0^t \mu(s, T)\,ds + \int_0^t \sigma(s, T)\,dW_s\right]\bigg|_{T=t}\right\}dt + \mu(t, t)\,dt + \sigma(t, t)\,dW_t\\ &= \left[\mu(t, t) + \frac{\partial f(t, T)}{\partial T}\bigg|_{T=t}\right]dt + \sigma(t, t)\,dW_t\\ &= \left[\sigma(t, t)\,\lambda(t) + \frac{\partial f(t, T)}{\partial T}\bigg|_{T=t}\right]dt + \sigma(t, t)\,dW_t.\end{aligned}$$

See [515, Appendix 1].

Exercise 26.5.4: From Eq. (26.21),

$$C_t = P(t, T)\,N(d_t) - XP(t, s)\,N(d_t - \sigma_t),$$

where $d_t \equiv (1/\sigma_t)\,\ln(P(t, T)/[\,XP(t, s)\,]) + (\sigma_t/2)$. (1) Let $\sigma_t \equiv \sigma(T-s)\sqrt{s-t}$. (2) Let $\sigma_t \equiv \beta(s-t, T-t)\sqrt{\phi(s-t)}$.

Exercise 26.5.5: From Example 26.5.3 it is $\frac{\partial r(t,T)}{\partial r(t)}\,\sigma = -\frac{1}{T-t}\,\frac{\partial \ln P(t,T)}{\partial r(t)}\,\sigma = \beta(t, T)\,\frac{\sigma}{T-t}$.

Exercise 26.5.7: If every node contains forward rates up to the maturity of the underlying bond, a term structure up to that maturity can be constructed. The tree then needs to extend to only the maturity of the claim.

Exercise 26.5.8: Use Eqs. (24.5), (26.23), and (26.24) to arrive at

$$1 = \frac{1}{2} \times \exp\left[-\left\{ \Delta t \int_{t+\Delta t}^{T} \mu(t, u)\, du + \sqrt{\Delta t} \int_{t+\Delta t}^{T} \sigma(t, u)\, du \right\} \right] + \frac{1}{2} \times \exp\left[-\left\{ \Delta t \int_{t+\Delta t}^{T} \mu(t, u)\, du - \sqrt{\Delta t} \int_{t+\Delta t}^{T} \sigma(t, u)\, du \right\} \right].$$

Exercise 26.5.9: Equation (26.25) can be applied iteratively to obtain

$$\begin{aligned}
(\Delta t)^2 \mu(t, T) &\approx \ln \frac{e^x + e^{-x}}{e^y + e^{-y}} \\
&\approx \frac{e^x - e^y + e^{-x} - e^{-y}}{e^y + e^{-y}} \\
&\approx \frac{[\, e^{(\Delta t)^{1.5} \sigma(t,T)} - 1\,] e^y - [\, 1 - e^{-(\Delta t)^{1.5} \sigma(t,T)}\,] e^{-y}}{e^y + e^{-y}} \\
&\approx \frac{(\Delta t)^{1.5} \sigma(t, T)\, e^y - (\Delta t)^{1.5} \sigma(t, T)\, e^{-y}}{e^y + e^{-y}} \\
&= (\Delta t)^{1.5} \sigma(t, T)\, \frac{e^y - e^{-y}}{e^y + e^{-y}} \\
&\approx (\Delta t)^{1.5} \sigma(t, T)\, \frac{e^x - e^{-x}}{e^x + e^{-x}} \\
&= (\Delta t)^{1.5} \sigma(t, T)\, \tanh(x),
\end{aligned}$$

where $x \equiv \sqrt{\Delta t} \int_{t+\Delta t}^{T} \sigma(t, u)\, du$ and $y \equiv \sqrt{\Delta t} \int_{t+\Delta t}^{T-\Delta t} \sigma(t, u)\, du$.

Chapter 27

Exercise 27.3.1: Without loss of generality, assume that the par value of the bond is \$1. Let the call price at time t be $C(t)$. Suppose the sinking-fund provision requires the issuer to retire the principal by F_i dollars at time t_i for $i = 1, 2, \ldots, n$, where $\sum_{i=1}^{n} F_i = 1$. As before, it is best to view the embedded option as n separate options. In particular, the ith option has the right to buy a face value F_i callable bond at par at maturity (time t_i) with the call schedule equal to $F_i C(t)$ for $0 \le t \le t_i$. The underlying bond is otherwise identical to the original bond. The option value can be calculated by backward induction. The desired price is that of the otherwise identical straight bond minus the n options. See [848].

Exercise 27.3.2: Suppose the bond were trading below its conversion value. Consider the following strategy: Buy one CB and sell short the stock with the number of shares equal to the conversion ratio, and cover the short position by conversion. Note that there is a positive initial cash flow, and the number of shares exactly covers the short position. See [325, p. 374].

Exercise 27.3.3: As bondholders convert, the price of the stock will decline because of dilution. The correct definition for the conversion value would use the stock price *after* conversion, taking into account the dilution issue. However, here we are assuming that the CBs in questions are the only securities to affect the number of outstanding shares, which is rarely true. See [221, Chapter 8] and [325, p. 373n].

Exercise 27.3.4: A CB's market value is at least its conversion value, but early conversion generates exactly the conversion value.

Exercise 27.3.5: If the current stock price is so high that the conversion option always expires in the money and is exercised, every node will be an exercise node by induction from the terminal nodes. As a result, every node's CB value equals the stock price.

Exercise 27.3.6: On a per-share basis, the desired relation is

$$\text{CB value}\,(B) = \text{warrant value}\,(W) + \text{straight value}\,(s), \tag{33.30}$$

where the warrant's strike price X equals the CB's conversion price P plus the coupon payment. Assume that the stock price of the issuer follows the binomial model and that there are n periods before maturity. Let c_i be the coupon payment at time, P be the par value of the option-free bond, R be the riskless gross return per period, and $s_i \equiv PR^{-(n-i)} + \sum_{j=i}^{n} c_j R^{-(j-i)}$, the straight value at time i. Identity (33.30) can be proved by induction as follows. At maturity, $W = \max(S - X, 0)$ and $B = \max(S, X) = W + (P + c_n) = W + s_n$. Inductively, at time i,

$$B = \frac{pB_{\mathrm{u}} + (1-p)\, B_{\mathrm{d}}}{R} + c_i = \frac{p(W_{\mathrm{u}} + s_{i+1}) + (1-p)(W_{\mathrm{d}} + s_{i+1})}{R} + c_i = W + s_i.$$

See [221, Chapter 8].

Exercise 27.4.1: It is identical to the valuation of riskless bonds by use of backward induction in a risk-neutral economy except that at every node at time i, the bond may branch with probability p_i into a default state and thus pay zero dollar at maturity. In other words, the risk-neutral probabilities at time i are those for the riskless bond multiplied by $1 - p_i$. See [514, p. 570] or [535] for the more general case in which the firm pays a fraction of the par value instead of zero.

Exercise 27.4.2: When interest rates rise, both the price of the noncallable bond component and the price of the embedded call option fall. The fall in price of the embedded option will offset some of the fall in price of the noncallable bond. See [325, p. 324].

Exercise 27.4.3: (1) From Eq. (27.1), $\text{price}_{\mathrm{c}} = \text{price}_{\mathrm{nc}} - \text{call price}$. Differentiate both sides and then divide by the price of the callable bond, $\text{price}_{\mathrm{c}}$, to obtain

$$-\mathrm{OAD}_{\mathrm{c}} = \frac{\partial(\text{price}_{\mathrm{nc}})}{\partial y} \frac{1}{\text{price}_{\mathrm{c}}} - \frac{\partial(\text{call price})}{\partial y} \frac{1}{\text{price}_{\mathrm{c}}}. \tag{33.31}$$

The first term on the right-hand side of Eq. (33.31) is

$$\frac{\partial(\text{price}_{\mathrm{nc}})}{\partial y} \frac{1}{\text{price}_{\mathrm{nc}}} \frac{\text{price}_{\mathrm{nc}}}{\text{price}_{\mathrm{c}}} = -\text{duration}_{\mathrm{nc}} \times \frac{\text{price}_{\mathrm{nc}}}{\text{price}_{\mathrm{c}}}. \tag{33.32}$$

From the chain rule and delta's definition,

$$\frac{\partial(\text{call price})}{\partial y} = \frac{\partial(\text{call price})}{\partial(\text{price}_{\mathrm{nc}})} \frac{\partial(\text{price}_{\mathrm{nc}})}{\partial y} = \Delta \times \frac{\partial(\text{price}_{\mathrm{nc}})}{\partial y}.$$

The second term on the right-hand side of Eq. (33.31) thus is

$$\begin{aligned} \frac{\partial(\text{call price})}{\partial y} \frac{1}{\text{price}_{\mathrm{nc}}} \frac{\text{price}_{\mathrm{nc}}}{\text{price}_{\mathrm{c}}} &= \Delta \times \frac{\partial(\text{price}_{\mathrm{nc}})}{\partial y} \frac{1}{\text{price}_{\mathrm{nc}}} \frac{\text{price}_{\mathrm{nc}}}{\text{price}_{\mathrm{c}}} \\ &= -\Delta \times \text{duration}_{\mathrm{nc}} \times \frac{\text{price}_{\mathrm{nc}}}{\text{price}_{\mathrm{c}}}. \end{aligned}$$

Combine Eq. (33.32) and the preceding equation to yield the desired result. See [325, pp. 332–334].

Exercise 27.4.4: See Exercise 23.3.1.

Exercise 27.4.5: Increased volatility makes the bond more likely to be called, reducing the investor's return.

Exercise 27.4.6: (1) The OAS increases. The price of course cannot drop below par. (2) The OAS decreases but will level off at zero after a certain point.

Exercise 27.4.7: Consider a T-period zero-coupon bond callable at time m with call price X. For ease of comparison, assume a binomial interest rate process in which $r_n(i)$ denotes the short rate for period i on the nth path. The true value is derived by backward induction thus:

$$P_{\mathrm{b}} \equiv \frac{1}{2^T} \sum_{n_1=1}^{2^m} \times \min\left(\sum_{n_2=1}^{2^{T-m}} \frac{100}{\prod_{i=1}^{T}[1 + r_{n_1 n_2}(i) + \mathrm{OAS}]}, \frac{X}{\prod_{i=1}^{m}[1 + r_{n_1 n_2}(i) + \mathrm{OAS}]} \right),$$

whereas the Monte Carlo method calculates

$$P_{\mathrm{MC}} \equiv \frac{1}{2^T} \sum_{n=1}^{2^T} \frac{1}{\prod_{i=1}^{m} [\,1 + r_n(i) + \mathrm{OAS}\,]} \times \min\left(\frac{100}{\prod_{i=m+1}^{T} [\,1 + r_n(i) + \mathrm{OAS}\,]}, X \right)$$

$$= \frac{1}{2^T} \sum_{n_1=1}^{2^m} \sum_{n_2=1}^{2^{T-m}} \frac{1}{\prod_{i=1}^{m} [\,1 + r_{n_1 n_2}(i) + \mathrm{OAS}\,]} \times \min\left(\frac{100}{\prod_{i=m+1}^{T} [\,1 + r_{n_1 n_2}(i) + \mathrm{OAS}\,]}, X \right)$$

$$= \frac{1}{2^T} \sum_{n_1=1}^{2^m} \sum_{n_2=1}^{2^{T-m}} \times \min\left(\frac{100}{\prod_{i=1}^{T} [\,1 + r_{n_1 n_2}(i) + \mathrm{OAS}\,]}, \frac{X}{\prod_{i=1}^{m} [\,1 + r_{n_1 n_2}(i) + \mathrm{OAS}\,]} \right).$$

From Jensen's inequality $\sum_{n=1}^{l} \min(x_n, c) \leq \min(\sum_{n=1}^{l} x_n, c)$, we conclude that $P_{\mathrm{MC}} \leq P_{\mathrm{b}}$. Similarly the Monte Carlo method tends to overestimate the value of putable bonds. See [349].

Exercise 27.4.11: In Fig. 27.7(d), (1) change 0.936 to 1.754 because $1.754 = (106.754 - 5) - 100$ and (2) change 0.482 to 0.873 because $0.873 = (1.754 + 0.071)/1.045$.

Exercise 27.4.12: We prove the claim for binomial interest rate trees with reference to Fig. 23.1. The same argument can be applied to any tree. Let s represent the OAS. Consider zero-coupon bonds first. Backward induction performs

$$P_{\mathrm{A}} = \frac{P_{\mathrm{B}} + P_{\mathrm{C}}}{2e^{r+s}} = e^{-s}\, \frac{P_{\mathrm{B}} + P_{\mathrm{C}}}{2e^{r}}$$

at each node. It follows by mathematical induction that the n-period benchmark bond price $P_{\mathrm{b}}(n)$ and the n-period nonbenchmark bond price $P_{\mathrm{nb}}(n)$ are related by $P_{\mathrm{b}}(n) = P_{\mathrm{nb}}(n)\, e^{sn}$. Let $r_{\mathrm{b}}(n)$ and $r_{\mathrm{nb}}(n)$ denote their respective yields. Then $e^{-r_{\mathrm{b}}(n)\, n} = e^{-r_{\mathrm{nb}}(n)\, n} e^{sn}$, i.e., $r_{\mathrm{b}}(n) = r_{\mathrm{nb}}(n) - s$. By assumption, $r_{\mathrm{b}}(n)$ are all identical. We can hence drop the dependence on the maturity n, and the identity becomes $r_{\mathrm{b}} = r_{\mathrm{nb}} - s$. Consider coupon bonds with cash flow $C_1, C_2, \ldots, C_n$. Then

$$P_{\mathrm{nb}} \equiv \sum_{i=1}^{n} C_i\, P_{\mathrm{nb}}(n) = \sum_{i=1}^{n} C_i e^{-r_{\mathrm{nb}} i} = \sum_{i=1}^{n} C_i e^{-(r_{\mathrm{b}}+s)\, i}.$$

The yield spread is hence s, as claimed.

CHAPTER 28

Exercise 28.5.1: It is because the servicing and guaranteeing fee is a percentage of the remaining principal, and the principal is paid down through time.

Exercise 28.6.1: The investor first has a capital gain. Now a pass-through trades at a discount because its coupon rate is lower than the current coupon rate of new issues. The prepayments therefore can be reinvested at a higher coupon rate. See [325, p. 254].

Exercise 28.6.2: Given the premise, there is a smaller incentive to refinance a loan with a lower remaining balance, other things being equal. Because 30-year loans amortize more slowly than 15-year loans, MBSs backed by 30-year mortgages should prepay faster. See [433].

Exercise 28.6.3: Even without the transactions costs, expression (28.2) says that refinancing does not make economic sense unless the old rate exceeds the new one.

Exercise 28.6.4: Expression (28.1) says that, at the first refinancing, the remaining balance is

$$C\, \frac{1 - (1+r)^{-n+a}}{r},$$

where C is the monthly payment. However, it is easy to show (or from Eq. (29.4)) that

$$C = \mathcal{O} \times \frac{r(1+r)^n}{(1+r)^n - 1}, \tag{33.33}$$

where $\mathcal{O}$ denotes the original balance. When the preceding expressions are combined, the remaining balance is

$$\mathcal{O}' = \mathcal{O} \times \frac{(1+r)^n}{(1+r)^n - 1}\, [\,1 - (1+r)^{-n+a}\,] = \mathcal{O} \times \frac{(1+r)^n - (1+r)^a}{(1+r)^n - 1}.$$

The new monthly payment is

$$C' = \mathcal{O}' \times \frac{r(1+r)^n}{(1+r)^n - 1} = \mathcal{O} \times \frac{(1+r)^n - (1+r)^a}{(1+r)^n - 1}\, \frac{r(1+r)^n}{(1+r)^n - 1}$$

by Eq. (33.33). After the preceding process is iterated, the monthly payment after the ith refinancing is

$$\mathcal{O} \times \left[\frac{(1+r)^n - (1+r)^a}{(1+r)^n - 1} \right]^i \frac{r(1+r)^n}{(1+r)^n - 1}.$$

Exercise 28.6.5: Although both scenarios have the same rate difference of 2%, the second deal is better because of a higher refinancing incentive, as prescribed by expression (28.3).

CHAPTER 29

Exercise 29.1.2: From Eq. (29.2) with $k = 1$, the remaining principal balance per $1 of original principal balance is

$$1 - \frac{x-1}{x^n - 1} = 1 - \frac{1}{x^{n-1} + x^{n-2} + \cdots + 1},$$

where $x \equiv 1 + r/m \geq 1$. The preceding equation is monotonically increasing in x. Hence different r's yield different balances as they give rise to different x's.

Programming Assignment 29.1.3: See Fig. 33.8.

Exercise 29.1.5: First, run the algorithm in Fig. 33.8 but on a short rate tree with $n + k - 1$ periods to obtain all the desired spot rates up to time n when the swap expires. Only the nodes up to time $n - 1$ are needed. Now use backward induction to derive at each node the price of a swap with $1 of notional principal initiated at that node and ending at time n. The swap's amortization amount follows the original swap's. Specifically, let a be the amortizing amount at the current node, which is determined uniquely by the node's k-period spot rate. Then the value at the node equals $C + (1 - a)(p_u P_u + p_d P_d)/(1 + r)$, where r is the node's short rate, C is its cash flow, p_u and p_d are the branching probabilities, and P_u and P_d are the values of the swaps as described above at the two successor nodes. Note the use of scaling. The running time is $O(kn^2)$.

Exercise 29.1.6:

Month	*6*	*12*	*18*	*24*	*30*	*36*
CPR	1.2	3.12	5.544	11.04	8.1	7.5

with the help of Eq. (29.5).

Exercise 29.1.7: No. See [323, p. 362] for an example.

Algorithm for generating spot rate dynamics:

```
input:   n, k, r[n][n];
real     s[n-k+1][n-k+1], P[n];
integer  i, j, l;
for (i = n-k down to 0) {
         // Backward induction to obtain discount factors.
         for (j = 0 to i+k-1)
            P[j] := 1/(1+r[i+k-1][j]);
         for (l = i+k-2 down to i)
            for (j = 0 to l)
               P[j] := 0.5(P[j] + P[j+1])/(r[l][j]+1);
         // Turn discount factors into spot rates.
         for (j = 0 to i) s[i][j] := P[j]^(-1/k) - 1;
}
return s[ ][ ];
```

Figure 33.8: Algorithm for generating spot rate dynamics. $r[\,i\,][\,j\,]$ is the $(j+1)$th short rate for period $i+1$, the short rate tree covers n periods, and k is the maturity of the desired spot rates. The spot rates are stored in $s[\,][\,]$. Specifically, $s[\,i\,][\,j\,]$ refers to the desired spot rate at the same node as $r[\,i\,][\,j\,]$, where $0 \leq i \leq n-k$ and $0 \leq j \leq i$. All rates are measured by the period.

Exercise 29.1.8: Sum $\overline{P_i}$ of Eq. (29.6) and $\overline{I_i}$ of Eq. (29.7) with $\alpha = 0$.

Exercise 29.1.9: (1) Substitute Eq. (29.8) into the formula's B_i (2)

$$\begin{aligned}\overline{P_i} + \mathrm{PP}_i &= b_{i-1} P_i + B_{i-1} \frac{\mathrm{RB}_i}{\mathrm{RB}_{i-1}} \times \mathrm{SMM}_i \\ &= b_{i-1} P_i + \mathrm{RB}_{i-1} \times b_{i-1} \frac{\mathrm{RB}_i}{\mathrm{RB}_{i-1}} \times \mathrm{SMM}_i \\ &= b_{i-1}(P_i + \mathrm{RB}_i \times \mathrm{SMM}_i).\end{aligned}$$

Exercise 29.1.10: For Eq. (29.9), just observe that $B_i/\mathrm{Bal}_i = (B_{i-1}/\mathrm{Bal}_{i-1})(1 - \mathrm{SMM}_i)$. As for Eqs. (29.10), use Eqs. (29.2) and (29.3) to prove the formula for $\overline{P_i}$. The formula for $\overline{I_i}$ is due to amortization.

Exercise 29.1.12: Let s denote the SMM. For simplicity, assume that the original balance is \$1 and r is the period yield:

$$\begin{aligned}\mathrm{PO} &= \sum_{i=1}^{n} \frac{\overline{P_i} + \mathrm{PP}_i}{(1+r)^i} \\ &= \sum_{i=1}^{n} \frac{(1-s)^{i-1} \frac{r(1+r)^{i-1}}{(1+r)^n - 1} + B_i \frac{s}{1-s}}{(1+r)^i} \quad \text{from Eqs. (29.3) and (29.10) and Exercise 29.1.9(1)} \\ &= \sum_{i=1}^{n} \frac{(1-s)^{i-1} \frac{r(1+r)^{i-1}}{(1+r)^n - 1} + \frac{(1+r)^n - (1+r)^i}{(1+r)^n - 1} s(1-s)^{i-1}}{(1+r)^i} \quad \text{from Eqs. (29.2) and (29.9)} \\ &= \frac{1}{(1+r)^n - 1} \left\{ \frac{r}{1+r} \frac{1-(1-s)^n}{s} + s \frac{(1+r)^n - (1-s)^n}{r+s} - [\,1 - (1-s)^n\,] \right\}.\end{aligned}$$

Similarly,

$$\begin{aligned}\mathrm{IO} &= \sum_{i=1}^{n} \frac{\overline{I_i}}{(1+r)^i} = \sum_{i=1}^{n} \frac{(1-s)^{i-1} \times \mathrm{RB}_{i-1} \times r}{(1+r)^i} \\ &= r \sum_{i=1}^{n} \frac{(1-s)^{i-1} \frac{(1+r)^n - (1+r)^{i-1}}{(1+r)^n - 1}}{(1+r)^i} \quad \text{from (29.2)} \\ &= \frac{r}{(1+r)^n - 1} \left[\frac{(1+r)^n - (1-s)^n}{r+s} - \frac{1}{1+r} \frac{1-(1-s)^n}{s} \right].\end{aligned}$$

Exercise 29.1.13: Let the principal payments be $P_1, P_2, \ldots, P_n$ and the interest payment at time i be $I_i \equiv r(P - \sum_{j=1}^{i-1} P_j)$ by the principle of amortization, where r is the period yield and $P = P_1 + P_2 + \cdots + P_n$ is the original principal amount. The FV of the combined cash flow $P_i + I_i$, $i = 1, 2, \ldots, n$, is

$$\begin{aligned}\sum_{i=1}^{n} \left[P_i + r\left(P - \sum_{j=1}^{i-1} P_j \right) \right] (1+r)^{n-i} &= \sum_{i=1}^{n} \left(P_i + r \sum_{j=i}^{n} P_j \right) (1+r)^{n-i} \\ &= \sum_{i=1}^{n} P_i (1+r)^{n-i} \left[1 + \sum_{j=1}^{i} r(1+r)^{i-j} \right] \\ &= \sum_{i=1}^{n} P_i (1+r)^{n-i} (1+r)^i \\ &= P(1+r)^n,\end{aligned}$$

independent of how P is distributed among the P_is.

Exercise 29.1.15: They are \$125,618, \$115,131, \$67,975, \$34,612, \$20,870, and \$0 for months 1–6. See the formula in Exercise 29.1.9(1).

Programming Assignment 29.1.16: Validate your program by running it with zero original balance for tranche C (that is, $\mathcal{O}[\,3\,] = 0$) and comparing the output with those in Fig. 29.11.

Exercise 29.2.1: Because the fast refinancers will exit the pool at a faster rate, an increasingly larger proportion of the remaining population will be the slow refinancers. Consequently, the refinancing rate of the pool will move toward that of the slow refinancers. See [433].

Exercise 29.2.2: PO $\approx s/(r+s)$ and IO $\approx r/(r+s)$. Hence, $\frac{\partial(\text{PO})/\partial s}{\text{PO}} \approx \frac{r}{s(r+s)}$ and $\frac{\partial(\text{IO})/\partial s}{\text{IO}} \approx -\frac{1}{r+s}$.

Exercise 29.2.3: Like interests, servicing fees are a percentage of the principal [829, p. 105].

Exercise 29.3.1: A premium-priced MBS must have a coupon rate exceeding the market discount rate. This means that the prepaid principal is less than the PV of the future cash flow foregone by such a prepayment. Hence prepayment depresses the value of the cash flow. The argument for discount MBSs is symmetric.

Exercise 29.3.2: The formula is valid only if the cash flow is independent of yields. For MBSs, this does not hold. See [55, p. 132].

Exercise 29.3.3: Refer to Fig. 28.8. In a bull market, Treasury securities' prices go up, whereas the MBSs' prices go down. This hedger therefore has the worst of both worlds, losing money on both. See [54, p. 204].

Exercise 29.3.4: This model does not take into account the negative convexity of the security [304].

CHAPTER 30

Exercise 30.2.1:

LIBOR Change (Basis Points)	*−300*	*−200*	*−100*	*0*	*+100*	*+200*	*+300*
Conventional floater	4.5	5.5	6.5	7.5	8.5	9.5	10.5
Superfloater	2.0	3.5	5.0	6.5	8.0	9.5	11.0

CHAPTER 31

Exercise 31.1.1: Let $\boldsymbol{\omega} \equiv [\,\omega_1, \omega_2, \ldots, \omega_n\,]^{\mathrm{T}}$ be the vector of portfolio weights, $\bar{\boldsymbol{r}} \equiv [\,\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_n\,]^{\mathrm{T}}$ be the vector of mean security returns, and Q be the covariance matrix of security returns. The efficient portfolio for a target return of $\bar{r}$ is determined by

$$\begin{array}{ll} \text{minimize} & \min_{\boldsymbol{\omega}} \boldsymbol{\omega}^{\mathrm{T}} Q \boldsymbol{\omega}, \\ \text{subject to} & \boldsymbol{\omega}^{\mathrm{T}} \bar{\boldsymbol{r}} = \bar{r}, \\ & \boldsymbol{\omega}^{\mathrm{T}} \mathbf{1} = 1, \end{array}$$

where $\mathbf{1} \equiv [\,1, 1, \ldots, 1\,]^{\mathrm{T}}$. See [3].

Exercise 31.1.2: The variance is $\omega_1^2 \sigma_1^2 + (1-\omega_1)^2 \sigma_2^2 - 2\omega_1(1-\omega_1)\,\sigma_1\sigma_2$. Pick $\omega_1 = \sigma_2/(\sigma_1+\sigma_2)$, which lies between zero and one, to make the variance zero. See [317, p. 74].

Exercise 31.1.3: (1) The problem is

$$\begin{array}{ll} \text{minimize} & -(1/2) \sum_i \sum_j \omega_i \omega_j \sigma_{ij}, \\ \text{subject to} & \sum_i \omega_i = 1, \\ & \omega_i \geq 0 \quad \text{for all } i. \end{array}$$

Let i^* be such that $\sigma_{i^*i^*} = \max_i \sigma_{ii}$. It is not hard to show that $\sigma_{ii}\sigma_{jj} > \sigma_{ij}^2$ for all i, j [465, p. 398]; hence $\sigma_{i^*i^*} > \sigma_{i^*i}$ for all $i \neq i^*$. The **Kuhn–Tucker conditions** for the optimal solution say that there exist μ and $\nu_1, \nu_2, \ldots$ such that [721, pp. 522–523]:

$$\begin{aligned} & -\sum_j \sigma_{ij}\omega_j + \mu - \nu_i = 0 \quad \text{for all } i, \\ & \sum_i \omega_i = 1, \\ & \omega_i \geq 0 \quad \text{for all } i, \\ & \nu_i \geq 0 \quad \text{for all } i, \\ & \nu_i \omega_i = 0 \quad \text{for all } i. \end{aligned}$$

Observe that $\sum_i (\mu - v_i)\,\omega_i = \mu > 0$ is the desired value. The feasible solution in which $\omega_{i^*} = 1$ and $\omega_i = 0$ for $i \neq i^*$ satisfies the above by the selection of $\mu = \sigma_{i^*i^*}$ and $v_i = \sigma_{i^*i^*} - \sigma_{i^*i} \geq 0$ for all i.

(2) The problem is

$$\begin{aligned} &\text{minimize} && (1/2)\sum_i \sum_j \omega_i \omega_j \sigma_{ij}, \\ &\text{subject to} && \sum_i \omega_i = 1, \\ &&& \omega_i \geq 0 \quad \text{for all } i. \end{aligned}$$

Use the Lagrange multiplier to show that the optimal solution satisfies

$$\begin{aligned} &\sum_j \sigma_{ij}\omega_j + \mu = 0 \quad \text{for all } i, \\ &\sum_i \omega_i = 1, \\ &\omega_i \geq 0 \quad \text{for all } i. \end{aligned}$$

Observe that $-\mu \sum_i \omega_i = -\mu > 0$ is the desired value. Let $\boldsymbol{\omega} \equiv [\,\omega_1, \omega_2, \ldots, \omega_n\,]^{\mathrm{T}}$ and $\mathbf{1} \equiv [\,1, 1, \ldots, 1\,]^{\mathrm{T}}$. Hence, $\boldsymbol{\omega} = -\mu C^{-1}\mathbf{1}$, which satisfies the conditions $\omega_i \geq 0$ for all i by the assumption. Combining this with $\mathbf{1}^{\mathrm{T}}\boldsymbol{\omega} = 1$, we have $\mathbf{1}^{\mathrm{T}}(-\mu C^{-1}\mathbf{1}) = 1$, implying that $-\mu = 1/(\mathbf{1}^{\mathrm{T}} C^{-1}\mathbf{1})$. If we let $C^{-1} \equiv [\,a_{ij}\,]$, then $-\mu = 1/(\sum_i \sum_j a_{ij})$.

Exercise 31.1.4: Note that $r(T) + 1$ is lognormally distributed and $r_{\mathrm{c}}(T)$ is normally distributed. The mean and the variance of $r(T)$ and those of $r_{\mathrm{c}}(T)$ are therefore related by

$$\begin{aligned} \mu(T) &= e^{\mu_{\mathrm{c}}(T) + 0.5\sigma_{\mathrm{c}}^2(T)} - 1, \\ \sigma^2(T) &= e^{2\mu_{\mathrm{c}}(T) + 2\sigma_{\mathrm{c}}^2(T)} - e^{2\mu_{\mathrm{c}}(T) + \sigma_{\mathrm{c}}^2(T)}, \end{aligned}$$

according to Eq. (6.11). Alternative formulations are

$$\begin{aligned} \mu_{\mathrm{c}}(T) &= 2\ln(1 + \mu(T)) - \frac{\ln(\sigma^2(T) + [\,1 + \mu(T)\,]^2)}{2}, \\ \sigma_{\mathrm{c}}^2(T) &= \ln\left(1 + \left[\frac{\sigma(T)}{1 + \mu(T)}\right]^2\right). \end{aligned}$$

See [646, p. 105].

Exercise 31.1.5: The return rate of each individual asset is $\mu - \sigma^2/2$ (see Exercise 13.3.8(2)). The variance of the portfolio is σ^2/n; the portfolio's return rate is hence $\mu - \sigma^2/(2n)$. Their difference is $\sigma^2/2 - \sigma^2/(2n) = (1 - 1/n)(\sigma^2/2)$. See [623, p. 429].

Exercise 31.1.6: Because all investors trade the same fund of risky assets, trading activity in each stock as a fraction of its shares outstanding is identical across all stocks [614].

Exercise 31.1.7: Replace every variable ω_i with $Y_i - Z_i$ and add the requirements $Y_i \geq 0$ and $Z_i \geq 0$. Under the new model, the equation about means remains a homogeneous linear form in the variables and the equation about covariances remains a homogeneous quadratic form. It is thus a general portfolio selection model. The new model is strictly equivalent to the original model for the following reasons. If $(Y_1, Z_1, Y_2, Z_2, \ldots, Y_n, Z_n)$ is a feasible portfolio for the new model, then $(\omega_1, \omega_2, \ldots, \omega_n)$ with $\omega_i \equiv Y_i - Z_i$ is feasible for the original model with the same mean and standard deviation. Conversely, if $(\omega_1, \omega_2, \ldots, \omega_n)$ is feasible for the original model, then $(Y_1, Z_1, Y_2, Z_2, \ldots, Y_n, Z_n)$ with $Y_i \equiv \max(\omega_i, 0)$ and $Z_i \equiv \max(-\omega_i, 0)$ is feasible for the new model, and with the same mean and standard deviation. See [642, p. 26].

Exercise 31.2.1: It has been shown in the text that the combination of minimum-variance portfolios results in a minimum-variance portfolio. In fact, the combination of efficient portfolios results in an efficient portfolio if the weights applied to the portfolios are all nonnegative. This is because the resulting portfolio is a minimum-variance portfolio with an expected rate of return at or above the MVP's, making it efficient by definition. Because investors hold only efficient portfolios, the market is a combination of efficient portfolios and thus is efficient as well.

Exercise 31.2.2: Let the market return be $r_M = \sum_i \omega'_i r_i$, where ω'_i are market proportions. The fact that every investor holds the same market portfolio implies that the market portfolio satisfies Eq. (31.1), or $\lambda \sum_i \sigma_{ij} \omega'_i = \overline{r}_j - r_{\text{f}}$. Note that λ is independent of the choice of j. Because

$$\mathrm{Cov}[\, r_j, r_M \,] = E\left[\, (r_j - \overline{r}_j) \sum_i \omega'_i (r_i - \overline{r}_i) \,\right] = \sum_i \omega'_i E[\, (r_j - \overline{r}_j)(r_i - \overline{r}_i) \,] = \sum_i \omega'_i \sigma_{ij},$$

we recognize that $\lambda \,\mathrm{Cov}[\, r_j, r_M \,] = \overline{r}_j - r_{\text{f}}$. For the market portfolio, $\lambda \,\mathrm{Cov}[\, r_M, r_M \,] = \overline{r}_M - r_{\text{f}}$, implying $\lambda = (\overline{r}_M - r_{\text{f}})/\sigma_M^2$. See [317, pp. 303–304], [623, p. 177], and [799, pp. 287–289].

Exercise 31.2.3: The risk of an asset that is uncorrelated with the market can be diversified away because purchasing many such assets that are also mutually uncorrelated results in a small total variance [623, p. 179].

Exercise 31.2.4: Such an asset reduces the overall portfolio risk when combined with the market. Investors must pay for this risk-reducing benefit. Another example is insurance. See [424, p. 213], [623, p. 179], and [799, p. 271].

Exercise 31.2.5: They all have the same level of systematic risk, thus beta. The part of the total risk that is specific will not be priced. See [88, p. 300].

Exercise 31.2.6: (1) We need to show that if $P_1 = \frac{\overline{Q}_1}{1+r_{\text{f}}+\beta_1(\overline{r}_M - r_{\text{f}})}$ and $P_2 = \frac{\overline{Q}_2}{1+r_{\text{f}}+\beta_2(\overline{r}_M - r_{\text{f}})}$, then

$$P_1 + P_2 = \frac{\overline{Q}_1 + \overline{Q}_2}{1 + r_{\text{f}} + \beta'(\overline{r}_M - r_{\text{f}})},$$

where β' is the beta of the asset which is the sum of assets 1 and 2. Because both terms within the braces of Eq. (31.5) depend linearly on Q, the claim holds. (2) If otherwise, then we can buy the cheaper portfolio and sell the more expensive portfolio to earn arbitrage profits in a perfect market. See [623, p. 188].

Exercise 31.2.7:

Index Value in a Year	*1200*	*1100*	*1000*	*900*	*800*
Portfolio Value in a Year	1.36	1.16	0.96	0.76	0.56

Exercise 31.2.8: If the index is S, the portfolio will be worth $\$1,000 \times S$. However, the payoff of the options will be

$$10 \times \max(1{,}000 - S, 0) \times 100 \geq 1{,}000{,}000 - 1{,}000 \times S$$

dollars. (Recall that the size of a stock index option is \$100 times the index.) Add them up to obtain a lower bound of \$1,000,000.

Exercise 31.2.10: It is a bull call spread.

Exercise 31.2.11: Because the continuous compounded return $\ln S(t)/S(0)$ is a $(\mu - \sigma^2/2, \sigma)$ Brownian motion by Example 14.3.3, its mean is proportional to t, whereas its volatility or noise is proportional to $\sqrt{t}$. Shorter-term returns (i.e., small t's) are therefore dominated by noise.

Exercise 31.2.12: See [317, p. 359] or [424, pp. 132–135]. Whether riskless assets exist is not essential.

Exercise 31.3.1: $\mathrm{Cov}[\, r_i, r_M \,] = b_i \times \mathrm{Var}[\, r_M \,]$. The CAPM predicts that $\alpha_i = 0$. See [623, p. 205].

Exercise 31.3.2: Consider a k-factor model $r_i = a_i + b_{i1} f_1 + b_{i2} f_2 + \cdots + b_{ik} f_k + \epsilon_i$. Define $\boldsymbol{f} \equiv [\, f_1, f_2, \ldots, f_k \,]^{\text{T}}$ and let C be the covariance matrix of the factors $\boldsymbol{f}$. Assume that $E[\, \boldsymbol{f} \,] = \boldsymbol{0}$ or apply the procedure below to $\boldsymbol{f} - E[\, \boldsymbol{f} \,]$ instead. By Eq. (19.1), there exists a real orthogonal matrix B such that $\Sigma \equiv B^{\text{T}} C B$ is a diagonal matrix. Because $[\, p_1, p_2, \ldots, p_k \,]^{\text{T}} \equiv B^{\text{T}} \boldsymbol{f}$ has the covariance matrix Σ, $\{\, p_1, p_2, \ldots, p_k \,\}$ can be used as the desired set of uncorrelated factors.

Exercise 31.3.3: When the CAPM holds, Eq. (31.7) becomes $\overline{r} = r_{\text{f}} + b_1 \lambda_1$. Equate it with the security market line $\overline{r} - r_{\text{f}} = \beta(\overline{r}_M - r_{\text{f}})$ in Theorem 31.2.1 to obtain $\lambda_1 = \beta(\overline{r}_M - r_{\text{f}})/b_1$. Because $\beta = \mathrm{Cov}[\, r, r_M \,]/\sigma_M^2 = b_1 \times \mathrm{Cov}[\, f, r_M \,]/\sigma_M^2$, we obtain $\lambda_1 = (\mathrm{Cov}[\, f, r_M \,]/\sigma_M^2)(\overline{r}_M - r_{\text{f}})$. See [799, p. 335].

Exercise 31.3.4: Suppose we invest ω_i dollars in asset i, $i = 1, 2, \ldots, n$, in order to satisfy $\sum_{i=1}^{n} \omega_i = 0$, $\sum_{i=1}^{n} \omega_i b_{i1} = 0, \ldots, \sum_{i=1}^{n} \omega_i b_{im} = 0$. This portfolio requires zero net investment and has zero risk. Therefore its expected payoff must be zero, or $\sum_{i=1}^{n} \omega_i \bar{r}_i = 0$. Define $\boldsymbol{\omega} \equiv (\omega_1, \omega_2, \ldots, \omega_n)^{\mathrm{T}}$, $\boldsymbol{b}_j \equiv (b_{1j}, b_{2j}, \ldots, b_{nj})^{\mathrm{T}}$ for $j = 1, 2, \ldots, m$, $\mathbf{1} \equiv (1, 1, \ldots, 1)^{\mathrm{T}}$, and $\bar{\boldsymbol{r}} = (\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_n)^{\mathrm{T}}$. We can restate the conclusion as follows: For any $\boldsymbol{\omega}$ satisfying $\boldsymbol{\omega}^{\mathrm{T}}\mathbf{1} = 0$, $\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{b}_1 = 0, \ldots, \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{b}_m = 0$, it holds that $\boldsymbol{\omega}^{\mathrm{T}}\bar{\boldsymbol{r}} = 0$. That is, any $\boldsymbol{\omega}$ orthogonal to $\mathbf{1}, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$ is also orthogonal to $\bar{\boldsymbol{r}}$. By a standard result in linear algebra, $\bar{\boldsymbol{r}}$ must be a linear combination of the vectors $\mathbf{1}, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$. Thus there are constants $\lambda_0, \lambda_1, \ldots, \lambda_m$ such that $\bar{\boldsymbol{r}} = \lambda_0 \mathbf{1} + \lambda_1 \boldsymbol{b}_1 + \cdots + \lambda_m \boldsymbol{b}_m$.

Exercise 31.3.5: The proof of Exercise 31.3.4 goes through for the well-diversified portfolio as it is riskless.

Exercise 31.4.2: As $\ln S_\tau - \ln S \sim N((\mu - \sigma^2/2)\,\tau, \sigma^2\tau)$ (see Comment 14.4.1), with probability c, the return $\ln S_\tau / S$ is at least $(\mu - \sigma^2/2)\,\tau + N^{-1}(1-c)\,\sigma\sqrt{\tau}$. The desired VaR thus equals $Se^{\mu\tau} - Se^{(\mu-\sigma^2/2)\,\tau + N^{-1}(1-c)\,\sigma\sqrt{\tau}}$. See [8].

Exercise 31.4.3: See Eq. (19.1) for (1) and (2). (3) It follows from

$$[\, dZ_1, dZ_2, \ldots, dZ_n \,]^{\mathrm{T}} = \mathrm{diag}[\, \lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1} \,]\, P^{\mathrm{T}}\, [\, dW_1, dW_2, \ldots, dW_n \,]^{\mathrm{T}}$$

and (2). (4) It results from the construction of the u_is.

NOTE

1. This can be verified formally as follows. Put in matrix terms, Eqs. (33.14) and (33.15) say that $Ak = 0$, where k is a column vector whose jth element is k_j and A is the $(n+1) \times (n+1)$ matrix whose ith row is $[\, \sigma_{i1} f_1, \sigma_{i2} f_2, \ldots, \sigma_{in} f_n \,]$ for $1 \le i \le n$ and whose $(n+1)$th row is $[\, f_1(\mu_1 - r),\ f_2(\mu_2 - r), \ldots,\ f_n(\mu_n - r) \,]$. For this set of equations to have a nontrivial solution, it is necessary that the determinant of A be zero. In particular, its last row must be a linear combination of the first n rows. This implies that $f_j(\mu_j - r) = \sum_i \lambda_i \sigma_{ij} f_j$, $j = 1, 2, \ldots, n+1$, for some $\lambda_1, \lambda_2, \ldots, \lambda_n$, which depend on only $S_1, S_2, \ldots, S_n$, and t.