PRACTICE PROBLEMS

1 The table below gives current information on the interest rates for two two-year and two eight-year maturity investments. The table also gives the maturity, liquidity, and default risk characteristics of a new investment possibility (Investment 3). All investments promise only a single payment (a payment at maturity). Assume that premiums relating to inflation, liquidity, and default risk are constant across all time horizons.

Investment	Maturity (in Years)	Liquidity	Default Risk	Interest Rate (%)
1	2	High	Low	2.0
2	2	Low	Low	2.5
3	7	Low	Low	r_3
4	8	High	Low	4.0
5	8	Low	High	6.5

Based on the information in the above table, address the following:

A Explain the difference between the interest rates on Investment 1 and Investment 2.

B Estimate the default risk premium.

C Calculate upper and lower limits for the interest rate on Investment 3, r_3.

2 A couple plans to set aside $20,000 per year in a conservative portfolio projected to earn 7 percent a year. If they make their first savings contribution one year from now, how much will they have at the end of 20 years?

3 Two years from now, a client will receive the first of three annual payments of $20,000 from a small business project. If she can earn 9 percent annually on her investments and plans to retire in six years, how much will the three business project payments be worth at the time of her retirement?

4 To cover the first year's total college tuition payments for his two children, a father will make a $75,000 payment five years from now. How much will he need to invest today to meet his first tuition goal if the investment earns 6 percent annually?

5 A client can choose between receiving 10 annual $100,000 retirement payments, starting one year from today, or receiving a lump sum today. Knowing that he can invest at a rate of 5 percent annually, he has decided to take the lump sum. What lump sum today will be equivalent to the future annual payments?

6 You are considering investing in two different instruments. The first instrument will pay nothing for three years, but then it will pay $20,000 per year for four years. The second instrument will pay $20,000 for three years and $30,000 in the fourth year. All payments are made at year-end. If your required rate of return on these investments is 8 percent annually, what should you be willing to pay for:

A The first instrument?

B The second instrument (use the formula for a four-year annuity)?

7 Suppose you plan to send your daughter to college in three years. You expect her to earn two-thirds of her tuition payment in scholarship money, so you estimate that your payments will be $10,000 a year for four years. To estimate whether you have set aside enough money, you ignore possible inflation in tuition payments and assume that you can earn 8 percent annually on your investments. How much should you set aside now to cover these payments?

8 A client plans to send a child to college for four years starting 18 years from now. Having set aside money for tuition, she decides to plan for room and board also. She estimates these costs at $20,000 per year, payable at the beginning of each year, by the time her child goes to college. If she starts next year and makes 17 payments into a savings account paying 5 percent annually, what annual payments must she make?

9 A couple plans to pay their child's college tuition for 4 years starting 18 years from now. The current annual cost of college is C$7,000, and they expect this cost to rise at an annual rate of 5 percent. In their planning, they assume that they can earn 6 percent annually. How much must they put aside each year, starting next year, if they plan to make 17 equal payments?

10 The nominal risk-free rate is *best* described as the sum of the real risk-free rate and a premium for:

A maturity.

B liquidity.

C expected inflation.

11 Which of the following risk premiums is most relevant in explaining the difference in yields between 30-year bonds issued by the US Treasury and 30-year bonds issued by a small private issuer?

A Inflation

B Maturity

C Liquidity

12 A bank quotes a stated annual interest rate of 4.00%. If that rate is equal to an effective annual rate of 4.08%, then the bank is compounding interest:

A daily.

B quarterly.

C semiannually.

13 The value in six years of $75,000 invested today at a stated annual interest rate of 7% compounded quarterly is *closest* to:

A $112,555.

B $113,330.

C $113,733.

14 A client requires £100,000 one year from now. If the stated annual rate is 2.50% compounded weekly, the deposit needed today is *closest* to:

A £97,500.

B £97,532.

C £97,561.

15 For a lump sum investment of ¥250,000 invested at a stated annual rate of 3% compounded daily, the number of months needed to grow the sum to ¥1,000,000 is *closest* to:

A 555.

B 563.

C 576.

16 Given a €1,000,000 investment for four years with a stated annual rate of 3% compounded continuously, the difference in its interest earnings compared with the same investment compounded daily is *closest* to:

A €1.

B €6.

C €455.

17 An investment pays €300 annually for five years, with the first payment occurring today. The present value (PV) of the investment discounted at a 4% annual rate is *closest* to:

A €1,336.

B €1,389.

C €1,625.

18 A perpetual preferred stock makes its first quarterly dividend payment of $2.00 in five quarters. If the required annual rate of return is 6% compounded quarterly, the stock's present value is *closest* to:

A $31.

B $126.

C $133.

19 A saver deposits the following amounts in an account paying a stated annual rate of 4%, compounded semiannually:

Year	End of Year Deposits ($)
1	4,000
2	8,000
3	7,000
4	10,000

At the end of Year 4, the value of the account is *closest* to:

A $30,432

B $30,447

C $31,677

20 An investment of €500,000 today that grows to €800,000 after six years has a stated annual interest rate *closest* to:

A 7.5% compounded continuously.

B 7.7% compounded daily.

C 8.0% compounded semiannually.

21 A sweepstakes winner may select either a perpetuity of £2,000 a month beginning with the first payment in one month or an immediate lump sum payment of £350,000. If the annual discount rate is 6% compounded monthly, the present value of the perpetuity is:

A less than the lump sum.

B equal to the lump sum.

C greater than the lump sum.

22 At a 5% interest rate per year compounded annually, the present value (PV) of a 10-year ordinary annuity with annual payments of $2,000 is $15,443.47. The PV of a 10-year annuity due with the same interest rate and payments is *closest* to:

A $14,708.

B $16,216.

C $17,443.

23 Grandparents are funding a newborn's future university tuition costs, estimated at $50,000/year for four years, with the first payment due as a lump sum in 18 years. Assuming a 6% effective annual rate, the required deposit today is *closest* to:

A $60,699.

B $64,341.

C $68,201.

24 The present value (PV) of an investment with the following year-end cash flows (CF) and a 12% required annual rate of return is *closest* to:

Year	Cash Flow (€)
1	100,000
2	150,000
5	–10,000

A €201,747.

B €203,191.

C €227,573.

25 A sports car, purchased for £200,000, is financed for five years at an annual rate of 6% compounded monthly. If the first payment is due in one month, the monthly payment is *closest* to:

A £3,847.

B £3,867.

C £3,957.

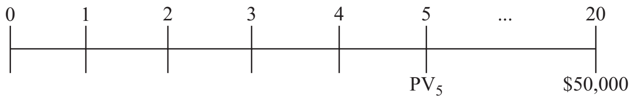
26 Given a stated annual interest rate of 6% compounded quarterly, the level amount that, deposited quarterly, will grow to £25,000 at the end of 10 years is *closest* to:

A £461.

B £474.

C £836.

27 Given the following timeline and a discount rate of 4% a year compounded annually, the present value (PV), as of the end of Year 5 (PV_5), of the cash flow received at the end of Year 20 is *closest* to:



A $22,819.

B $27,763.

C $28,873.

28 A client invests €20,000 in a four-year certificate of deposit (CD) that annually pays interest of 3.5%. The annual CD interest payments are automatically reinvested in a separate savings account at a stated annual interest rate of 2% compounded monthly. At maturity, the value of the combined asset is *closest* to:

A €21,670.

B €22,890.

C €22,950.

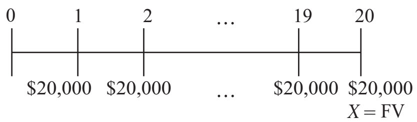
SOLUTIONS

1 A Investment 2 is identical to Investment 1 except that Investment 2 has low liquidity. The difference between the interest rate on Investment 2 and Investment 1 is 0.5 percentage point. This amount represents the liquidity premium, which represents compensation for the risk of loss relative to an investment's fair value if the investment needs to be converted to cash quickly.

B To estimate the default risk premium, find the two investments that have the same maturity but different levels of default risk. Both Investments 4 and 5 have a maturity of eight years. Investment 5, however, has low liquidity and thus bears a liquidity premium. The difference between the interest rates of Investments 5 and 4 is 2.5 percentage points. The liquidity premium is 0.5 percentage point (from Part A). This leaves 2.5 – 0.5 = 2.0 percentage points that must represent a default risk premium reflecting Investment 5's high default risk.

C Investment 3 has liquidity risk and default risk comparable to Investment 2, but with its longer time to maturity, Investment 3 should have a higher maturity premium. The interest rate on Investment 3, r_3, should thus be above 2.5 percent (the interest rate on Investment 2). If the liquidity of Investment 3 were high, Investment 3 would match Investment 4 except for Investment 3's shorter maturity. We would then conclude that Investment 3's interest rate should be less than the interest rate on Investment 4, which is 4 percent. In contrast to Investment 4, however, Investment 3 has low liquidity. It is possible that the interest rate on Investment 3 exceeds that of Investment 4 despite 3's shorter maturity, depending on the relative size of the liquidity and maturity premiums. However, we expect r_3 to be less than 4.5 percent, the expected interest rate on Investment 4 if it had low liquidity. Thus 2.5 percent < r_3 < 4.5 percent.

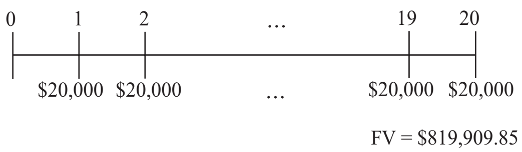
2 i. Draw a time line.



ii. Identify the problem as the future value of an annuity.

iii. Use the formula for the future value of an annuity.

$$\begin{aligned} FV_N &= A\left[\frac{(1+r)^N - 1}{r}\right] \\ &= \$20{,}000\left[\frac{(1+0.07)^{20} - 1}{0.07}\right] \\ &= \$819{,}909.85 \end{aligned}$$



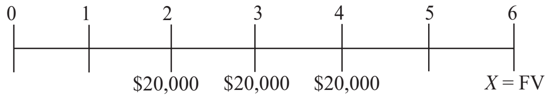
iv. Alternatively, use a financial calculator.

Notation Used on Most Calculators	Numerical Value for This Problem
N	20
$\%i$	7
PV	n/a (= 0)
FV **compute**	X
PMT	$20,000

Enter 20 for N, the number of periods. Enter 7 for the interest rate and 20,000 for the payment size. The present value is not needed, so enter 0. Calculate the future value. Verify that you get $819,909.85 to make sure you have mastered your calculator's keystrokes.

In summary, if the couple sets aside $20,000 each year (starting next year), they will have $819,909.85 in 20 years if they earn 7 percent annually.

3 **i.** Draw a time line.



ii. Recognize the problem as the future value of a delayed annuity. Delaying the payments requires two calculations.

iii. Use the formula for the future value of an annuity (Equation 7).

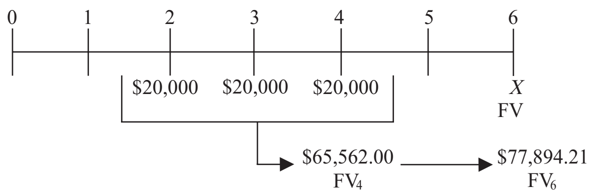
$$\mathrm{FV}_N = A\left[\frac{(1+r)^N - 1}{r}\right]$$

to bring the three $20,000 payments to an equivalent lump sum of $65,562.00 four years from today.

Notation Used on Most Calculators	Numerical Value for This Problem
N	3
$\%i$	9
PV	n/a (= 0)
FV **compute**	X
PMT	$20,000

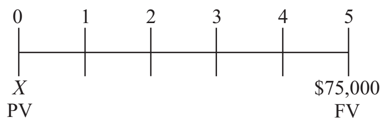
iv. Use the formula for the future value of a lump sum (Equation 2), $\mathrm{FV}_N = \mathrm{PV}(1 + r)^N$, to bring the single lump sum of $65,562.00 to an equivalent lump sum of $77,894.21 six years from today.

Notation Used on Most Calculators	Numerical Value for This Problem
N	2
$\%i$	9
PV	$65,562.00
FV **compute**	X
PMT	n/a (= 0)



In summary, your client will have \$77,894.21 in six years if she receives three yearly payments of \$20,000 starting in Year 2 and can earn 9 percent annually on her investments.

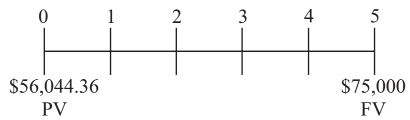
4 **i.** Draw a time line.



ii. Identify the problem as the present value of a lump sum.

iii. Use the formula for the present value of a lump sum.

$$\begin{aligned} \text{PV} &= \text{FV}_N(1+r)^{-N} \\ &= \$75{,}000(1+0.06)^{-5} \\ &= \$56{,}044.36 \end{aligned}$$



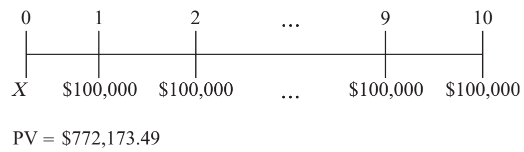
In summary, the father will need to invest \$56,044.36 today in order to have \$75,000 in five years if his investments earn 6 percent annually.

5 **i.** Draw a time line for the 10 annual payments.

ii. Identify the problem as the present value of an annuity.

iii. Use the formula for the present value of an annuity.

$$\begin{aligned} \text{PV} &= A\left[\frac{1-\dfrac{1}{(1+r)^N}}{r}\right] \\ &= \$100{,}000\left[\frac{1-\dfrac{1}{(1+0.05)^{10}}}{0.05}\right] \\ &= \$772{,}173.49 \end{aligned}$$



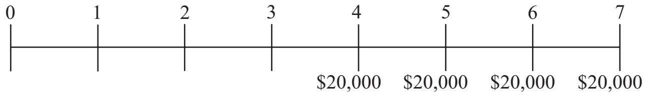
iv. Alternatively, use a financial calculator.

Notation Used on Most Calculators	Numerical Value for This Problem
N	10
$\%i$	5
PV **compute**	X
FV	n/a (= 0)
PMT	$100,000

In summary, the present value of 10 payments of $100,000 is $772,173.49 if the first payment is received in one year and the rate is 5 percent compounded annually. Your client should accept no less than this amount for his lump sum payment.

6 A To evaluate the first instrument, take the following steps:

i. Draw a time line.



ii.

$$PV_3 = A\left[\frac{1-\frac{1}{(1+r)^N}}{r}\right]$$

$$= \$20{,}000\left[\frac{1-\frac{1}{(1+0.08)^4}}{0.08}\right]$$

$$= \$66{,}242.54$$

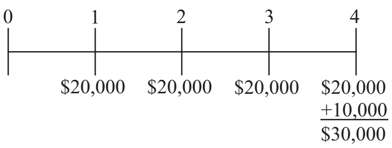
iii.

$$PV_0 = \frac{PV_3}{(1+r)^N} = \frac{\$66{,}242.54}{1.08^3} = \$52{,}585.46$$

You should be willing to pay $52,585.46 for this instrument.

B To evaluate the second instrument, take the following steps:

i. Draw a time line.



The time line shows that this instrument can be analyzed as an ordinary annuity of $20,000 with four payments (valued in Step ii below) and a $10,000 payment to be received at t = 4 (valued in Step iii below).

ii.

$$PV = A\left[\frac{1-\frac{1}{(1+r)^N}}{r}\right]$$
$$= \$20{,}000\left[\frac{1-\frac{1}{(1+0.08)^4}}{0.08}\right]$$
$$= \$66{,}242.54$$

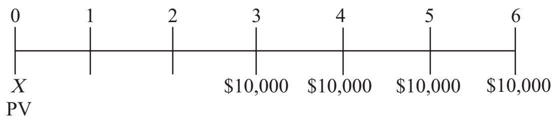
iii.

$$PV = \frac{FV_4}{(1+r)^N} = \frac{\$10{,}000}{(1+0.08)^4} = \$7{,}350.30$$

iv. Total = \$66,242.54 + \$7,350.30 = \$73,592.84

You should be willing to pay \$73,592.84 for this instrument.

7 **i.** Draw a time line.



ii. Recognize the problem as a delayed annuity. Delaying the payments requires two calculations.

iii. Use the formula for the present value of an annuity (Equation 11).

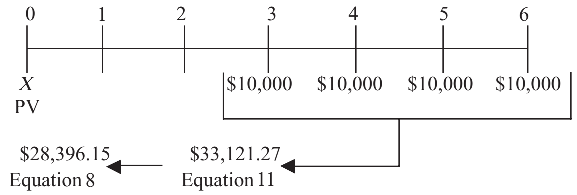
$$PV = A\left[\frac{1-\frac{1}{(1+r)^N}}{r}\right]$$

to bring the four payments of \$10,000 back to a single equivalent lump sum of \$33,121.27 at t = 2. Note that we use t = 2 because the first annuity payment is then one period away, giving an ordinary annuity.

Notation Used on Most Calculators	Numerical Value for This Problem
N	4
$\%i$	8
PV **compute**	X
PMT	\$10,000

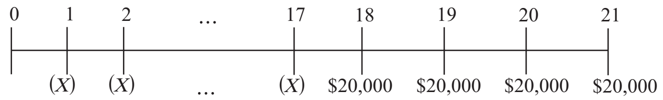
iv. Then use the formula for the present value of a lump sum (Equation 8), PV = $FV_N(1 + r)^{-N}$, to bring back the single payment of \$33,121.27 (at t = 2) to an equivalent single payment of \$28,396.15 (at t = 0).

Notation Used on Most Calculators	Numerical Value for This Problem
N	2
$\%i$	8
PV **compute**	X
FV	\$33,121.27
PMT	n/a (= 0)



In summary, you should set aside $28,396.15 today to cover four payments of $10,000 starting in three years if your investments earn a rate of 8 percent annually.

8 **i.** Draw a time line.



ii. Recognize that you need to equate the values of two annuities.

iii. Equate the value of the four $20,000 payments to a single payment in Period 17 using the formula for the present value of an annuity (Equation 11), with $r = 0.05$. The present value of the college costs as of $t = 17$ is $70,919.

$$\mathrm{PV} = \$20{,}000\left[\frac{1-\dfrac{1}{(1.05)^4}}{0.05}\right] = \$70{,}919$$

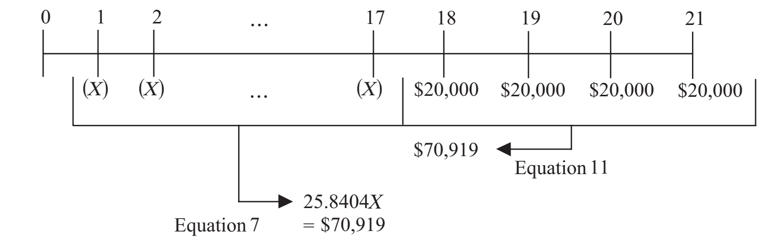
Notation Used on Most Calculators	**Numerical Value for This Problem**
N	4
%*i*	5
PV **compute**	*X*
FV	n/a (= 0)
PMT	$20,000

iv. Equate the value of the 17 investments of X to the amount calculated in Step iii, college costs as of $t = 17$, using the formula for the future value of an annuity (Equation 7). Then solve for X.

$$\$70{,}919 = \left[\frac{(1.05)^{17} - 1}{0.05}\right] = 25.840366X$$

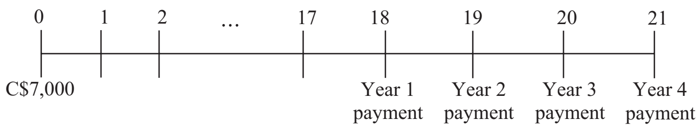
$$X = \$2{,}744.50$$

Notation Used on Most Calculators	**Numerical Value for This Problem**
N	17
%*i*	5
PV	n/a (= 0)
FV	$70,919
PMT **compute**	*X*



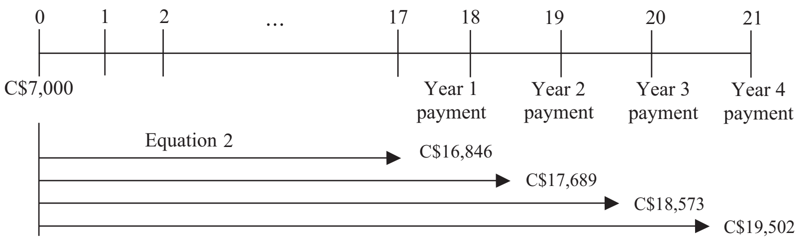
In summary, your client will have to save $2,744.50 each year if she starts next year and makes 17 payments into a savings account paying 5 percent annually.

9 **i.** Draw a time line.



ii. Recognize that the payments in Years 18, 19, 20, and 21 are the future values of a lump sum of C$7,000 in Year 0.

iii. With r = 5%, use the formula for the future value of a lump sum (Equation 2), $FV_N = PV(1 + r)^N$, four times to find the payments. These future values are shown on the time line below.



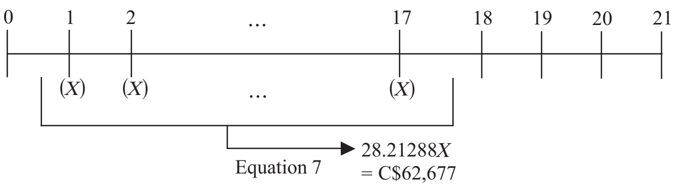
iv. Using the formula for the present value of a lump sum (r = 6%), equate the four college payments to single payments as of t = 17 and add them together. C\$16,846(1.06)$^{-1}$ + C\$17,689(1.06)$^{-2}$ + C\$18,573(1.06)$^{-3}$ + C\$19,502(1.06)$^{-4}$ = C$62,677

v. Equate the sum of C$62,677 at t = 17 to the 17 payments of X, using the formula for the future value of an annuity (Equation 7). Then solve for X.

$$\text{C\$62,677} = X\left[\frac{(1.06)^{17} - 1}{0.06}\right] = 28.21288X$$

$$X = \text{C\$2,221.58}$$

Notation Used on Most Calculators	**Numerical Value for This Problem**
N	17
%i	6
PV	n/a (= 0)
FV	C$62,677
PMT **compute**	X



In summary, the couple will need to put aside C$2,221.58 each year if they start next year and make 17 equal payments.

10 C is correct. The sum of the real risk-free interest rate and the inflation premium is the nominal risk-free rate.

11 C is correct. US Treasury bonds are highly liquid, whereas the bonds of small issuers trade infrequently and the interest rate includes a liquidity premium. This liquidity premium reflects the relatively high costs (including the impact on price) of selling a position.

12 A is correct. The effective annual rate (EAR) when compounded daily is 4.08%.

$$\text{EAR} = (1 + \text{Periodic interest rate})^m - 1$$

$$\text{EAR} = (1 + 0.04/365)^{365} - 1$$

$$\text{EAR} = (1.0408) - 1 = 0.04081 \approx 4.08\%.$$

13 C is correct, as shown in the following (where FV is future value and PV is present value):

$$\text{FV} = \text{PV}\left(1 + \frac{r_s}{m}\right)^{mN}$$

$$\text{FV}_6 = \$75{,}000\left(1 + \frac{0.07}{4}\right)^{(4\times6)}$$

$$\text{FV}_6 = \$113{,}733.21.$$

14 B is correct because £97,531 represents the present value (PV) of £100,000 received one year from today when today's deposit earns a stated annual rate of 2.50% and interest compounds weekly, as shown in the following equation (where FV is future value):

$$\text{PV} = \text{FV}_N\left(1 + \frac{r_s}{m}\right)^{-mN}$$

$$\text{PV} = £100{,}000\left(1 + \frac{0.025}{52}\right)^{-52}$$

$$\text{PV} = £97{,}531.58.$$

15 A is correct. The effective annual rate (EAR) is calculated as follows:

$$\text{EAR} = (1 + \text{Periodic interest rate})^m - 1$$

$$\text{EAR} = (1 + 0.03/365)^{365} - 1$$

$$\text{EAR} = (1.03045) - 1 = 0.030453 \approx 3.0453\%.$$

Solving for N on a financial calculator results in (where FV is future value and PV is present value):

$(1 + 0.030453)^N = FV_N/PV = ¥1,000,000/¥250,000$

= 46.21 years, which multiplied by 12 to convert to months results in 554.5, or ≈ 555 months.

16 B is correct. The difference between continuous compounding and daily compounding is

€127,496.85 – €127,491.29 = €5.56, or ≈ €6, as shown in the following calculations.

With continuous compounding, the investment earns (where PV is present value)

$$\begin{aligned} PVe^{r_s N} - PV &= €1{,}000{,}000e^{0.03(4)} - €1{,}000{,}000 \\ &= €1{,}127{,}496.85 - €1{,}000{,}000 \\ &= €127{,}496.85 \end{aligned}$$

With daily compounding, the investment earns:

$€1,000,000(1 + 0.03/365)^{365(4)}$ – €1,000,000 = €1,127,491.29 – €1,000,000 = €127,491.29.

17 B is correct, as shown in the following calculation for an annuity (A) due:

$$PV = A\left[\frac{1 - \frac{1}{(1+r)^N}}{r}\right](1+r)$$

where A = €300, r = 0.04, and N = 5.

$$PV = €300\left[\frac{1 - \frac{1}{(1+.04)^5}}{.04}\right](1.04)$$

PV = €1,388.97, or ≈ €1,389.

18 B is correct. The value of the perpetuity one year from now is calculated as:

PV = A/r, where PV is present value, A is annuity, and r is expressed as a quarterly required rate of return because the payments are quarterly.

PV = \$2.00/(0.06/4)

PV = \$133.33.

The value today is (where FV is future value)

$PV = FV_N(1 + r)^{-N}$

$PV = \$133.33(1 + 0.015)^{-4}$

PV = \$125.62 ≈ \$126.

19 B is correct. To solve for the future value of unequal cash flows, compute the future value of each payment as of Year 4 at the semiannual rate of 2%, and then sum the individual future values, as follows:

Year	End of Year Deposits ($)	Factor	Future Value ($)
1	4,000	$(1.02)^6$	4,504.65
2	8,000	$(1.02)^4$	8,659.46
3	7,000	$(1.02)^2$	7,282.80
4	10,000	$(1.02)^0$	10,000.00
		Sum =	30,446.91

20 C is correct, as shown in the following (where FV is future value and PV is present value):

If:

$$FV_N = PV\left(1 + \frac{r_s}{m}\right)^{mN}$$

Then:

$$\left(\frac{FV_N}{PV}\right)^{\frac{1}{mN}} - 1 = \frac{r_s}{m}$$

$$\left(\frac{800{,}000}{500{,}000}\right)^{\frac{1}{2\times 6}} - 1 = \frac{r_s}{2}$$

r_s = 0.07988 (rounded to 8.0%).

21 C is correct. As shown below, the present value (PV) of a £2,000 per month perpetuity is worth approximately £400,000 at a 6% annual rate compounded monthly. Thus, the present value of the annuity (A) is worth more than the lump sum offer.

A = £2,000

r = (6%/12) = 0.005

PV = (A/r)

PV = (£2,000/0.005)

PV = £400,000

22 B is correct.

The present value of a 10-year annuity (A) due with payments of $2,000 at a 5% discount rate is calculated as follows:

$$PV = A\left[\frac{1 - \frac{1}{(1+r)^N}}{r}\right] + \$2{,}000$$

$$PV = \$2{,}000\left[\frac{1 - \dfrac{1}{(1 + 0.05)^9}}{0.05}\right] + \$2{,}000$$

PV = $16,215.64.

Alternatively, the PV of a 10-year annuity due is simply the PV of the ordinary annuity multiplied by 1.05:

PV = $15,443.47 × 1.05

PV = $16,215.64.

23 B is correct. First, find the present value (PV) of an ordinary annuity in Year 17 that represents the tuition costs:

$$\$50{,}000\left[\frac{1 - \dfrac{1}{(1 + 0.06)^4}}{0.06}\right]$$

= $50,000 × 3.4651

= $173,255.28.

Then, find the PV of the annuity in today's dollars (where FV is future value):

$$PV_0 = \frac{FV}{(1 + 0.06)^{17}}$$

$$PV_0 = \frac{\$173{,}255.28}{(1 + 0.06)^{17}}$$

PV_0 = $64,340.85 ≈ $64,341.

24 B is correct, as shown in the following table.

Year	Cash Flow (€)	Formula CF × (1 + *r*)^t	PV at Year 0
1	100,000	$100{,}000(1.12)^{-1}$ =	89,285.71
2	150,000	$150{,}000(1.12)^{-2}$ =	119,579.08
5	–10,000	$-10{,}000(1.12)^{-5}$ =	–5,674.27
			203,190.52

25 B is correct, calculated as follows (where A is annuity and PV is present value):

$$A = (\text{PV of annuity}) \Bigg/ \left[\frac{1 - \dfrac{1}{(1 + r_s/m)^{mN}}}{r_s/m} \right]$$

$$= (£200{,}000) \Bigg/ \left[\frac{1 - \dfrac{1}{(1 + r_s/m)^{mN}}}{r_s/m} \right]$$

$$(£200{,}000) \Bigg/ \left[\frac{1 - \dfrac{1}{(1 + 0.06/12)^{12(5)}}}{0.06/12} \right]$$

$$= (£200{,}000)/51.72556$$

$$= £3{,}866.56$$

26 A is correct. To solve for an annuity (A) payment, when the future value (FV), interest rate, and number of periods is known, use the following equation:

$$\text{FV} = A \left[\frac{\left(1 + \dfrac{r_s}{m}\right)^{mN} - 1}{\dfrac{r}{m}} \right]$$

$$£25{,}000 = A \left[\frac{\left(1 + \dfrac{0.06}{4}\right)^{4 \times 10} - 1}{\dfrac{0.06}{4}} \right]$$

$A = £460.68$

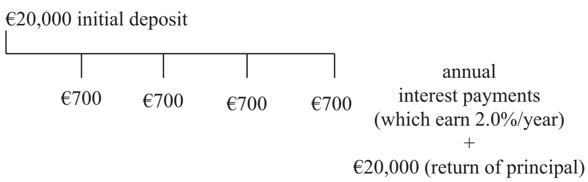
27 B is correct. The PV in Year 5 of a \$50,000 lump sum paid in Year 20 is \$27,763.23 (where FV is future value):

$\text{PV} = \text{FV}_N(1 + r)^{-N}$

$\text{PV} = \$50{,}000(1 + 0.04)^{-15}$

$\text{PV} = \$27{,}763.23$

28 B is correct, as the following cash flows show:



The four annual interest payments are based on the CD's 3.5% annual rate.

The first payment grows at 2.0% compounded monthly for three years (where FV is future value):

$$\mathrm{FV}_N = €700\left(1+\frac{0.02}{12}\right)^{3\times 12}$$

$$\mathrm{FV}_N = 743.25$$

The second payment grows at 2.0% compounded monthly for two years:

$$\mathrm{FV}_N = €700\left(1+\frac{0.02}{12}\right)^{2\times 12}$$

$$\mathrm{FV}_N = 728.54$$

The third payment grows at 2.0% compounded monthly for one year:

$$\mathrm{FV}_N = €700\left(1+\frac{0.02}{12}\right)^{1\times 12}$$

$$\mathrm{FV}_N = 714.13$$

The fourth payment is paid at the end of Year 4. Its future value is €700.

The sum of all future value payments is as follows:

€20,000.00	CD
€743.25	First payment's *FV*
€728.54	Second payment's *FV*
€714.13	Third payment's *FV*
€700.00	Fourth payment's *FV*
€22,885.92	Total *FV*

PRACTICE PROBLEMS

1 The net present value (NPV) of an investment is equal to the sum of the expected cash flows discounted at the:

A internal rate of return.

B risk-free rate.

C opportunity cost of capital.

2 A $2.2 million investment will result in the cash flows shown below:

Year	Year-End Cash Flow (millions)
1	$1.3
2	$1.6
3	$1.9
4	$0.8

Using an 8% opportunity cost of capital, the project's net present value (NPV) is *closest* to:

A $2.47 million.

B $3.40 million.

C $4.67 million.

3 A firm is considering three projects as shown below.

	Net Present Value (NPV)	Internal Rate of Return (IRR)	Hurdle Rate
Project A	$47,000	10%	5%
Project B	$58,000	20%	12%
Project C	$52,000	22%	12%

If the firm can only accept one project, to maximize shareholder wealth, the firm is *most likely* to select:

A Project A.

B Project B.

C Project C.

4 The internal rate of return (IRR) is *best* described as the:

A opportunity cost of capital.

B time-weighted rate of return.

C discount rate that makes the net present value equal to zero.

5 A three-year investment requires an initial outlay of £1,000. It is expected to provide three year-end cash flows of £200 plus a net salvage value of £700 at the end of three years. Its internal rate of return (IRR) is *closest* to:

A 10%.

B 11%.

C 20%.

6 The internal rate of return (IRR) rule indicates acceptance of a project when the IRR is:

A greater than zero.

B less than the opportunity cost of capital.

C greater than the opportunity cost of capital.

7 Suppose a company has only €1,000,000 available to invest. The three projects available are described in the table:

Year	Project A Cash Flow (€)	Project B Cash Flow (€)	Project C Cash Flow (€)
0	–1,000,000	–1,000,000	–500,000
1	1,200,000	0	0
2	0	0	0
3	0	1,600,000	850,000
Internal rate of return (IRR)	20.00%	16.96%	19.35%

If the opportunity cost of capital is 12%, which project should be accepted?

A Project A.

B Project B.

C Project C.

8 An investor buys a share of stock for \$52.68 and receives an \$0.88 dividend one year later. If the share sells for \$57.50 just after the dividend payment, the holding period return is *closest* to:

A 9.1%.

B 9.9%.

C 10.8%.

9 An investor performs the following transactions on the shares of a firm.

- At t = 0, she purchases a share for \$1,000.
- At t = 1, she receives a dividend of \$25 and then purchases three additional shares for \$1,055 each.
- At t = 2, she receives a total dividend of \$100 and then sells the four shares for \$1,100 each.

The money-weighted rate of return is *closest* to:

A 4.5%.

B 6.9%.

C 7.3%.

10 A fund receives investments at the beginning of each year and generates returns as shown in the table.

Year of Investment	Amount of Investment	Return during Year of Investment
1	\$1,000	15%
2	\$4,000	14%
3	\$45,000	–4%

Which return measure over the three-year period is negative?

A Geometric mean return

B Time-weighted rate of return

C Money-weighted rate of return

11 At the beginning of Year 1, a fund has \$10 million under management; it earns a return of 14% for the year. The fund attracts another \$100 million at the start of Year 2 and earns a return of 8% for that year. The money-weighted rate of return is *most likely*:

A less than the time-weighted rate of return.

B the same as the time-weighted rate of return.

C greater than the time-weighted rate of return.

12 An investor buys a bond for \$980. After six months, she collects a semiannual coupon of \$30 and sells the bond for \$990. Her six-month holding period yield (HPY) is *closest* to:

A 1.0%.

B 4.1%.

C 8.3%.

13 A portfolio manager pays \$99,500 for a 182-day US T-bill with face value of \$100,000. The T-bill will be held to maturity. A yield of 0.5025% calculated for this T-bill when it is purchased is *most* accurately described as the:

A bank discount yield.

B money market yield.

C holding period yield.

14 A 123-day T-bill with a maturity value of \$100,000 is priced at \$99,620. The bill's effective annual yield is *closest* to:

A 0.38%.

B 1.12%.

C 1.14%.

15 For a T-bill purchased at \$97,000 that matures at \$100,000 in 300 days, which of the following yields is *closest* to 3.71%?

A Money market yield

B Holding period yield (HPY)

C Effective annual yield

16 A 223-day T-bill with a maturity value of \$100,000 has a bank discount yield of 2.05%. The bill's holding period yield is *closest* to:

A 1.29%.

B 2.08%.

C 2.11%.

17 Given a 300-day holding period yield (HPY) of 7%, the effective annual yield (EAY) is *closest* to:

A 8.4%.

B 8.5%.

C 8.6%.

SOLUTIONS

1 C is correct. The NPV sums the project's expected cash flows (CF) discounted at the opportunity cost of capital. The NPV calculation is

$$\text{NPV} = \sum_{t=0}^{N} \frac{\text{CF}_t}{(1+r)^t}$$

where

CF_t = the expected net cash flow at time t
N = the investment's projected life
r = the discount rate or opportunity cost of capital

2 A is correct.

$$\text{The NPV} = -\$2.2 + \frac{\$1.3}{(1.08)} + \frac{\$1.6}{(1.08)^2} + \frac{\$1.9}{(1.08)^3} + \frac{\$0.8}{(1.08)^4} = \$2.47\text{ million.}$$

3 B is correct. According to the NPV rule, shareholder wealth is maximized by selecting a project with the highest NPV. The IRR rule also signals acceptance; however, the IRR rule should not be used to rank projects. In this case, Project B adds the most value to the firm.

4 C is correct. The internal rate of return is computed by identifying all cash flows and solving for the rate that makes the net present value of those cash flows equal to zero.

5 B is correct. IRR is determined by setting the net present value equal to zero for the cash flows shown in the table.

Year	Cash Flow (£)
0	–1,000
1	200
2	200
3	900

6 C is correct. The IRR investment decision rule states, "Accept projects or investments for which the IRR is greater than the opportunity cost of capital."

7 B is correct. The projects are mutually exclusive because the amount to invest is constrained to €1,000,000. Therefore, the net present value (NPV) rule should be used to choose among them when the IRR rule and NPV rule conflict. Based on the opportunity cost of capital of 12%, the NPV of Project B is €138,848, which is higher than the NPV of €71,429 for Project A and the NPV of $105,013 for Project C.

Project A NPV = –€1,000,000 + €1,200,000/(1.12) = €71,429

Project B NPV = –€1,000,000 + €1,600,000/$(1.12)^3$ = €138,848

Project C NPV = –€500,000 + €850,000/$(1.12)^3$ = €105,013

8 C is correct. The formula for the holding period return is

$$\frac{P_1 - P_0 + D_1}{P_0}$$

where P_0 is the initial investment, P_1 is the final value at the end of the holding period, and D_1 is the cash dividend paid at the end of the holding period. This investment results in a holding period return (HPR) of

$$\text{HPR} = \frac{\$57.50 - \$52.68 + \$0.88}{\$52.68} = 10.82\%$$

9 B is correct. Computation of the money-weighted return, r, requires finding the discount rate that sets the present value (outflows) equal to the present value (inflows).

Solving for r,

$$\$1{,}000 + \frac{\$3{,}165}{(1+r)} = \frac{\$25}{(1+r)} + \frac{\$4{,}500}{(1+r)^2}$$

results in a value of r = 6.91%

10 C is correct. The money-weighted rate of return considers both the timing and amounts of investments into the fund. The investment at the beginning of Year 1 will be worth \$1,000(1.15)(1.14)(0.96) = \$1,258.56 at the end of Year 3. The investment made at the beginning of Year 2 will be worth \$4,377.60 = \$4,000(1.14)(0.96) at the end of Year 3. The investment of \$45,000 at the beginning of Year 3 decreases to a value of \$45,000(0.96) = \$43,200 at the end of Year 3.

Solving for r,

$$\$1{,}000 + \frac{\$4{,}000}{(1+r)} + \frac{\$45{,}000}{(1+r)^2} = \frac{\$1{,}258 + \$4{,}377.60 + \$43{,}200}{(1+r)^3}$$

results in r = –2.08%

Note that B is incorrect because the time-weighted rate of return (TWR) of the fund is the same as the geometric mean return of the fund and is thus positive:

$$\text{TWR} = \sqrt[3]{(1.15)(1.14)(0.96)} - 1 = 7.97\%$$

11 A is correct. The money-weighted rate of return is found by setting the present value (PV) of investments into the fund equal to the PV of the fund's terminal value. Because most of the investment came during Year 2, the measure will be biased toward the performance of Year 2. Set the PV of investments equal to the PV of the fund's terminal value:

$$\$10 + \frac{100}{(1+r)} = \frac{10 \times 1.14 \times 1.08 + 100 \times 1.08}{(1+r)^2}$$

Solving for r results in r = 8.53%.

The time-weighted return of the fund is = $\sqrt[2]{(1.14)(1.08)} - 1 = 10.96\%$.

12 B is correct. The HPY for the bond is

$$\frac{P_1 - P_0 + D_1}{P_0} = \frac{\$990 - \$980 + \$30}{\$980} = 4.08\%$$

13 C is correct. The 182-day holding period yield (HPY) is calculated as follows:

$$\text{HPY} = \frac{P_1 - P_0 + D_1}{P_0} = \frac{\$100{,}000 - \$99{,}500}{\$99{,}500} = 0.5025\%$$

14 C is correct. The effective annual yield (EAY) is calculated as follows (where HPY is holding period yield):

$$\text{HPY} = \frac{P_1 - P_0 + D_1}{P_0} = \frac{\$100{,}00 - \$99{,}620}{\$99{,}620} = 0.3814\%$$

$$\text{EAY} = (1 + \text{HPY})^{365/123} - 1 = (1 + .003814)^{365/123} - 1 = 1.1362\%$$

15 A is correct. The money market yield is equal to the annualized HPY, assuming a 360-day year. r_{MM} = (HPY)(360/t). A T-bill purchased at \$97,000 has a HPY of [(\$100,000 – \$97,000)/\$97,000] = 3.09%. The money market yield for this T-bill is (3.09%)(360/300) = 3.71%. The effective annual yield for this T-bill is $(1 + 0.0309)^{365/300} - 1$ = 3.77%.

16 A is correct. The holding period yield of a bill can be computed from the bank discount yield by finding the bill's discount D:

$$r_{BD} = \frac{D}{F} \times \frac{360}{t}, \text{ so } 0.0205 = \frac{D}{\$100{,}000} \times \frac{360}{223}$$

Solving for D results in a discount of \$1269.86. This result implies a purchase price of \$100,000 – \$1,269.86 = \$98,730.14.

The holding period yield (HPY) then computes as:

$$\text{HPY} = \frac{\$1{,}269.86}{\$98{,}730.14} = 1.286\%$$

B is incorrect because it is the money market yield and C is incorrect because it is the equivalent annual yield.

17 C is correct. The EAY is one plus the HPY, compounded forward one year, and then subtract one:

$$\text{EAY} = (1 + \text{HPY})^{365/t} - 1$$

$$(1 + 0.07)^{365/300} - 1 = 8.58\% \approx 8.6\%$$

PRACTICE PROBLEMS

1 Which of the following groups *best* illustrates a sample?

A The set of all estimates for Exxon Mobil's FY2015 EPS

B The FTSE Eurotop 100 as a representation of the European stock market

C UK shares traded on 13 August 2015 that also closed above £120/share on the London Stock Exchange

2 Published ratings on stocks ranging from 1 (strong sell) to 5 (strong buy) are examples of which measurement scale?

A Ordinal

B Interval

C Nominal

3 In descriptive statistics, an example of a parameter is the:

A median of a population.

B mean of a sample of observations.

C standard deviation of a sample of observations.

4 A mutual fund has the return frequency distribution shown in the following table.

Return Interval (%)	Absolute Frequency
–10.0 to –7.0	3
–7.0 to –4.0	7
–4.0 to –1.0	10
–1.0 to +2.0	12
+2.0 to +5.0	23
+5.0 to +8.0	5

Which of the following statements is correct?

A The relative frequency of the interval "–1.0 to +2.0" is 20%.

B The relative frequency of the interval "+2.0 to +5.0" is 23%.

C The cumulative relative frequency of the interval "+5.0 to +8.0" is 91.7%.

5 An analyst is using the data in the following table to prepare a statistical report.

Portfolio's Deviations from Benchmark Return, 2003–2014 (%)			
2003	2.48	**2009**	–9.19
2004	–2.59	**2010**	–5.11
2005	9.47	**2011**	1.33
2006	–0.55	**2012**	6.84
2007	–1.69	**2013**	3.04
2008	–0.89	**2014**	4.72

The cumulative relative frequency for the interval $-1.71\% \leq x < 2.03\%$ is *closest* to:

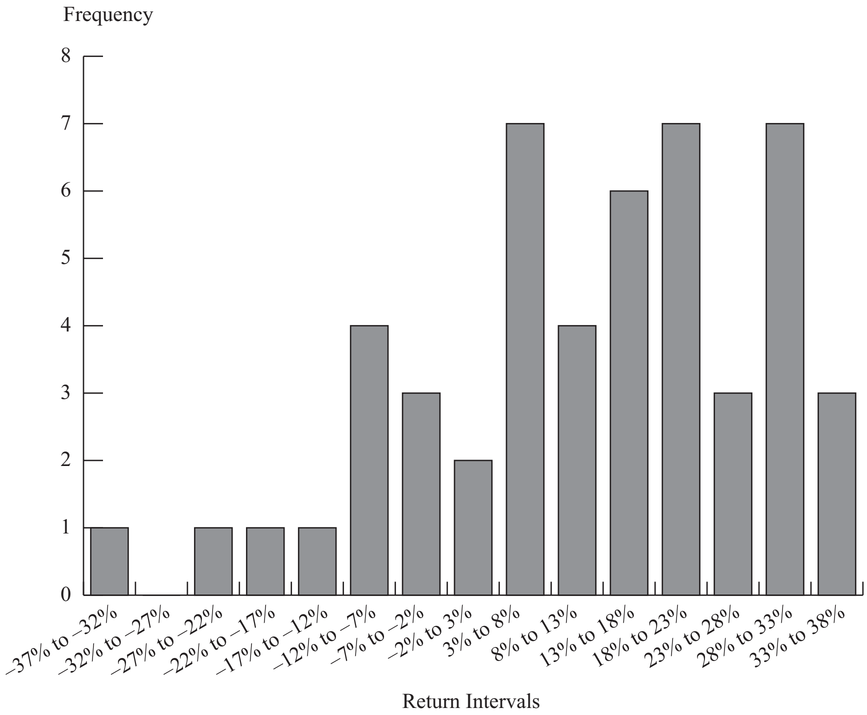
A 0.250.

B 0.333.

C 0.583.

The following information relates to Questions 6–7

The following histogram shows a distribution of the S&P 500 Index annual returns from 1964 to 2013:



6 The interval containing the median return is:

A 3% to 8%.

B 8% to 13%.

C 13% to 18%.

7 Based on the previous histogram, the distribution is *best* described as having:

A one mode.

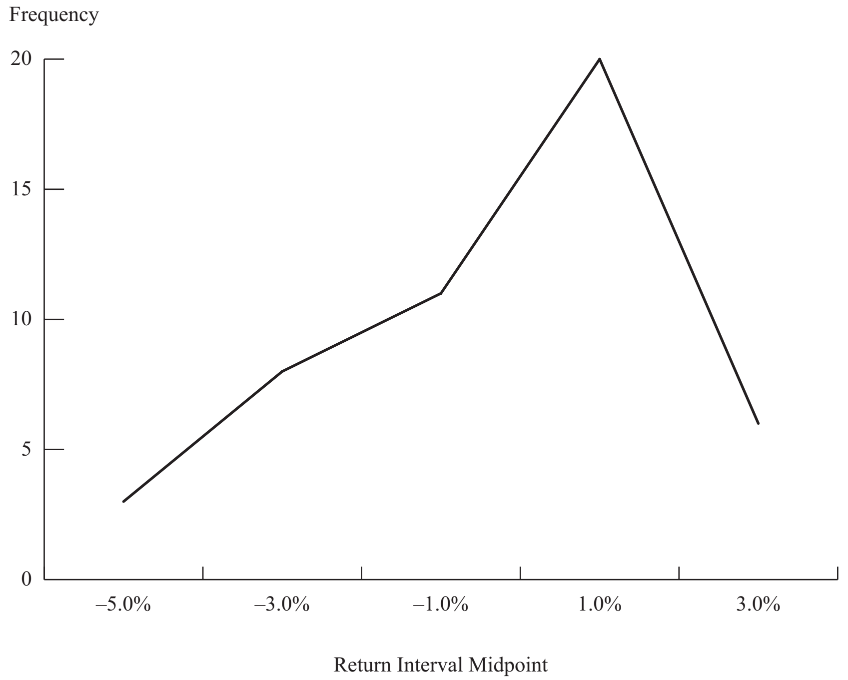
B two modes.

C three modes.

---

8 The following is a frequency polygon of monthly exchange rate changes in the US dollar/Japanese yen spot exchange rate from January 2010 to December 2013. A positive change represents yen appreciation (the yen buys more dollars), and a negative change represents yen depreciation (the yen buys fewer dollars).

Monthly Changes in the US Dollar/Japanese Yen Spot Exchange Rate



Based on the chart, yen appreciation:

A occurred more than 50% of the time.

B was less frequent than yen depreciation.

C in the 0.0 to 2.0 interval occurred 20% of the time.

9 The annual returns for three portfolios are shown in the following table. Portfolios P and R were created in 2009, Portfolio Q in 2010.

	Annual Portfolio Returns (%)				
	2009	**2010**	**2011**	**2012**	**2013**
Portfolio P	–3.0	4.0	5.0	3.0	7.0
Portfolio Q		–3.0	6.0	4.0	8.0
Portfolio R	1.0	–1.0	4.0	4.0	3.0

The median annual return from portfolio creation to 2013 for:

A Portfolio P is 4.5%.

B Portfolio Q is 4.0%.

C Portfolio R is higher than its arithmetic mean annual return.

10 In 2015, an investor allocated his retirement savings in the asset classes shown in the following table.

Asset Class	**Asset Allocation (%)**	**Asset Class Return (%)**
Large-cap US equities	20.0	8.0
Small-cap US equities	40.0	12.0
Emerging market equities	25.0	–3.0
High-yield bonds	15.0	4.0

The portfolio return in 2015 is *closest to*:

A 5.1%.

B 5.3%.

C 6.3%.

11 The following table shows the annual returns for Fund Y.

	Fund Y (%)
2010	19.5
2011	–1.9
2012	19.7
2013	35.0
2014	5.7

The geometric mean for Fund Y is *closest* to:

A 14.9%.

B 15.6%.

C 19.5%.

12 A manager invests €5,000 annually in a security for four years at the prices shown in the following table.

	Purchase Price of Security (€)
Year 1	62.00
Year 2	76.00
Year 3	84.00
Year 4	90.00

The average price paid for the security is *closest* to:

A €76.48.

B €77.26.

C €78.00.

The following information relates to Questions 13–14

The following table shows the annual MSCI World Index total returns for 2004–2013.

2004	15.25%	**2009**	30.79%
2005	10.02%	**2010**	12.34%
2006	20.65%	**2011**	–5.02%
2007	9.57%	**2012**	16.54%
2008	–40.33%	**2013**	27.37%

13 The fourth quintile return for the MSCI World Index is *closest* to:

A 20.65%.

B 26.03%.

C 27.37%.

14 For 2009–2013, the mean absolute deviation of the MSCI World Index total returns is *closest* to:

A 10.20%.

B 12.74%.

C 16.40%.

---

15 Annual returns and summary statistics for three funds are listed in the following table:

	Annual Returns (%)		
Year	**Fund ABC**	**Fund XYZ**	**Fund PQR**
2009	−20.0	−33.0	−14.0
2010	23.0	−12.0	−18.0
2011	−14.0	−12.0	6.0
2012	5.0	−8.0	−2.0
2013	−14.0	11.0	3.0
Mean	−4.0	−10.8	−5.0
Standard deviation	17.8	15.6	10.5

The fund that shows the highest dispersion is:

A Fund PQR if the measure of dispersion is the range.

B Fund XYZ if the measure of dispersion is the variance.

C Fund ABC if the measure of dispersion is the mean absolute deviation.

16 Over the past 240 months, an investor's portfolio had a mean monthly return of 0.79%, with a standard deviation of monthly returns of 1.16%. According to Chebyshev's inequality, the minimum number of the 240 monthly returns that fall into the range of −0.95% to 2.53% is *closest* to:

A 80.

B 107.

C 133.

17 The mean monthly return and the standard deviation for three industry sectors are shown in the following table.

Sector	**Mean Monthly Return (%)**	**Standard Deviation of Return (%)**
Utilities (UTIL)	2.10	1.23
Materials (MATR)	1.25	1.35
Industrials (INDU)	3.01	1.52

Based on the coefficient of variation, the riskiest sector is:

A utilities.

B materials.

C industrials.

18 Three equity fund managers have performance records summarized in the following table:

	Mean Annual Return (%)	Standard Deviation of Return (%)
Manager 1	14.38	10.53
Manager 2	9.25	6.35
Manager 3	13.10	8.23

Given a risk-free rate of return of 2.60%, which manager performed best based on the Sharpe ratio?

A Manager 1

B Manager 2

C Manager 3

The following information relates to Questions 19–21

The following table shows various statistics for Portfolios 1, 2, and 3.

	Mean Return (%)	Standard Deviation of Returns (%)	Skewness	Excess Kurtosis
Portfolio 1	7.8	15.1	0.0	0.7
Portfolio 2	10.2	20.5	0.9	–1.8
Portfolio 3	12.9	29.3	–1.5	6.2

19 An investment adviser bases his allocation on the Sharpe ratio. Assuming a risk-free rate of 1.5%, which portfolio is he *most likely* to recommend?

A Portfolio 1

B Portfolio 2

C Portfolio 3

20 The skewness of Portfolio 1 indicates its mean return is *most likely*:

A less than its median.

B equal to its median.

C greater than its median.

21 Compared with a normal distribution, the distribution of returns for Portfolio 3 *most likely*:

A has less weight in the tails.

B has a greater number of extreme returns.

C has fewer small deviations from its mean.

---

22 Two portfolios have unimodal return distributions. Portfolio 1 has a skewness of 0.77, and Portfolio 2 has a skewness of –1.11.

Which of the following is correct?

A For Portfolio 1, the median is less than the mean.

B For Portfolio 1, the mode is greater than the mean.

C For Portfolio 2, the mean is greater than the median.

23 When analyzing investment returns, which of the following statements is correct?

A The geometric mean will exceed the arithmetic mean for a series with non-zero variance.

B The geometric mean measures an investment's compound rate of growth over multiple periods.

C The arithmetic mean accurately estimates an investment's terminal value over multiple periods.

SOLUTIONS

1 B is correct. The FTSE Eurotop 100 represents a sample of all European stocks. It is a subset of the population of all European stocks.

2 A is correct. Ordinal scales sort data into categories that are ordered with respect to some characteristic and may involve numbers to identify categories but do not assure that the differences between scale values are equal. The buy rating scale indicates that a stock ranked 5 is expected to perform better than a stock ranked 4, but it tells us nothing about the performance difference between stocks ranked 4 and 5 compared with the performance difference between stocks ranked 1 and 2, and so on.

3 A is correct. Any descriptive measure of a population characteristic is referred to as a parameter.

4 A is correct. The relative frequency is the absolute frequency of each interval divided by the total number of observations. Here, the relative frequency is calculated as: (12/60) × 100 = 20%. B is incorrect because the relative frequency of this interval is (23/60) × 100 = 38.33%. C is incorrect because the cumulative relative frequency of the last interval must equal 100%.

5 C is correct. The cumulative relative frequency of an interval identifies the fraction of observations that are less than the upper limit of the given interval. It is determined by summing the relative frequencies from the lowest interval up to and including the given interval. The following table shows the relative frequencies for all the intervals of the data from the previous table:

Lower Limit (%)	Upper Limit (%)	Absolute Frequency	Relative Frequency	Cumulative Relative Frequency
−9.19 ≤	< −5.45	1	0.083	0.083
−5.45 ≤	< −1.71	2	0.167	0.250
−1.71 ≤	< 2.03	4	0.333	0.583
2.03 ≤	< 5.77	3	0.250	0.833
5.77 ≤	≥ 9.51	2	0.167	1.000

The interval −1.71% ≤ x < 2.03% has a cumulative relative frequency of 0.583.

6 C is correct. Because there are 50 data points in the histogram, the median return would be the mean of the 50/2 = 25th and (50 + 2)/2 = 26th positions. The sum of the return interval frequencies to the left of the 13% to 18% interval is 24. As a result, the 25th and 26th returns will fall in the 13% to 18% interval.

7 C is correct. The mode of a distribution with data grouped in intervals is the interval with the highest frequency. The three intervals of 3% to 8%, 18% to 23%, and 28% to 33% all have a high frequency of 7.

8 A is correct. Twenty observations lie in the interval "0.0 to 2.0," and six observations lie in the 2.0 to 4.0 interval. Together, they represent 26/48, or 54.17% of all observations, which is more than 50%.

9 C is correct. The median of Portfolio R is 0.8% higher than the mean for Portfolio R.

10 C is correct. The portfolio return must be calculated as the weighted mean return, where the weights are the allocations in each asset class:

$$(0.20 \times 8\%) + (0.40 \times 12\%) + (0.25 \times -3\%) + (0.15 \times 4\%) = 6.25\%, \text{ or } \approx 6.3\%.$$

11 A is correct. The geometric mean return for Fund Y is found as follows:

$$\text{Fund Y} = [(1 + 0.195) \times (1 - 0.019) \times (1 + 0.197) \times (1 + 0.350) \times (1 + 0.057)]^{(1/5)} - 1$$
$$= 14.9\%.$$

12 A is correct. The harmonic mean is appropriate for determining the average price per unit. It is calculated by summing the reciprocals of the prices; then averaging that sum by dividing by the number of prices; and finally, taking the reciprocal of the average:

$$4/[(1/62.00) + (1/76.00) + (1/84.00) + (1/90.00)] = €76.48.$$

13 B is correct. Quintiles divide a distribution into fifths, with the fourth quintile occurring at the point at which 80% of the observations lie below it. The fourth quintile is equivalent to the 80th percentile. To find the yth percentile (P_y), we first must determine its location. The formula for the location (L_y) of a yth percentile in an array with n entries sorted in ascending order is $L_y = (n + 1) \times (y/100)$. In this case, $n = 10$ and $y = 80\%$, so

$$L_{80} = (10 + 1) \times (80/100) = 11 \times 0.8 = 8.8.$$

With the data arranged in ascending order (–40.33%, –5.02%, 9.57%, 10.02%, 12.34%, 15.25%, 16.54%, 20.65%, 27.37%, and 30.79%), the 8.8th position would be between the 8th and 9th entries, 20.65% and 27.37%, respectively. Using linear interpolation, $P_{80} = X_8 + (L_y - 8) \times (X_9 - X_8)$,

$$\begin{aligned} P_{80} &= 20.65 + (8.8 - 8) \times (27.37 - 20.65) \\ &= 20.65 + (0.8 \times 6.72) = 20.65 + 5.38 \\ &= 26.03\%. \end{aligned}$$

14 A is correct. The formula for mean absolute deviation (MAD) is

$$\text{MAD} = \frac{\sum_{i=1}^{n} \left| X_i - \bar{X} \right|}{n}$$

Column 1: Sum annual returns and divide by n to find the arithmetic mean $\left(\bar{X}\right)$ of 16.40%.

Column 2: Calculate the absolute value of the difference between each year's return and the mean from Column 1. Sum the results and divide by n to find the MAD.

These calculations are shown in the following table:

	Column 1		Column 2
Year	**Return**		$\left\lvert X_i - \bar{X} \right\rvert$
2009	30.79%		14.39%
2010	12.34%		4.06%
2011	–5.02%		21.42%
2012	16.54%		0.14%
2013	27.37%		10.97%
Sum:	82.02%	Sum:	50.98%

	Column 1		Column 2
Year	Return		$\lvert X_i - \bar{X} \rvert$
n:	5	n:	5
$\bar{X}$:	16.40%	MAD:	10.20%

15 C is correct. The mean absolute deviation (MAD) of Fund ABC's returns is greater than the MAD of both of the other funds.

$$\text{MAD} = \frac{\sum_{i=1}^{n} \lvert X_i - \bar{X} \rvert}{n}, \text{ where } \bar{X} \text{ is the arithmetic mean of the series.}$$

MAD for Fund ABC =

$$\frac{\lvert -20-(-4)\rvert + \lvert 23-(-4)\rvert + \lvert -14-(-4)\rvert + \lvert 5-(-4)\rvert + \lvert -14-(-4)\rvert}{5} = 14.4\%$$

MAD for Fund XYZ =

$$\frac{\lvert -33-(-10.8)\rvert + \lvert -12-(-10.8)\rvert + \lvert -12-(-10.8)\rvert + \lvert -8-(-10.8)\rvert + \lvert 11-(-10.8)\rvert}{5} = 9.8\%$$

MAD for Fund PQR =

$$\frac{\lvert -14-(-5)\rvert + \lvert -18-(-5)\rvert + \lvert 6-(-5)\rvert + \lvert -2-(-5)\rvert + \lvert 3-(-5)\rvert}{5} = 8.8\%$$

A and B are incorrect because the range and variance of the three funds are as follows:

	Fund ABC	Fund XYZ	Fund PQR
Range	43%	44%	24%
Variance	317	243	110

The numbers shown for variance are understood to be in "percent squared" terms so that when taking the square root, the result is standard deviation in percentage terms. Alternatively, by expressing standard deviation and variance in decimal form, one can avoid the issue of units; in decimal form, the variances for Fund ABC, Fund XYZ, and Fund PQR are 0.0317, 0.0243, and 0.0110, respectively.

16 C is correct. According to Chebyshev's inequality, the proportion of the observations within k standard deviations of the arithmetic mean is at least $1 - 1/k^2$ for all $k > 1$.

The upper limit of the range is 2.53%, which is 2.53 – 0.79 = 1.74% above the mean. The lower limit is –0.95, which is 0.79 – (–0.95) = 1.74% below the mean. As a result, k = 1.74/1.16 = 1.50 standard deviations.

Because k = 1.50, the proportion of observations within the interval is at least $1 - 1/1.5^2 = 1 - 0.444 = 0.556$, or 55.6%. Thus, the number of observations in the given range is at least 240 × 55.6%, which is ≈ 133.

17 B is correct. The coefficient of variation (CV) is the ratio of the standard deviation to the mean, where a higher CV implies greater risk per unit of return.

$$CV_{UTIL} = \frac{s}{\overline{X}} = \frac{1.23\%}{2.10\%} = 0.59$$

$$CV_{MATR} = \frac{s}{\overline{X}} = \frac{1.35\%}{1.25\%} = 1.08$$

$$CV_{INDU} = \frac{s}{\overline{X}} = \frac{1.52\%}{3.01\%} = 0.51$$

18 C is correct. The Sharpe ratio (S) is the mean excess portfolio return per unit of risk, where a higher Sharpe ratio indicates better performance:

$$S_1 = \frac{\overline{R}_p - \overline{R}_F}{s_p} = \frac{14.38 - 2.60}{10.53} = 1.12$$

$$S_2 = \frac{\overline{R}_p - \overline{R}_F}{s_p} = \frac{9.25 - 2.60}{6.35} = 1.05$$

$$S_3 = \frac{\overline{R}_p - \overline{R}_F}{s_p} = \frac{13.10 - 2.60}{8.23} = 1.28$$

19 B is correct. The Sharpe ratio measures a portfolio's return per unit of risk and is defined as $S_1 = \frac{\overline{R}_p - \overline{R}_F}{s_p}$, where $\overline{R}_p$ is the mean return for the portfolio, $\overline{R}_F$ is the mean return to a risk-free asset, and s_p is the standard deviation of return on the portfolio. The Sharpe ratios for the three portfolios are as follows:

Portfolio 1 = (7.8 – 1.5)/15.1 = 6.3/15.1 = 0.417

Portfolio 2 = (10.2 – 1.5)/20.5 = 8.7/20.5 = 0.424

Portfolio 3 = (12.9 – 1.5)/29.3 = 11.4/29.3 = 0.389

So Portfolio 2 has the highest return per unit of risk.

20 B is correct. Portfolio 1 has a skewness of 0.0, which indicates that the portfolio's return distribution is symmetrical and thus its mean and median are equal.

21 B is correct. Portfolio 3 has positive excess kurtosis (i.e., kurtosis greater than 3), which indicates that its return distribution is leptokurtic and has fatter tails than the normal. The fatter tails mean Portfolio 3 has a greater number of extreme returns.

22 A is correct. Portfolio 1 is positively skewed, so the mean is greater than the median, which is greater than the mode.

23 B is correct. The geometric mean compounds the periodic returns of every period, giving the investor a more accurate measure of the terminal value of an investment.

READING

9

Probability Concepts

by Richard A. DeFusco, PhD, CFA, Dennis W. McLeavey, CFA, Jerald E. Pinto, PhD, CFA, and David E. Runkle, PhD, CFA

Richard A. DeFusco, PhD, CFA, is at the University of Nebraska-Lincoln (USA). Dennis W. McLeavey, CFA, is at the University of Rhode Island (USA). Jerald E. Pinto, PhD, CFA, is at CFA Institute (USA). David E. Runkle, PhD, CFA, is at Trilogy Global Advisors (USA).

LEARNING OUTCOMES

Mastery	*The candidate should be able to:*
☐	**a.** define a random variable, an outcome, an event, mutually exclusive events, and exhaustive events;
☐	**b.** state the two defining properties of probability and distinguish among empirical, subjective, and a priori probabilities;
☐	**c.** state the probability of an event in terms of odds for and against the event;
☐	**d.** distinguish between unconditional and conditional probabilities;
☐	**e.** explain the multiplication, addition, and total probability rules;
☐	**f.** calculate and interpret 1) the joint probability of two events, 2) the probability that at least one of two events will occur, given the probability of each and the joint probability of the two events, and 3) a joint probability of any number of independent events;
☐	**g.** distinguish between dependent and independent events;
☐	**h.** calculate and interpret an unconditional probability using the total probability rule;
☐	**i.** explain the use of conditional expectation in investment applications;
☐	**j.** explain the use of a tree diagram to represent an investment problem;
☐	**k.** calculate and interpret covariance and correlation;
☐	**l.** calculate and interpret the expected value, variance, and standard deviation of a random variable and of returns on a portfolio;

(continued)

LEARNING OUTCOMES

Mastery	The candidate should be able to:
☐	**m.** calculate and interpret covariance given a joint probability function;
☐	**n.** calculate and interpret an updated probability using Bayes' formula;
☐	**o.** identify the most appropriate method to solve a particular counting problem and solve counting problems using factorial, combination, and permutation concepts.

1 INTRODUCTION

All investment decisions are made in an environment of risk. The tools that allow us to make decisions with consistency and logic in this setting come under the heading of probability. This reading presents the essential probability tools needed to frame and address many real-world problems involving risk. We illustrate how these tools apply to such issues as predicting investment manager performance, forecasting financial variables, and pricing bonds so that they fairly compensate bondholders for default risk. Our focus is practical. We explore in detail the concepts that are most important to investment research and practice. One such concept is independence, as it relates to the predictability of returns and financial variables. Another is expectation, as analysts continually look to the future in their analyses and decisions. Analysts and investors must also cope with variability. We present variance, or dispersion around expectation, as a risk concept important in investments. The reader will acquire specific skills in using portfolio expected return and variance.

The basic tools of probability, including expected value and variance, are set out in Section 2 of this reading. Section 3 introduces covariance and correlation (measures of relatedness between random quantities) and the principles for calculating portfolio expected return and variance. Two topics end the reading: Bayes' formula and outcome counting. Bayes' formula is a procedure for updating beliefs based on new information. In several areas, including a widely used option-pricing model, the calculation of probabilities involves defining and counting outcomes. The reading ends with a discussion of principles and shortcuts for counting.

2 PROBABILITY, EXPECTED VALUE, AND VARIANCE

The probability concepts and tools necessary for most of an analyst's work are relatively few and simple but require thought to apply. This section presents the essentials for working with probability, expectation, and variance, drawing on examples from equity and fixed income analysis.

An investor's concerns center on returns. The return on a risky asset is an example of a **random variable**, a quantity whose **outcomes** (possible values) are uncertain. For example, a portfolio may have a return objective of 10 percent a year. The portfolio manager's focus at the moment may be on the likelihood of earning a return that is less than 10 percent over the next year. Ten percent is a particular value or outcome of the random variable "portfolio return." Although we may be concerned about a single outcome, frequently our interest may be in a set of outcomes: The concept of "event" covers both.

- **Definition of Event.** An **event** is a specified set of outcomes.

We may specify an event to be a single outcome—for example, *the portfolio earns a return of 10 percent.* (We use italics to highlight statements that define events.) We can capture the portfolio manager's concerns by defining the event as *the portfolio earns a return below 10 percent.* This second event, referring as it does to all possible returns greater than or equal to −100 percent (the worst possible return) but less than 10 percent, contains an infinite number of outcomes. To save words, it is common to use a capital letter in italics to represent a defined event. We could define A = *the portfolio earns a return of 10 percent* and B = *the portfolio earns a return below 10 percent.*

To return to the portfolio manager's concern, how likely is it that the portfolio will earn a return below 10 percent?

The answer to this question is a **probability**: a number between 0 and 1 that measures the chance that a stated event will occur. If the probability is 0.40 that the portfolio earns a return below 10 percent, there is a 40 percent chance of that event happening. If an event is impossible, it has a probability of 0. If an event is certain to happen, it has a probability of 1. If an event is impossible or a sure thing, it is not random at all. So, 0 and 1 bracket all the possible values of a probability.

Probability has two properties, which together constitute its definition.

- **Definition of Probability.** The two defining properties of a probability are as follows:
  1. The probability of any event E is a number between 0 and 1: $0 \le P(E) \le 1$.
  2. The sum of the probabilities of any set of mutually exclusive and **exhaustive** events equals 1.

P followed by parentheses stands for "the probability of (the event in parentheses)," as in $P(E)$ for "the probability of event E." We can also think of P as a rule or function that assigns numerical values to events consistent with Properties 1 and 2.

In the above definition, the term mutually exclusive means that only one event can occur at a time; **exhaustive** means that the events cover all possible outcomes. The events A = *the portfolio earns a return of 10 percent* and B = *the portfolio earns a return below 10 percent* are mutually exclusive because A and B cannot both occur at the same time. For example, a return of 8.1 percent means that B has occurred and A has not occurred. Although events A and B are mutually exclusive, they are not exhaustive because they do not cover outcomes such as a return of 11 percent. Suppose we define a third event: C = *the portfolio earns a return above 10 percent.* Clearly, A, B, and C are mutually exclusive and exhaustive events. Each of $P(A)$, $P(B)$, and $P(C)$ is a number between 0 and 1, and $P(A) + P(B) + P(C) = 1$.

The most basic kind of mutually exclusive and exhaustive events is the set of all the distinct possible outcomes of the random variable. If we know both that set and the assignment of probabilities to those outcomes—the probability distribution of the random variable—we have a complete description of the random variable, and we can assign a probability to any event that we might describe.[1] The probability of any event is the sum of the probabilities of the distinct outcomes included in the definition of the event. Suppose the event of interest is D = *the portfolio earns a return above the risk-free rate,* and we know the probability distribution of portfolio returns. Assume

1 In the reading on common probability distributions, we describe some of the probability distributions most frequently used in investment applications.

the risk-free rate is 4 percent. To calculate $P(D)$, the probability of D, we would sum the probabilities of the outcomes that satisfy the definition of the event; that is, we would sum the probabilities of portfolio returns greater than 4 percent.

Earlier, to illustrate a concept, we assumed a probability of 0.40 for a portfolio earning less than 10 percent, without justifying the particular assumption. We also talked about using a probability distribution of outcomes to calculate the probability of events, without explaining how a probability distribution might be estimated. Making actual financial decisions using inaccurate probabilities might have grave consequences. How, in practice, do we estimate probabilities? This topic is a field of study in itself, but there are three broad approaches to estimating probabilities. In investments, we often estimate the probability of an event as a relative frequency of occurrence based on historical data. This method produces an **empirical probability**. For example, Thanatawee (2013) reports that of his sample of 1,927 yearly observations for nonfinancial SET (Stock Exchange of Thailand) firms during the years 2002 to 2010, 1,382 were dividend paying firms and 545 were non dividend paying firms. The empirical probability of a Thai firm paying a dividend is thus 1,382/1,927 = 0.72, approximately. We will point out empirical probabilities in several places as they appear in this reading.

Relationships must be stable through time for empirical probabilities to be accurate. We cannot calculate an empirical probability of an event not in the historical record or a reliable empirical probability for a very rare event. There are cases, then, in which we may adjust an empirical probability to account for perceptions of changing relationships. In other cases, we have no empirical probability to use at all. We may also make a personal assessment of probability without reference to any particular data. Each of these three types of probability is a **subjective probability**, one drawing on personal or subjective judgment. Subjective probabilities are of great importance in investments. Investors, in making buy and sell decisions that determine asset prices, often draw on subjective probabilities. Subjective probabilities appear in various places in this reading, notably in our discussion of Bayes' formula.

In a more narrow range of well-defined problems, we can sometimes deduce probabilities by reasoning about the problem. The resulting probability is an **a priori probability**, one based on logical analysis rather than on observation or personal judgment. We will use this type of probability in Example 6. The counting methods we discuss later are particularly important in calculating an a priori probability. Because a priori and empirical probabilities generally do not vary from person to person, they are often grouped as **objective probabilities**.

In business and elsewhere, we often encounter probabilities stated in terms of odds—for instance, "the odds for E" or the "odds against E." For example, as of November 2013, analysts' fiscal year 2014 EPS forecasts for JetBlue Airways ranged from \$0.55 to \$0.69. Suppose one analyst asserts that the odds for the company beating the highest estimate, \$0.69, are 1 to 7. Suppose a second analyst argues that the odds against that happening are 15 to 1. What do those statements imply about the probability of the company's EPS beating the highest estimate? We interpret probabilities stated in terms of odds as follows:

- **Probability Stated as Odds.** Given a probability $P(E)$,
  1. Odds for $E = P(E)/[1 - P(E)]$. The odds for E are the probability of E divided by 1 minus the probability of E. Given odds for E of "a to b," the implied probability of E is $a/(a + b)$.

In the example, the statement that the odds for *the company's EPS for FY2014 beating \$0.69* are 1 to 7 means that the speaker believes the probability of the event is $1/(1 + 7) = 1/8 = 0.125$.

2 Odds against $E = [1 - P(E)]/P(E)$, the reciprocal of odds for E. Given odds against E of "a to b," the implied probability of E is $b/(a + b)$.

The statement that the odds against *the company's EPS for FY2014 beating $0.69* are 15 to 1 is consistent with a belief that the probability of the event is 1/(1 + 15) = 1/16 = 0.0625.

To further explain odds for an event, if $P(E) = 1/8$, the odds for E are (1/8)/(7/8) = (1/8)(8/7) = 1/7, or "1 to 7." For each occurrence of E, we expect seven cases of non-occurrence; out of eight cases in total, therefore, we expect E to happen once, and the probability of E is 1/8. In wagering, it is common to speak in terms of the odds against something, as in Statement 2. For odds of "15 to 1" against E (an implied probability of E of 1/16), a $1 wager on E, if successful, returns $15 in profits plus the $1 staked in the wager. We can calculate the bet's anticipated profit as follows:

Win: Probability = 1/16; Profit =$15

Loss: Probability = 15/16; Profit =−$1

Anticipated profit = (1/16)($15) + (15/16)(−$1) = $0

Weighting each of the wager's two outcomes by the respective probability of the outcome, if the odds (probabilities) are accurate, the anticipated profit of the bet is $0.

EXAMPLE 1

Profiting from Inconsistent Probabilities

You are examining the common stock of two companies in the same industry in which an important antitrust decision will be announced next week. The first company, SmithCo Corporation, will benefit from a governmental decision that there is no antitrust obstacle related to a merger in which it is involved. You believe that SmithCo's share price reflects a 0.85 probability of such a decision. A second company, Selbert Corporation, will equally benefit from a "go ahead" ruling. Surprisingly, you believe Selbert stock reflects only a 0.50 probability of a favorable decision. Assuming your analysis is correct, what investment strategy would profit from this pricing discrepancy?

Consider the logical possibilities. One is that the probability of 0.50 reflected in Selbert's share price is accurate. In that case, Selbert is fairly valued but SmithCo is overvalued, as its current share price overestimates the probability of a "go ahead" decision. The second possibility is that the probability of 0.85 is accurate. In that case, SmithCo shares are fairly valued, but Selbert shares, which build in a lower probability of a favorable decision, are undervalued. You diagram the situation as shown in Table 1.

Table 1 Worksheet for Investment Problem

	True Probability of a "Go Ahead" Decision	
	0.50	**0.85**
SmithCo	Shares Overvalued	Shares Fairly Valued
Selbert	Shares Fairly Valued	Shares Undervalued

The 0.50 probability column shows that Selbert shares are a better value than SmithCo shares. Selbert shares are also a better value if a 0.85 probability is accurate. Thus SmithCo shares are overvalued relative to Selbert shares.

Your investment actions depend on your confidence in your analysis and on any investment constraints you face (such as constraints on selling stock short).[2] A conservative strategy would be to buy Selbert shares and reduce or eliminate any current position in SmithCo. The most aggressive strategy is to short SmithCo stock (relatively overvalued) and simultaneously buy the stock of Selbert (relatively undervalued). This strategy is known as **pairs arbitrage trade**: a trade in two closely related stocks involving the short sale of one and the purchase of the other.

The prices of SmithCo and Selbert shares reflect probabilities that are not **consistent**. According to one of the most important probability results for investments, the **Dutch Book Theorem**,[3] inconsistent probabilities create profit opportunities. In our example, investors, by their buy and sell decisions to exploit the inconsistent probabilities, should eliminate the profit opportunity and inconsistency.

To understand the meaning of a probability in investment contexts, we need to distinguish between two types of probability: unconditional and conditional. Both unconditional and conditional probabilities satisfy the definition of probability stated earlier, but they are calculated or estimated differently and have different interpretations. They provide answers to different questions.

The probability in answer to the straightforward question "What is the probability of this event A?" is an **unconditional probability**, denoted $P(A)$. Unconditional probability is also frequently referred to as **marginal probability**.[4]

Suppose the question is "What is the probability that *the stock earns a return above the risk-free rate* (event A)?" The answer is an unconditional probability that can be viewed as the ratio of two quantities. The numerator is the sum of the probabilities of stock returns above the risk-free rate. Suppose that sum is 0.70. The denominator is 1, the sum of the probabilities of all possible returns. The answer to the question is $P(A) = 0.70$.

Contrast the question "What is the probability of A?" with the question "What is the probability of A, given that B has occurred?" The probability in answer to this last question is a **conditional probability**, denoted $P(A \mid B)$ (read: "the probability of A given B").

Suppose we want to know the probability that *the stock earns a return above the risk-free rate* (event A), given that *the stock earns a positive return* (event B). With the words "given that," we are restricting returns to those larger than 0 percent—a new element in contrast to the question that brought forth an unconditional probability. The conditional probability is calculated as the ratio of two quantities. The numerator is the sum of the probabilities of stock returns above the risk-free rate; in this particular case, the numerator is the same as it was in the unconditional case, which we gave as 0.70. The denominator, however, changes from 1 to the sum of the probabilities for all outcomes (returns) above 0 percent. Suppose that number is 0.80, a

2 *Selling short* or *shorting stock* means selling borrowed shares in the hope of repurchasing them later at a lower price.

3 The theorem's name comes from the terminology of wagering. Suppose someone places a \$100 bet on X at odds of 10 to 1 against X, and later he is able to place a \$600 bet against X at odds of 1 to 1 against X. Whatever the outcome of X, that person makes a riskless profit (equal to \$400 if X occurs or \$500 if X does not occur) because the implied probabilities are inconsistent. He is said to have made a *Dutch book* in X. Ramsey (1931) presented the problem of inconsistent probabilities. See also Lo (1999).

4 In analyses of probabilities presented in tables, unconditional probabilities usually appear at the ends or *margins* of the table, hence the term *marginal probability*. Because of possible confusion with the way *marginal* is used in economics (roughly meaning *incremental*), we use the term *unconditional probability* throughout this discussion.

larger number than 0.70 because returns between 0 and the risk-free rate have some positive probability of occurring. Then $P(A \mid B) = 0.70/0.80 = 0.875$. If we observe that the stock earns a positive return, the probability of a return above the risk-free rate is greater than the unconditional probability, which is the probability of the event given no other information. The result is intuitive.[5] To review, an unconditional probability is the probability of an event without any restriction; it might even be thought of as a stand-alone probability. A conditional probability, in contrast, is a probability of an event given that another event has occurred.

In discussing approaches to calculating probability, we gave one empirical estimate of the probability that a change in dividends is a dividend decrease. That probability was an unconditional probability. Given additional information on company characteristics, could an investor refine that estimate? Investors continually seek an information edge that will help improve their forecasts. In mathematical terms, they are attempting to frame their view of the future using probabilities conditioned on relevant information or events. Investors do not ignore useful information; they adjust their probabilities to reflect it. Thus, the concepts of conditional probability (which we analyze in more detail below), as well as related concepts discussed further on, are extremely important in investment analysis and financial markets.

To state an exact definition of conditional probability, we first need to introduce the concept of joint probability. Suppose we ask the question "What is the probability of both A and B happening?" The answer to this question is a **joint probability**, denoted $P(AB)$ (read: "the probability of A and B"). If we think of the probability of A and the probability of B as sets built of the outcomes of one or more random variables, the joint probability of A and B is the sum of the probabilities of the outcomes they have in common. For example, consider two events: *the stock earns a return above the risk-free rate* (A) and *the stock earns a positive return* (B). The outcomes of A are contained within (a subset of) the outcomes of B, so $P(AB)$ equals $P(A)$. We can now state a formal definition of conditional probability that provides a formula for calculating it.

- **Definition of Conditional Probability.** The conditional probability of A given that B has occurred is equal to the joint probability of A and B divided by the probability of B (assumed not to equal 0).

$$P(A \mid B) = P(AB)/P(B),\ P(B) \neq 0 \tag{1}$$

Sometimes we know the conditional probability $P(A \mid B)$ and we want to know the joint probability $P(AB)$. We can obtain the joint probability from the following **multiplication rule for probabilities**, which is Equation 1 rearranged.

- **Multiplication Rule for Probability.** The joint probability of A and B can be expressed as

$$P(AB) = P(A \mid B)P(B) \tag{2}$$

5 In this example, the conditional probability is greater than the unconditional probability. The conditional probability of an event may, however, be greater than, equal to, or less than the unconditional probability, depending on the facts. For instance, the probability that *the stock earns a return above the risk-free rate* given that *the stock earns a negative return* is 0.

EXAMPLE 2

Conditional Probabilities and Predictability of Mutual Fund Performance (1)

Vidal-Garcia (2013) examined whether historical performance predicts future performance for a sample of mutual funds that included 1,050 actively managed equity funds in six European countries during the period of 1988 through 2010. Funds were classified into nine investment styles based on combinations of investment focus (growth, blend, and value) and fund's market capitalization (small, mid, and large cap). One approach Vidal-Garcia used involved calculating each fund's annual benchmark-adjusted return by subtracting a benchmark return from the annual return of the fund. MSCI (Morgan Stanley Capital International) style indexes were used as benchmarks. For each style of fund in each country, funds were classified as winners or losers for each of two consecutive years. The top 50 percent of funds by benchmark-adjusted return for a given year were labeled winners; the bottom 50 percent were labeled losers. An excerpt from the results of the study for 135 French funds classified as large value funds is given in Table 2. It shows the percentage of those funds that were winners in two consecutive years, winner in one year and then loser in the next year, losers then winners, and losers in both years. The winner–winner entry, for example, shows that 65.5% of the first-year winner funds were also winners in the second year. Note that the four entries in the table can be viewed as conditional probabilities.

Table 2 Persistence of Returns for Large Value Funds in France: 1988 through 2010

	Year 2 Winner	Year 2 Loser
Year 1 winner	65.5%	34.5%
Year 1 loser	15.5%	84.5%

Source: Vidal-Garcia (2013), Table 4.

Based on the data in Table 2, answer the following questions:

1 State the four events needed to define the four conditional probabilities.

2 State the four entries of the table as conditional probabilities using the form P(*this event* | *that event*) = number.

3 Are the conditional probabilities in Part 2 empirical, a priori, or subjective probabilities?

4 Using information in the table, calculate the probability of the event a *fund is a loser in both Year 1 and Year 2.* (Note that because 50 percent of funds are categorized as losers in each year, the unconditional probability that a fund is labeled a loser in either year is 0.5.)

Solution to 1:

The four events needed to define the conditional probabilities are as follows:

Fund is a Year 1 winner

Fund is a Year 1 loser

Fund is a Year 2 loser

Fund is a Year 2 winner

Solution to 2:

From Row 1:

P(*fund is a Year 2 winner* | *fund is a Year 1 winner*) = 0.655

P(*fund is a Year 2 loser* | *fund is a Year 1 winner*) = 0.345

From Row 2:

P(*fund is a Year 2 winner* | *fund is a Year 1 loser*) = 0.155

P(*fund is a Year 2 loser* | *fund is a Year 1 loser*) = 0.845

Solution to 3:

These probabilities are calculated from data, so they are empirical probabilities.

Solution to 4:

The estimated probability is 0.423. With A the event that a *fund is a Year 2 loser* and B the event that a *fund is a Year 1 loser, AB* is the event that a *fund is a loser in both Year 1 and Year 2.* From Table 2, $P(A \mid B) = 0.845$ and $P(B) = 0.50$. Thus, using Equation 2, we find that

$$P(AB) = P(A \mid B)P(B) = 0.845(0.50) = 0.4225$$

or a probability of approximately 0.423.

Equation 2 states that the joint probability of A and B equals the probability of A given B times the probability of B. Because $P(AB) = P(BA)$, the expression $P(AB) = P(BA) = P(B \mid A)P(A)$ is equivalent to Equation 2.

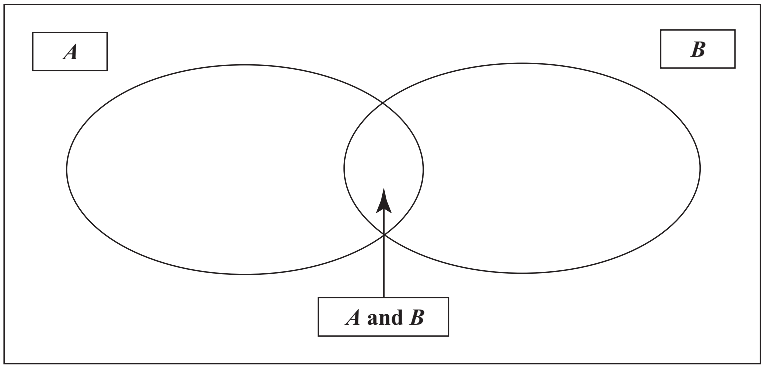
When we have two events, A and B, that we are interested in, we often want to know the probability that either A or B occurs. Here the word "or" is inclusive, meaning that either A or B occurs or that both A and B occur. Put another way, the probability of A or B is the probability that at least one of the two events occurs. Such probabilities are calculated using the **addition rule for probabilities**.

- **Addition Rule for Probabilities.** Given events A and B, the probability that A or B occurs, or both occur, is equal to the probability that A occurs, plus the probability that B occurs, minus the probability that both A and B occur.

$$P(A \text{ or } B) = P(A) + P(B) - P(AB) \tag{3}$$

If we think of the individual probabilities of A and B as sets built of outcomes of one or more random variables, the first step in calculating the probability of A or B is to sum the probabilities of the outcomes in A to obtain $P(A)$. If A and B share any outcomes, then if we now added $P(B)$ to $P(A)$, we would count twice the probabilities of those shared outcomes. So we add to $P(A)$ the quantity $[P(B) - P(AB)]$, which is the probability of outcomes in B net of the probability of any outcomes already counted when we computed $P(A)$. Figure 1 illustrates this process; we avoid double-counting the outcomes in the intersection of A and B by subtracting $P(AB)$. As an example of the calculation, if $P(A) = 0.50$, $P(B) = 0.40$, and $P(AB) = 0.20$, then $P(A \text{ or } B) = 0.50 + 0.40 - 0.20 = 0.70$. Only if the two events A and B were mutually exclusive, so that $P(AB) = 0$, would it be correct to state that $P(A \text{ or } B) = P(A) + P(B)$.

Figure 1 Addition Rule for Probabilities



The next example shows how much useful information can be obtained using the few probability rules presented to this point.

EXAMPLE 3

Probability of a Limit Order Executing

You have two buy limit orders outstanding on the same stock. A limit order to buy stock at a stated price is an order to buy at that price or lower. A number of vendors, including an internet service that you use, supply the estimated probability that a limit order will be filled within a stated time horizon, given the current stock price and the price limit. One buy order (Order 1) was placed at a price limit of \$10. The probability that it will execute within one hour is 0.35. The second buy order (Order 2) was placed at a price limit of \$9.75; it has a 0.25 probability of executing within the same one-hour time frame.

1 What is the probability that either Order 1 or Order 2 will execute?

2 What is the probability that Order 2 executes, given that Order 1 executes?

Solution to 1:

The probability is 0.35. The two probabilities that are given are $P(\textit{Order 1 executes}) = 0.35$ and $P(\textit{Order 2 executes}) = 0.25$. Note that if Order 2 executes, it is certain that Order 1 also executes because the price must pass through \$10 to reach \$9.75. Thus,

$$P(\textit{Order 1 executes} \mid \textit{Order 2 executes}) = 1$$

and

$$P(\textit{Order 1 executes} \text{ and } \textit{Order 2 executes}) = P(\textit{Order 1 executes} \mid \textit{Order 2 executes})P(\textit{Order 2 executes}) = 1(0.25) = 0.25$$

To answer the question, we use the addition rule for probabilities:

$$P(\textit{Order 1 executes} \text{ or } \textit{Order 2 executes}) = P(\textit{Order 1 executes}) + P(\textit{Order 2 executes}) - P(\textit{Order 1 executes} \text{ and } \textit{Order 2 executes}) = 0.35 + 0.25 - 0.25 = 0.35$$

Note that the outcomes for which Order 2 executes are a subset of the outcomes for which Order 1 executes. After you count the probability that Order 1 executes, you have counted the probability of the outcomes for which Order 2 also executes. Therefore, the answer to the question is the probability that Order 1 executes, 0.35.

Solution to 2:

If the first order executes, the probability that the second order executes is 0.714. In the solution to Part 1, you found that $P(\textit{Order 1 executes} \text{ and } \textit{Order 2 executes}) = P(\textit{Order 1 executes} \mid \textit{Order 2 executes})P(\textit{Order 2 executes}) = 1(0.25) = 0.25$. An equivalent way to state this joint probability is useful here:

$$P(\textit{Order 1 executes} \text{ and } \textit{Order 2 executes}) = 0.25$$
$$= P(\textit{Order 2 executes} \mid \textit{Order 1 executes})P(\textit{Order 1 executes})$$

Because $P(\textit{Order 1 executes}) = 0.35$ was a given, you have one equation with one unknown:

$$0.25 = P(\textit{Order 2 executes} \mid \textit{Order 1 executes})(0.35)$$

You conclude that $P(\textit{Order 2 executes} \mid \textit{Order 1 executes}) = 0.25/0.35 = 5/7$, or about 0.714. You can also use Equation 1 to obtain this answer.

Of great interest to investment analysts are the concepts of independence and dependence. These concepts bear on such basic investment questions as which financial variables are useful for investment analysis, whether asset returns can be predicted, and whether superior investment managers can be selected based on their past records.

Two events are independent if the occurrence of one event does not affect the probability of occurrence of the other event.

- **Definition of Independent Events.** Two events A and B are **independent** if and only if $P(A \mid B) = P(A)$ or, equivalently, $P(B \mid A) = P(B)$.

When two events are not independent, they are **dependent**: The probability of occurrence of one is related to the occurrence of the other. If we are trying to forecast one event, information about a dependent event may be useful, but information about an independent event will not be useful.

When two events are independent, the multiplication rule for probabilities, Equation 2, simplifies because $P(A \mid B)$ in that equation then equals $P(A)$.

- **Multiplication Rule for Independent Events.** When two events are independent, the joint probability of A and B equals the product of the individual probabilities of A and B.

$$P(AB) = P(A)P(B) \tag{4}$$

Therefore, if we are interested in two independent events with probabilities of 0.75 and 0.50, respectively, the probability that both will occur is $0.375 = 0.75(0.50)$. The multiplication rule for independent events generalizes to more than two events; for example, if A, B, and C are independent events, then $P(ABC) = P(A)P(B)P(C)$.

EXAMPLE 4

BankCorp's Earnings per Share (1)

As part of your work as a banking industry analyst, you build models for forecasting earnings per share of the banks you cover. Today you are studying BankCorp. The historical record shows that in 55 percent of recent quarters BankCorp's

EPS has increased sequentially, and in 45 percent of quarters EPS has decreased or remained unchanged sequentially.[6] At this point in your analysis, you are assuming that changes in sequential EPS are independent.

Earnings per share for 2Q:2014 (that is, EPS for the second quarter of 2014) were larger than EPS for 1Q:2014.

1 What is the probability that 3Q:2014 EPS will be larger than 2Q:2014 EPS (a positive change in sequential EPS)?

2 What is the probability that EPS decreases or remains unchanged in the next two quarters?

Solution to 1:

Under the assumption of independence, the probability that 3Q:2014 EPS will be larger than 2Q:2014 EPS is the unconditional probability of positive change, 0.55. The fact that 2Q:2014 EPS was larger than 1Q:2014 EPS is not useful information, as the next change in EPS is independent of the prior change.

Solution to 2:

The probability is 0.2025 = 0.45(0.45).

The following example illustrates how difficult it is to satisfy a set of independent criteria even when each criterion by itself is not necessarily stringent.

EXAMPLE 5

Screening Stocks for Investment

You have developed a stock screen—a set of criteria for selecting stocks. Your investment universe (the set of securities from which you make your choices) is the Russell 1000 Index, an index of 1,000 large-capitalization US equities. Your criteria capture different aspects of the selection problem; you believe that the criteria are independent of each other, to a close approximation.

Criterion	Fraction of Russell 1000 Stocks Meeting Criterion
First valuation criterion	0.50
Second valuation criterion	0.50
Analyst coverage criterion	0.25
Profitability criterion for company	0.55
Financial strength criterion for company	0.67

How many stocks do you expect to pass your screen?

Only 23 stocks out of 1,000 pass through your screen. If you define five events—*the stock passes the first valuation criterion, the stock passes the second valuation criterion, the stock passes the analyst coverage criterion, the company*

6 *Sequential* comparisons of quarterly EPS are with the immediate prior quarter. A sequential comparison stands in contrast to a comparison with the same quarter one year ago (another frequent type of comparison).

passes the profitability criterion, the company passes the financial strength criterion (say events *A, B, C, D,* and *E,* respectively)—then the probability that a stock will pass all five criteria, under independence, is

$$P(ABCDE) = P(A)P(B)P(C)P(D)P(E) = (0.50)(0.50)(0.25)(0.55)(0.67)$$
$$= 0.023031$$

Although only one of the five criteria is even moderately strict (the strictest lets 25 percent of stocks through), the probability that a stock can pass all five is only 0.023031, or about 2 percent. The size of the list of candidate investments is 0.023031(1,000) = 23.031, or 23 stocks.

An area of intense interest to investment managers and their clients is whether records of past performance are useful in identifying repeat winners and losers. The following example shows how this issue relates to the concept of independence.

EXAMPLE 6

Conditional Probabilities and Predictability of Mutual Fund Performance (2)

The purpose of the Vidal-Garcia (2013) study, introduced in Example 2, was to address the question of repeat European mutual fund winners and losers. If the status of a fund as a winner or a loser in one year is independent of whether it is a winner in the next year, the practical value of performance ranking is questionable. Using the four events defined in Example 2 as building blocks, we can define the following events to address the issue of predictability of mutual fund performance:

Fund is a Year 1 winner and *fund is a Year 2 winner*

Fund is a Year 1 winner and *fund is a Year 2 loser*

Fund is a Year 1 loser and *fund is a Year 2 winner*

Fund is a Year 1 loser and *fund is a Year 2 loser*

In Part 4 of Example 2, you calculated that

P(*fund is a Year 2 loser* and *fund is a Year 1 loser*) = 0.423

If the ranking in one year is independent of the ranking in the next year, what will you expect *P*(*fund is a Year 2 loser* and *fund is a Year 1 loser*) to be? Interpret the empirical probability 0.423.

By the multiplication rule for independent events, *P*(*fund is a Year 2 loser* and *fund is a Year 1 loser*) = *P*(*fund is a Year 2 loser*)*P*(*fund is a Year 1 loser*). Because 50 percent of funds are categorized as losers in each year, the unconditional probability that a fund is labeled a loser in either year is 0.50. Thus *P*(*fund is a Year 2 loser*)*P*(*fund is a Year 1 loser*) = 0.50(0.50) = 0.25. If the status of a fund as a loser in one year is independent of whether it is a loser in the prior year, we conclude that *P*(*fund is a Year 2 loser* and *fund is a Year 1 loser*) = 0.25. This probability is a priori because it is obtained from reasoning about the problem. You could also reason that the four events described above define categories and that if funds are randomly assigned to the four categories, there is a 1/4 probability of *fund is a Year 1 loser* and *fund is a Year 2 loser.* If the classifications in Year 1 and Year 2 were dependent, then the assignment of funds to categories would not be random. The empirical probability of 0.423 is above 0.25. Is this apparent predictability the result of chance? A test conducted by Vidal-Garcia indicated a less than 1 percent chance of observing the tabled data if the Year 1 and Year 2 rankings were independent.

In investments, the question of whether one event (or characteristic) provides information about another event (or characteristic) arises in both time-series settings (through time) and cross-sectional settings (among units at a given point in time). Examples 4 and 6 examined independence in a time-series setting. Example 5 illustrated independence in a cross-sectional setting. Independence/dependence relationships are often also explored in both settings using regression analysis, a technique we discuss in a later reading.

In many practical problems, we logically analyze a problem as follows: We formulate scenarios that we think affect the likelihood of an event that interests us. We then estimate the probability of the event, given the scenario. When the scenarios (conditioning events) are mutually exclusive and exhaustive, no possible outcomes are left out. We can then analyze the event using the **total probability rule**. This rule explains the unconditional probability of the event in terms of probabilities conditional on the scenarios.

The total probability rule is stated below for two cases. Equation 5 gives the simplest case, in which we have two scenarios. One new notation is introduced: If we have an event or scenario S, the event not-S, called the **complement** of S, is written S^C.[7] Note that $P(S) + P(S^C) = 1$, as either S or not-S must occur. Equation 6 states the rule for the general case of n mutually exclusive and exhaustive events or scenarios.

■ **The Total Probability Rule.**

$$\begin{aligned} P(A) &= P(AS) + P\left(AS^C\right) \\ &= P(A \mid S)P(S) + P\left(A \mid S^C\right)P\left(S^C\right) \end{aligned} \tag{5}$$

$$\begin{aligned} P(A) &= P(AS_1) + P(AS_2) + \ldots + P(AS_n) \\ &= P(A \mid S_1)P(S_1) + P(A \mid S_2)P(S_2) + \ldots + P(A \mid S_n)P(S_n) \end{aligned} \tag{6}$$

where S_1, S_2, …, S_n are mutually exclusive and exhaustive scenarios or events.

Equation 6 states the following: The probability of any event [$P(A)$] can be expressed as a weighted average of the probabilities of the event, given scenarios [terms such $P(A \mid S_1)$]; the weights applied to these conditional probabilities are the respective probabilities of the scenarios [terms such as $P(S_1)$ multiplying $P(A \mid S_1)$], and the scenarios must be mutually exclusive and exhaustive. Among other applications, this rule is needed to understand Bayes' formula, which we discuss later in the reading.

In the next example, we use the total probability rule to develop a consistent set of views about BankCorp's earnings per share.

EXAMPLE 7

BankCorp's Earnings per Share (2)

You are continuing your investigation into whether you can predict the direction of changes in BankCorp's quarterly EPS. You define four events:

7 For readers familiar with mathematical treatments of probability, S, a notation usually reserved for a concept called the sample space, is being appropriated to stand for *scenario*.

Event	**Probability**
A = *Change in sequential EPS is positive next quarter*	0.55
A^C = *Change in sequential EPS is 0 or negative next quarter*	0.45
S = *Change in sequential EPS is positive in the prior quarter*	0.55
S^C = *Change in sequential EPS is 0 or negative in the prior quarter*	0.45

On inspecting the data, you observe some persistence in EPS changes: Increases tend to be followed by increases, and decreases by decreases. The first probability estimate you develop is P(*change in sequential EPS is positive next quarter* | *change in sequential EPS is 0 or negative in the prior quarter*) = $P(A \mid S^C)$ = 0.40. The most recent quarter's EPS (2Q:2014) is announced, and the change is a positive sequential change (the event S). You are interested in forecasting EPS for 3Q:2014.

1. Write this statement in probability notation: "the probability that the change in sequential EPS is positive next quarter, given that the change in sequential EPS is positive the prior quarter."
2. Calculate the probability in Part 1. (Calculate the probability that is consistent with your other probabilities or beliefs.)

Solution to 1:

In probability notation, this statement is written $P(A \mid S)$.

Solution to 2:

The probability is 0.673 that the change in sequential EPS is positive for 3Q:2014, given the positive change in sequential EPS for 2Q:2014, as shown on the following page.

According to Equation 5, $P(A) = P(A \mid S)P(S) + P(A \mid S^C)P(S^C)$. The values of the probabilities needed to calculate $P(A \mid S)$ are already known: $P(A) = 0.55$, $P(S) = 0.55$, $P(S^C) = 0.45$, and $P(A \mid S^C) = 0.40$. Substituting into Equation 5,

$$0.55 = P(A \mid S)(0.55) + 0.40(0.45)$$

Solving for the unknown, $P(A \mid S) = [0.55 - 0.40(0.45)]/0.55 = 0.672727$, or 0.673.

You conclude that P(*change in sequential EPS is positive next quarter* | *change in sequential EPS is positive the prior quarter*) = 0.673. Any other probability is not consistent with your other estimated probabilities. Reflecting the persistence in EPS changes, this conditional probability of a positive EPS change, 0.673, is greater than the unconditional probability of an EPS increase, 0.55.

In the reading on statistical concepts and market returns, we discussed the concept of a weighted average or weighted mean. The example highlighted in that reading was that portfolio return is a weighted average of the returns on the individual assets in the portfolio, where the weight applied to each asset's return is the fraction of the portfolio invested in that asset. The total probability rule, which is a rule for stating an unconditional probability in terms of conditional probabilities, is also a weighted average. In that formula, probabilities of scenarios are used as weights. Part of the definition of weighted average is that the weights sum to 1. The probabilities of mutually exclusive and exhaustive events do sum to 1 (this is part of the definition of probability). The next weighted average we discuss, the expected value of a random variable, also uses probabilities as weights.

The expected value of a random variable is an essential quantitative concept in investments. Investors continually make use of expected values—in estimating the rewards of alternative investments, in forecasting EPS and other corporate financial variables and ratios, and in assessing any other factor that may affect their financial position. The expected value of a random variable is defined as follows:

- **Definition of Expected Value.** The **expected value** of a random variable is the probability-weighted average of the possible outcomes of the random variable. For a random variable *X*, the expected value of *X* is denoted *E*(*X*).

Expected value (for example, expected stock return) looks either to the future, as a forecast, or to the "true" value of the mean (the population mean, discussed in the reading on statistical concepts and market returns). We should distinguish expected value from the concepts of historical or sample mean. The sample mean also summarizes in a single number a central value. However, the sample mean presents a central value for a particular set of observations as an equally weighted average of those observations. To summarize, the contrast is forecast versus historical, or population versus sample.

EXAMPLE 8

BankCorp's Earnings per Share (3)

You continue with your analysis of BankCorp's EPS. In Table 3, you have recorded a probability distribution for BankCorp's EPS for the current fiscal year.

Table 3 Probability Distribution for BankCorp's EPS

Probability	EPS ($)
0.15	2.60
0.45	2.45
0.24	2.20
0.16	2.00
1.00	

What is the expected value of BankCorp's EPS for the current fiscal year?

Following the definition of expected value, list each outcome, weight it by its probability, and sum the terms.

$$E(\text{EPS}) = 0.15(\$2.60) + 0.45(\$2.45) + 0.24(\$2.20) + 0.16(\$2.00) = \$2.3405$$

The expected value of EPS is \$2.34.

An equation that summarizes your calculation in Example 8 is

$$E(X) = P(X_1)X_1 + P(X_2)X_2 + \ldots + P(X_n)X_n = \sum_{i=1}^{n} P(X_i)X_i \qquad (7)$$

where X_i is one of n possible outcomes of the random variable X.[8]

The expected value is our forecast. Because we are discussing random quantities, we cannot count on an individual forecast being realized (although we hope that, on average, forecasts will be accurate). It is important, as a result, to measure the risk we face. Variance and standard deviation measure the dispersion of outcomes around the expected value or forecast.

- **Definition of Variance.** The **variance** of a random variable is the expected value (the probability-weighted average) of squared deviations from the random variable's expected value:

$$\sigma^2(X) = E\left\{\left[X - E(X)\right]^2\right\} \quad (8)$$

- The two notations for variance are $\sigma^2(X)$ and Var(X).

Variance is a number greater than or equal to 0 because it is the sum of squared terms. If variance is 0, there is no dispersion or risk. The outcome is certain, and the quantity X is not random at all. Variance greater than 0 indicates dispersion of outcomes. Increasing variance indicates increasing dispersion, all else equal. Variance of X is a quantity in the squared units of X. For example, if the random variable is return in percent, variance of return is in units of percent squared. Standard deviation is easier to interpret than variance, as it is in the same units as the random variable. If the random variable is return in percent, standard deviation of return is also in units of percent. In the following example, when the variance of returns is stated as a percent or amount of money, to conserve space the reading may suppress showing the unit squared. Note that when the variance of returns is stated as a decimal, the complication of dealing with units of "percent squared" does not arise.

- **Definition of Standard Deviation.** **Standard deviation** is the positive square root of variance.

The best way to become familiar with these concepts is to work examples.

EXAMPLE 9

BankCorp's Earnings per Share (4)

In Example 8, you calculated the expected value of BankCorp's EPS as \$2.34, which is your forecast. Now you want to measure the dispersion around your forecast. Table 4 shows your view of the probability distribution of EPS for the current fiscal year.

Table 4 Probability Distribution for BankCorp's EPS

Probability	EPS (\$)
0.15	2.60
0.45	2.45
0.24	2.20

(continued)

8 For simplicity, we model all random variables in this reading as discrete random variables, which have a countable set of outcomes. For continuous random variables, which are discussed along with discrete random variables in the reading on common probability distributions, the operation corresponding to summation is integration.

Table 4 (Continued)

Probability	EPS ($)
0.16	2.00
1.00	

What are the variance and standard deviation of BankCorp's EPS for the current fiscal year?

The order of calculation is always expected value, then variance, then standard deviation. Expected value has already been calculated. Following the definition of variance above, calculate the deviation of each outcome from the mean or expected value, square each deviation, weight (multiply) each squared deviation by its probability of occurrence, and then sum these terms.

$$\begin{aligned}\sigma^2(\text{EPS}) &= P(\$2.60)[\$2.60 - E(\text{EPS})]^2 + P(\$2.45)[\$2.45 - E(\text{EPS})]^2 \\ &\quad + P(\$2.20)[\$2.20 - E(\text{EPS})]^2 + P(\$2.00)[\$2.00 - E(\text{EPS})]^2 \\ &= 0.15(2.60 - 2.34)^2 + 0.45(2.45 - 2.34)^2 \\ &\quad + 0.24(2.20 - 2.34)^2 + 0.16(2.00 - 2.34)^2 \\ &= 0.01014 + 0.005445 + 0.004704 + 0.018496 = 0.038785\end{aligned}$$

Standard deviation is the positive square root of 0.038785:

$$\sigma(\text{EPS}) = 0.038785^{1/2} = 0.196939, \text{ or approximately } 0.20.$$

An equation that summarizes your calculation of variance in Example 9 is

$$\begin{aligned}\sigma^2(X) &= P(X_1)[X_1 - E(X)]^2 + P(X_2)[X_2 - E(X)]^2 \\ &\quad + \ldots + P(X_n)[X_n - E(X)]^2 = \sum_{i=1}^{n} P(X_i)[X_i - E(X)]^2\end{aligned} \tag{9}$$

where X_i is one of n possible outcomes of the random variable X.

In investments, we make use of any relevant information available in making our forecasts. When we refine our expectations or forecasts, we are typically making adjustments based on new information or events; in these cases we are using **conditional expected values**. The expected value of a random variable X given an event or scenario S is denoted $E(X \mid S)$. Suppose the random variable X can take on any one of n distinct outcomes $X_1, X_2, \ldots, X_n$ (these outcomes form a set of mutually exclusive and exhaustive events). The expected value of X conditional on S is the first outcome, X_1, times the probability of the first outcome given S, $P(X_1 \mid S)$, plus the second outcome, X_2, times the probability of the second outcome given S, $P(X_2 \mid S)$, and so forth.

$$E(X \mid S) = P(X_1 \mid S)X_1 + P(X_2 \mid S)X_2 + \ldots + P(X_n \mid S)X_n \tag{10}$$

We will illustrate this equation shortly.

Parallel to the total probability rule for stating unconditional probabilities in terms of conditional probabilities, there is a principle for stating (unconditional) expected values in terms of conditional expected values. This principle is the **total probability rule for expected value**.

- **The Total Probability Rule for Expected Value.**

$$E(X) = E(X \mid S)P(S) + E(X \mid S^C)P(S^C) \tag{11}$$

$$E(X) = E(X \mid S_1)P(S_1) + E(X \mid S_2)P(S_2) + \ldots + E(X \mid S_n)P(S_n) \quad \textbf{(12)}$$

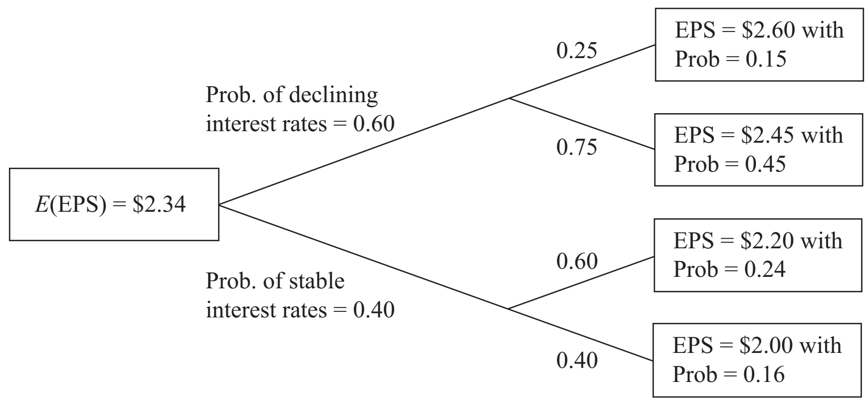
where $S_1, S_2, \ldots, S_n$ are mutually exclusive and exhaustive scenarios or events.

The general case, Equation 12, states that the expected value of X equals the expected value of X given Scenario 1, $E(X \mid S_1)$, times the probability of Scenario 1, $P(S_1)$, plus the expected value of X given Scenario 2, $E(X \mid S_2)$, times the probability of Scenario 2, $P(S_2)$, and so forth.

To use this principle, we formulate mutually exclusive and exhaustive scenarios that are useful for understanding the outcomes of the random variable. This approach was employed in developing the probability distribution of BankCorp's EPS in Examples 8 and 9, as we now discuss.

The earnings of BankCorp are interest rate sensitive, benefiting from a declining interest rate environment. Suppose there is a 0.60 probability that BankCorp will operate in a *declining interest rate environment* in the current fiscal year and a 0.40 probability that it will operate in a *stable interest rate environment* (assessing the chance of an increasing interest rate environment as negligible). If a *declining interest rate environment* occurs, the probability that EPS will be \$2.60 is estimated at 0.25, and the probability that EPS will be \$2.45 is estimated at 0.75. Note that 0.60, the probability of *declining interest rate environment,* times 0.25, the probability of \$2.60 EPS given a *declining interest rate environment,* equals 0.15, the (unconditional) probability of \$2.60 given in the table in Examples 8 and 9. The probabilities are consistent. Also, 0.60(0.75) = 0.45, the probability of \$2.45 EPS given in Tables 3 and 4. The **tree diagram** in Figure 2 shows the rest of the analysis.

Figure 2 BankCorp's Forecasted EPS



A declining interest rate environment points us to the **node** of the tree that branches off into outcomes of \$2.60 and \$2.45. We can find expected EPS given a declining interest rate environment as follows, using Equation 10:

$$\begin{aligned} E(\text{EPS} \mid \textit{declining interest rate environment}) &= 0.25(\$2.60) + 0.75(\$2.45) \\ &= \$2.4875 \end{aligned}$$

If interest rates are stable,

$$\begin{aligned} E(\text{EPS} \mid \textit{stable interest rate environment}) &= 0.60(\$2.20) + 0.40(\$2.00) \\ &= \$2.12 \end{aligned}$$

Once we have the new piece of information that interest rates are stable, for example, we revise our original expectation of EPS from \$2.34 downward to \$2.12. Now using the total probability rule for expected value,

$$E(\text{EPS}) = E(\text{EPS} \mid \textit{declining interest rate environment})P(\textit{declining interest rate environment}) + E(\text{EPS} \mid \textit{stable interest rate environment})P(\textit{stable interest rate environment})$$

So $E(\text{EPS}) = \$2.4875(0.60) + \$2.12(0.40) = \$2.3405$ or about \$2.34.

This amount is identical to the estimate of the expected value of EPS calculated directly from the probability distribution in Example 8. Just as our probabilities must be consistent, so must our expected values, unconditional and conditional; otherwise our investment actions may create profit opportunities for other investors at our expense.

To review, we first developed the factors or scenarios that influence the outcome of the event of interest. After assigning probabilities to these scenarios, we formed expectations conditioned on the different scenarios. Then we worked backward to formulate an expected value as of today. In the problem just worked, EPS was the event of interest, and the interest rate environment was the factor influencing EPS.

We can also calculate the variance of EPS given each scenario:

$$\begin{aligned}
&\sigma^2(\text{EPS} \mid \textit{declining interest rate environment}) \\
&\quad = P(\$2.60 \mid \textit{declining interest rate environment}) \\
&\quad \times[\$2.60 - E(\text{EPS} \mid \textit{declining interest rate environment})]^2 \\
&\quad + P(\$2.45 \mid \textit{declining interest rate environment}) \\
&\quad \times[\$2.45 - E(\text{EPS} \mid \textit{declining interest rate environment})]^2 \\
&\quad = 0.25(\$2.60 - \$2.4875)^2 + 0.75(\$2.45 - \$2.4875)^2 \\
&\quad = 0.004219
\end{aligned}$$

$$\begin{aligned}
&\sigma^2(\text{EPS} \mid \textit{stable interest rate environment}) \\
&\quad = P(\$2.20 \mid \textit{stable interest rate environment}) \\
&\quad \times[\$2.20 - E(\text{EPS} \mid \textit{stable interest rate environment})]^2 \\
&\quad + P(\$2.00 \mid \textit{stable interest rate environment}) \\
&\quad \times[\$2.00 - E(\text{EPS} \mid \textit{stable interest rate environment})]^2 \\
&\quad = 0.60(\$2.20 - \$2.12)^2 + 0.40(\$2.00 - \$2.12)^2 = 0.0096
\end{aligned}$$

These are **conditional variances**, the variance of EPS given a *declining interest rate environment* and the variance of EPS given a *stable interest rate environment*. The relationship between unconditional variance and conditional variance is a relatively advanced topic.[9] The main points are 1) that variance, like expected value, has a conditional counterpart to the unconditional concept and 2) that we can use conditional variance to assess risk given a particular scenario.

9 The unconditional variance of EPS is the sum of two terms: 1) the expected value (probability-weighted average) of the conditional variances (parallel to the total probability rules) and 2) the variance of conditional expected values of EPS. The second term arises because the variability in conditional expected value is a source of risk. Term 1 is $\sigma^2(\text{EPS}) = P(\textit{declining interest rate environment})\,\sigma^2(\text{EPS} \mid \textit{declining interest rate environment}) + P(\textit{stable interest rate environment})\,\sigma^2(\text{EPS} \mid \textit{stable interest rate environment}) = 0.60(0.004219) + 0.40(0.0096) = 0.006371$. Term 2 is $\sigma^2[E(\text{EPS} \mid \text{interest rate environment})] = 0.60(\$2.4875 - \$2.34)^2 + 0.40(\$2.12 - \$2.34)^2 = 0.032414$. Summing the two terms, unconditional variance equals $0.006371 + 0.032414 = 0.038785$.

EXAMPLE 10

BankCorp's Earnings per Share (5)

Continuing with BankCorp, you focus now on BankCorp's cost structure. One model you are researching for BankCorp's operating costs is

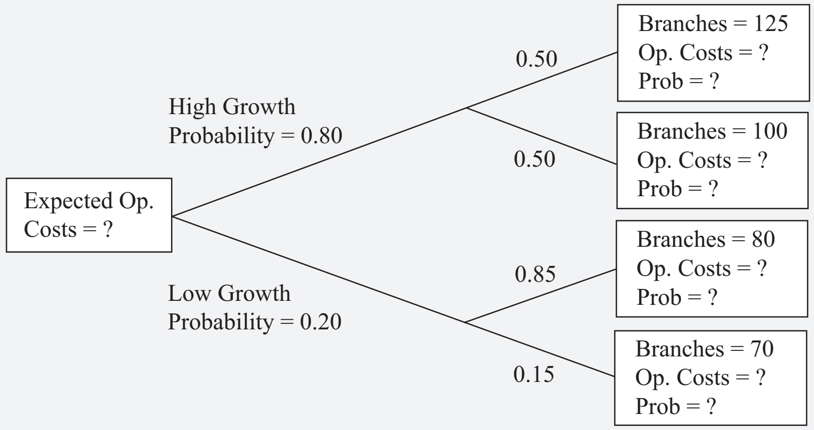
$$\hat{Y} = a + bX$$

where $\hat{Y}$ is a forecast of operating costs in millions of dollars and X is the number of branch offices. $\hat{Y}$ represents the expected value of Y given X, or $E(Y \mid X)$. ($\hat{Y}$ is a notation used in regression analysis, which we discuss in a later reading.) You interpret the intercept a as fixed costs and b as variable costs. You estimate the equation as

$$\hat{Y} = 12.5 + 0.65X$$

BankCorp currently has 66 branch offices, and the equation estimates that 12.5 + 0.65(66) = \$55.4 million. You have two scenarios for growth, pictured in the tree diagram in Figure 3.

Figure 3 BankCorp's Forecasted Operating Costs



1. Compute the forecasted operating costs given the different levels of operating costs, using $\hat{Y} = 12.5 + 0.65X$. State the probability of each level of the number of branch offices. These are the answers to the questions in the terminal boxes of the tree diagram.
2. Compute the expected value of operating costs under the high growth scenario. Also calculate the expected value of operating costs under the low growth scenario.
3. Answer the question in the initial box of the tree: What are BankCorp's expected operating costs?

Solution to 1:

Using $\hat{Y} = 12.5 + 0.65X$, from top to bottom, we have

Operating Costs	Probability
$\hat{Y} = 12.5 + 0.65(125) = \$93.75$ million	0.80(0.50) = 0.40
$\hat{Y} = 12.5 + 0.65(100) = \$77.50$ million	0.80(0.50) = 0.40
$\hat{Y} = 12.5 + 0.65(80) = \$64.50$ million	0.20(0.85) = 0.17
$\hat{Y} = 12.5 + 0.65(70) = \$58.00$ million	0.20(0.15) = 0.03
	Sum = 1.00

Solution to 2:

Dollar amounts are in millions.

$$\begin{aligned} E(\text{operating costs} \mid \textit{high growth}) &= 0.50(\$93.75) + 0.50(\$77.50) \\ &= \$85.625 \end{aligned}$$

$$\begin{aligned} E(\text{operating costs} \mid \textit{low growth}) &= 0.85(\$64.50) + 0.15(\$58.00) \\ &= \$63.525 \end{aligned}$$

Solution to 3:

Dollar amounts are in millions.

$$\begin{aligned} E(\text{operating costs}) &= E(\text{operating costs} \mid \textit{high growth})P(\textit{high growth}) \\ &\quad + E(\text{operating costs} \mid \textit{low growth})P(\textit{low growth}) \\ &= \$85.625(0.80) + \$63.525(0.20) = \$81.205 \end{aligned}$$

BankCorp's expected operating costs are \$81.205 million.

We will see conditional probabilities again when we discuss Bayes' formula. This section has introduced a few problems that can be addressed using probability concepts. The following problem draws on these concepts, as well as on analytical skills.

EXAMPLE 11

The Default Risk Premium for a One-Period Debt Instrument

As the co-manager of a short-term bond portfolio, you are reviewing the pricing of a speculative-grade, one-year-maturity, zero-coupon bond. For this type of bond, the return is the difference between the amount paid and the principal value received at maturity. Your goal is to estimate an appropriate default risk premium for this bond. You define the default risk premium as the extra return above the risk-free return that will compensate investors for default risk. If R is the promised return (yield-to-maturity) on the debt instrument and R_F is the risk-free rate, the default risk premium is $R - R_F$. You assess the probability that the bond defaults as P(*the bond defaults*) = 0.06. Looking at current money market yields, you find that one-year US Treasury bills (T-bills) are offering a return of 2 percent, an estimate of R_F. As a first step, you make the simplifying assumption that bondholders will recover nothing in the event of a default. What is the minimum default risk premium you should require for this instrument?

The challenge in this type of problem is to find a starting point. In many problems, including this one, an effective first step is to divide up the possible outcomes into mutually exclusive and exhaustive events in an economically logical way. Here, from the viewpoint of a bondholder, the two events that affect returns are *the bond defaults* and *the bond does not default*. These two events cover all outcomes. How do these events affect a bondholder's returns? A second step is to compute the value of the bond for the two events. We have no specifics on bond **face value**, but we can compute value per \$1 or one unit of currency invested.

	The Bond Defaults	*The Bond Does Not Default*
Bond value	\$0	\$(1 + R)

The third step is to find the expected value of the bond (per \$1 invested).

$$E(\text{bond}) = \$0 \times P(\textit{the bond defaults}) + \$(1 + R)[1 - P(\textit{the bond defaults})]$$

So $E(\text{bond}) = \$(1 + R)[1 - P(\textit{the bond defaults})]$. The expected value of the T-bill per \$1 invested is $(1 + R_F)$. In fact, this value is certain because the T-bill is risk free. The next step requires economic reasoning. You want the default premium to be large enough so that you expect to at least break even compared with investing in the T-bill. This outcome will occur if the expected value of the bond equals the expected value of the T-bill per \$1 invested.

$$\begin{aligned} \textit{Expected Value of Bond} &= \textit{Expected Value of T-Bill} \\ \$(1 + R)\left[1 - P(\textit{the bond defaults})\right] &= (1 + R_F) \end{aligned}$$

Solving for the promised return on the bond, you find $R = \{(1 + R_F)/[1 - P(\textit{the bond defaults})]\} - 1$. Substituting in the values in the statement of the problem, $R = [1.02/(1 - 0.06)] - 1 = 1.08511 - 1 = 0.08511$ or about 8.51 percent, and default risk premium is $R - R_F = 8.51\% - 2\% = 6.51\%$.

You require a default risk premium of at least 651 basis points. You can state the matter as follows: If the bond is priced to yield 8.51 percent, you will earn a 651 basis-point spread and receive the bond principal with 94 percent probability. If the bond defaults, however, you will lose everything. With a premium of 651 basis points, you expect to just break even relative to an investment in T-bills. Because an investment in the zero-coupon bond has variability, if you are risk averse you will demand that the premium be larger than 651 basis points.

This analysis is a starting point. Bondholders usually recover part of their investment after a default. A next step would be to incorporate a recovery rate.

In this section, we have treated random variables such as EPS as stand-alone quantities. We have not explored how descriptors such as expected value and variance of EPS may be functions of other random variables. Portfolio return is one random variable that is clearly a function of other random variables, the random returns on the individual securities in the portfolio. To analyze a portfolio's expected return and variance of return, we must understand these quantities are a function of characteristics of the individual securities' returns. Looking at the dispersion or variance of portfolio return, we see that the way individual security returns move together or covary is important. To understand the significance of these movements, we need to explore some new concepts, covariance and correlation. The next section, which deals with portfolio expected return and variance of return, introduces these concepts.

3 PORTFOLIO EXPECTED RETURN AND VARIANCE OF RETURN

Modern portfolio theory makes frequent use of the idea that investment opportunities can be evaluated using expected return as a measure of reward and variance of return as a measure of risk. The calculation and interpretation of portfolio expected return and variance of return are fundamental skills. In this section, we will develop an understanding of portfolio expected return and variance of return.[10] Portfolio return is determined by the returns on the individual holdings. As a result, the calculation of portfolio variance, as a function of the individual asset returns, is more complex than the variance calculations illustrated in the previous section.

We work with an example of a portfolio that is 50 percent invested in an S&P 500 Index fund, 25 percent invested in a US long-term corporate bond fund, and 25 percent invested in a fund indexed to the MSCI EAFE Index (representing equity markets in Europe, Australasia, and the Far East). Table 5 shows these weights.

Table 5 Portfolio Weights

Asset Class	Weights
S&P 500	0.50
US long-term corporate bonds	0.25
MSCI EAFE	0.25

We first address the calculation of the expected return on the portfolio. In the previous section, we defined the expected value of a random variable as the probability-weighted average of the possible outcomes. Portfolio return, we know, is a weighted average of the returns on the securities in the portfolio. Similarly, the expected return on a portfolio is a weighted average of the expected returns on the securities in the portfolio, using exactly the same weights. When we have estimated the expected returns on the individual securities, we immediately have portfolio expected return. This convenient fact follows from the properties of expected value.

- **Properties of Expected Value.** Let w_i be any constant and R_i be a random variable.
  1. The expected value of a constant times a random variable equals the constant times the expected value of the random variable.

     $$E(w_iR_i) = w_iE(R_i)$$

  2. The expected value of a weighted sum of random variables equals the weighted sum of the expected values, using the same weights.

     $$E(w_1R_1 + w_2R_2 + \ldots + w_nR_n) = w_1E(R_1) + w_2E(R_2) + \ldots + w_nE(R_n) \quad (13)$$

10 Although we outline a number of basic concepts in this section, we do not present mean–variance analysis per se. For a presentation of mean–variance analysis, see the readings on portfolio concepts, as well as the extended treatments in standard investment textbooks such as Bodie, Kane, and Marcus (2017), Elton, Gruber, Brown, and Goetzmann (2013), and Reilly and Brown (2018).

Suppose we have a random variable with a given expected value. If we multiply each outcome by 2, for example, the random variable's expected value is multiplied by 2 as well. That is the meaning of Part 1. The second statement is the rule that directly leads to the expression for portfolio expected return. A portfolio with n securities is defined by its portfolio weights, w_1, w_2, …, w_n, which sum to 1. So portfolio return, R_p, is $R_p = w_1R_1 + w_2R_2 + \ldots + w_nR_n$. We can state the following principle:

- **Calculation of Portfolio Expected Return.** Given a portfolio with n securities, the expected return on the portfolio is a weighted average of the expected returns on the component securities:

$$\begin{aligned} E\left(R_p\right) &= E\left(w_1R_1 + w_2R_2 + \ldots + w_nR_n\right) \\ &= w_1E\left(R_1\right) + w_2E\left(R_2\right) + \ldots + w_nE\left(R_n\right) \end{aligned}$$

Suppose we have estimated expected returns on the assets in the portfolio, as given in Table 6.

Table 6 Weights and Expected Returns

Asset Class	Weight	Expected Return (%)
S&P 500	0.50	13
US long-term corporate bonds	0.25	6
MSCI EAFE	0.25	15

We calculate the expected return on the portfolio as 11.75 percent:

$$\begin{aligned} E\left(R_p\right) &= w_1E\left(R_1\right) + w_2E\left(R_2\right) + w_3E\left(R_3\right) \\ &= 0.50(13\%) + 0.25(6\%) + 0.25(15\%) = 11.75\% \end{aligned}$$

In the previous section, we studied variance as a measure of dispersion of outcomes around the expected value. Here we are interested in portfolio variance of return as a measure of investment risk. Letting R_p stand for the return on the portfolio, portfolio variance is $\sigma^2(R_p) = E\{[R_p - E(R_p)]^2\}$ according to Equation 8. How do we implement this definition? In the reading on statistical concepts and market returns, we learned how to calculate a historical or sample variance based on a sample of returns. Now we are considering variance in a forward-looking sense. We will use information about the individual assets in the portfolio to obtain portfolio variance of return. To avoid clutter in notation, we write ER_p for $E(R_p)$. We need the concept of covariance.

- **Definition of Covariance.** Given two random variables R_i and R_j, the covariance between R_i and R_j is

$$\operatorname{Cov}\left(R_i, R_j\right) = E\left[\left(R_i - ER_i\right)\left(R_j - ER_j\right)\right] \tag{14}$$

Alternative notations are $\sigma(R_i,R_j)$ and σ_{ij}.

Equation 14 states that the covariance between two random variables is the probability-weighted average of the cross-products of each random variable's deviation from its own expected value. We will return to discuss covariance after we establish the need for the concept. Working from the definition of variance, we find

$$\begin{aligned}
\sigma^2(R_p) &= E\left[\left(R_p - ER_p\right)^2\right] \\
&= E\left\{\left[w_1R_1 + w_2R_2 + w_3R_3 - E\left(w_1R_1 + w_2R_2 + w_3R_3\right)\right]^2\right\} \\
&= E\left\{\left[w_1R_1 + w_2R_2 + w_3R_3 - w_1ER_1 - w_2ER_2 - w_3ER_3\right]^2\right\} \\
&\qquad \text{(using Equation 13)} \\
&= E\left\{\left[w_1(R_1 - ER_1) + w_2(R_2 - ER_2) + w_3(R_3 - ER_3)\right]^2\right\} \\
&\qquad \text{(rearranging)} \\
&= E\{\left[w_1(R_1 - ER_1) + w_2(R_2 - ER_2) + w_3(R_3 - ER_3)\right] \\
&\quad \times\left[w_1(R_1 - ER_1) + w_2(R_2 - ER_2) + w_3(R_3 - ER_3)\right]\} \\
&\qquad \text{(what squaring means)} \\
&= E[w_1w_1(R_1 - ER_1)(R_1 - ER_1) + w_1w_2(R_1 - ER_1)(R_2 - ER_2) \\
&\quad + w_1w_3(R_1 - ER_1)(R_3 - ER_3) + w_2w_1(R_2 - ER_2)(R_1 - ER_1) \\
&\quad + w_2w_2(R_2 - ER_2)(R_2 - ER_2) + w_2w_3(R_2 - ER_2)(R_3 - ER_3) \\
&\quad + w_3w_1(R_3 - ER_3)(R_1 - ER_1) + w_3w_2(R_3 - ER_3)(R_2 - ER_2) \\
&\quad + w_3w_3(R_3 - ER_3)(R_3 - ER_3)] \qquad \text{(doing the multiplication)} \\
&= w_1^2E\left[(R_1 - ER_1)^2\right] + w_1w_2E\left[(R_1 - ER_1)(R_2 - ER_2)\right] \\
&\quad + w_1w_3E\left[(R_1 - ER_1)(R_3 - ER_3)\right] + w_2w_1E\left[(R_2 - ER_2)(R_1 - ER_1)\right] \\
&\quad + w_2^2E\left[(R_2 - ER_2)^2\right] + w_2w_3E\left[(R_2 - ER_2)(R_3 - ER_3)\right] \\
&\quad + w_3w_1E\left[(R_3 - ER_3)(R_1 - ER_1)\right] + w_3w_2E\left[(R_3 - ER_3)(R_2 - ER_2)\right] \\
&\quad + w_3^2E\left[(R_3 - ER_3)^2\right] \qquad \text{(recalling that the } w_i \text{ terms are constants)}
\end{aligned}$$

$$\begin{aligned}
&= w_1^2\sigma^2(R_1) + w_1w_2\text{Cov}(R_1,R_2) + w_1w_3\text{Cov}(R_1,R_3) \\
&\quad + w_1w_2\mathit{Cov}(R_1,R_2) + w_2^2\sigma^2(R_2) + w_2w_3\text{Cov}(R_2,R_3) \qquad \text{(15)} \\
&\quad + w_1w_3\mathit{Cov}(R_1,R_3) + w_2w_3\mathit{Cov}(R_2,R_3) + w_3^2\sigma^2(R_3)
\end{aligned}$$

The last step follows from the definitions of variance and covariance.[11] For the italicized covariance terms in Equation 15, we used the fact that the order of variables in covariance does not matter: $\text{Cov}(R_2,R_1) = \text{Cov}(R_1,R_2)$, for example. As we will show, the diagonal variance terms $\sigma^2(R_1)$, $\sigma^2(R_2)$, and $\sigma^2(R_3)$ can be expressed as $\text{Cov}(R_1,R_1)$, $\text{Cov}(R_2,R_2)$, and $\text{Cov}(R_3,R_3)$, respectively. Using this fact, the most compact way to

11 Useful facts about variance and covariance include: 1) The variance of a constant *times* a random variable equals the constant squared times the variance of the random variable, or $\sigma^2(wR) = w^2\sigma^2(R)$; 2) The variance of a constant *plus* a random variable equals the variance of the random variable, or $\sigma^2(w + R) = \sigma^2(R)$ because a constant has zero variance; 3) The covariance between a constant and a random variable is zero.

state Equation 15 is $\sigma^2\left(R_p\right) = \sum_{i=1}^{3}\sum_{j=1}^{3} w_i w_j \mathrm{Cov}\left(R_i, R_j\right)$. The double summation signs say: "Set i = 1 and let j run from 1 to 3; then set i = 2 and let j run from 1 to 3; next set i = 3 and let j run from 1 to 3; finally, add the nine terms." This expression generalizes for a portfolio of any size n to

$$\sigma^2\left(R_p\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \mathrm{Cov}\left(R_i, R_j\right) \tag{16}$$

We see from Equation 15 that individual variances of return constitute part, but not all, of portfolio variance. The three variances are actually outnumbered by the six covariance terms off the diagonal. For three assets, the ratio is 1 to 2, or 50 percent. If there are 20 assets, there are 20 variance terms and 20(20) − 20 = 380 off-diagonal covariance terms. The ratio of variance terms to off-diagonal covariance terms is less than 6 to 100, or 6 percent. A first observation, then, is that as the number of holdings increases, covariance[12] becomes increasingly important, all else equal.

What exactly is the effect of covariance on portfolio variance? The covariance terms capture how the co-movements of returns affect portfolio variance. For example, consider two stocks: One tends to have high returns (relative to its expected return) when the other has low returns (relative to its expected return). The returns on one stock tend to offset the returns on the other stock, lowering the variability or variance of returns on the portfolio. Like variance, the units of covariance are hard to interpret, and we will introduce a more intuitive concept shortly. Meanwhile, from the definition of covariance, we can establish two essential observations about covariance.

1 We can interpret the sign of covariance as follows:

   Covariance of returns is negative if, when the return on one asset is above its expected value, the return on the other asset tends to be below its expected value (an average inverse relationship between returns).

   Covariance of returns is 0 if returns on the assets are unrelated.

   Covariance of returns is positive when the returns on both assets tend to be on the same side (above or below) their expected values at the same time (an average positive relationship between returns).

2 The covariance of a random variable with itself (*own covariance*) is its own variance: $\mathrm{Cov}(R,R) = E\{[R - E(R)][R - E(R)]\} = E\{[R - E(R)]^2\} = \sigma^2(R)$.

A complete list of the covariances constitutes all the statistical data needed to compute portfolio variance of return. Covariances are often presented in a square format called a **covariance matrix**. Table 7 summarizes the inputs for portfolio expected return and variance of return.

Table 7 Inputs to Portfolio Expected Return and Variance

A. Inputs to Portfolio Expected Return

Asset	A	B	C
	$E(R_A)$	$E(R_B)$	$E(R_C)$

(continued)

12 When the meaning of covariance as "off-diagonal covariance" is obvious, as it is here, we omit the qualifying words. Covariance is usually used in this sense.

Table 7 (Continued)

B. Covariance Matrix: The Inputs to Portfolio Variance of Return

Asset	A	B	C
A	$\mathbf{Cov}(\boldsymbol{R_A},\boldsymbol{R_A})$	$\text{Cov}(R_A,R_B)$	$\text{Cov}(R_A,R_C)$
B	$\text{Cov}(R_B,R_A)$	$\mathbf{Cov}(\boldsymbol{R_B},\boldsymbol{R_B})$	$\text{Cov}(R_B,R_C)$
C	$\text{Cov}(R_C,R_A)$	$\text{Cov}(R_C,R_B)$	$\mathbf{Cov}(\boldsymbol{R_C},\boldsymbol{R_C})$

With three assets, the covariance matrix has $3^2 = 3 \times 3 = 9$ entries, but it is customary to treat the diagonal terms, the variances, separately from the off-diagonal terms. These diagonal terms are bolded in Table 7. This distinction is natural, as security variance is a single-variable concept. So there are 9 − 3 = 6 covariances, excluding variances. But $\text{Cov}(R_B,R_A) = \text{Cov}(R_A,R_B)$, $\text{Cov}(R_C,R_A) = \text{Cov}(R_A,R_C)$, and $\text{Cov}(R_C,R_B) = \text{Cov}(R_B,R_C)$. The covariance matrix below the diagonal is the mirror image of the covariance matrix above the diagonal. As a result, there are only 6/2 = 3 distinct covariance terms to estimate. In general, for n securities, there are $n(n - 1)/2$ distinct covariances to estimate and n variances to estimate.

Suppose we have the covariance matrix shown in Table 8. We will be working in returns stated as percents and the table entries are in units of percent squared ($\%^2$). The terms $38\%^2$ and $400\%^2$ are 0.0038 and 0.0400, respectively, stated as decimals; correctly working in percents and decimals leads to identical answers.

Table 8 Covariance Matrix

	S&P 500	US Long-Term Corporate Bonds	MSCI EAFE
S&P 500	400	45	189
US long-term corporate bonds	45	81	38
MSCI EAFE	189	38	441

Taking Equation 15 and grouping variance terms together produces the following:

$$\begin{aligned}\sigma^2\left(R_p\right) &= w_1^2\sigma^2\left(R_1\right) + w_2^2\sigma^2\left(R_2\right) + w_3^2\sigma^2\left(R_3\right) + 2w_1w_2\text{Cov}\left(R_1,R_2\right)\\ &\quad + 2w_1w_3\text{Cov}\left(R_1,R_3\right) + 2w_2w_3\text{Cov}\left(R_2,R_3\right)\\ &= (0.50)^2(400) + (0.25)^2(81) + (0.25)^2(441)\\ &\quad + 2(0.50)(0.25)(45) + 2(0.50)(0.25)(189)\\ &\quad + 2(0.25)(0.25)(38)\\ &= 100 + 5.0625 + 27.5625 + 11.25 + 47.25 + 4.75 = 195.875\end{aligned} \tag{17}$$

The variance is 195.875. Standard deviation of return is $195.875^{1/2}$ = 14 percent. To summarize, the portfolio has an expected annual return of 11.75 percent and a standard deviation of return of 14 percent.

Let us look at the first three terms in the calculation above. Their sum, 100 + 5.0625 + 27.5625 = 132.625, is the contribution of the individual variances to portfolio variance. If the returns on the three assets were independent, covariances would be 0 and the standard deviation of portfolio return would be $132.625^{1/2}$ = 11.52 percent as compared to 14 percent before. The portfolio would have less risk. Suppose the covariance terms were negative. Then a negative number would be added to 132.625, so portfolio variance and risk would be even smaller. At the same time, we have not changed expected return. For the same expected portfolio return, the portfolio has less risk. This risk reduction is a diversification benefit, meaning a risk-reduction benefit from holding a portfolio of assets. The diversification benefit increases with decreasing covariance. This observation is a key insight of modern portfolio theory. It is even more intuitively stated when we can use the concept of **correlation**. Then we can say that as long as security returns are not perfectly positively correlated, diversification benefits are possible. Furthermore, the smaller the correlation between security returns, the greater the cost of not diversifying (in terms of risk-reduction benefits forgone), all else equal.

- **Definition of Correlation.** The correlation between two random variables, R_i and R_j, is defined as $\rho(R_i,R_j) = \text{Cov}(R_i,R_j)/\sigma(R_i)\sigma(R_j)$. Alternative notations are $\text{Corr}(R_i,R_j)$ and ρ_{ij}.

Frequently, covariance is substituted out using the relationship $\text{Cov}(R_i,R_j) = \rho(R_i,R_j)\sigma(R_i)\sigma(R_j)$. The division indicated in the definition makes correlation a pure number (one without a unit of measurement) and places bounds on its largest and smallest possible values. Using the above definition, we can state a correlation matrix from data in the covariance matrix alone. Table 9 shows the correlation matrix.

Table 9 Correlation Matrix of Returns

	S&P 500	US Long-Term Corporate Bonds	MSCI EAFE
S&P 500	1.00	0.25	0.45
US long-term corporate bonds	0.25	1.00	0.20
MSCI EAFE	0.45	0.20	1.00

For example, the covariance between long-term bonds and MSCI EAFE is 38, from Table 8. The standard deviation of long-term bond returns is $81^{1/2}$ = 9 percent, that of MSCI EAFE returns is $441^{1/2}$ = 21 percent, from diagonal terms in Table 8. The correlation ρ(Return on long-term bonds, Return on EAFE) is 38/(9%)(21%) = 0.201, rounded to 0.20. The correlation of the S&P 500 with itself equals 1: The calculation is its own covariance divided by its standard deviation squared.

- **Properties of Correlation.**

  1 Correlation is a number between −1 and +1 for two random variables, X and Y:

  $$-1 \le \rho(X,Y) \le +1$$

  2 A correlation of 0 (uncorrelated variables) indicates an absence of any linear (straight-line) relationship between the variables.[13] Increasingly positive correlation indicates an increasingly strong positive linear relationship (up to 1, which indicates a perfect linear relationship). Increasingly negative correlation indicates an increasingly strong negative (inverse) linear relationship (down to −1, which indicates a perfect inverse linear relationship).[14]

13 If the correlation is 0, $R_1 = a + bR_2$ + error, with $b = 0$.
14 If the correlation is positive, $R_1 = a + bR_2$ + error, with $b > 0$. If the correlation is negative, $b < 0$.

EXAMPLE 12

Portfolio Expected Return and Variance of Return

You have a portfolio of two mutual funds, A and B, 75 percent invested in A, as shown in Table 10.

Table 10 Mutual Fund Expected Returns, Return Variances, and Covariances

Fund	**A** $E(R_A) = 20\%$	**B** $E(R_B) = 12\%$
	Covariance Matrix	
Fund	A	B
A	625	120
B	120	196

1 Calculate the expected return of the portfolio.

2 Calculate the correlation matrix for this problem. Carry out the answer to two decimal places.

3 Compute portfolio standard deviation of return.

Solution to 1:

$E(R_p) = w_A E(R_A) + (1 - w_A)E(R_B) = 0.75(20\%) + 0.25(12\%) = 18\%$. Portfolio weights must sum to 1: $w_B = 1 - w_A$.

Solution to 2:

$\sigma(R_A) = 625^{1/2} = 25$ percent $\sigma(R_B) = 196^{1/2} = 14$ percent. There is one distinct covariance and thus one distinct correlation: $\rho(R_A,R_B) = \text{Cov}(R_A,R_B)/\sigma(R_A)\sigma(R_B) = 120/[25(14)] = 0.342857$, or 0.34 Table 11 shows the correlation matrix.

Table 11 Correlation Matrix

	A	**B**
A	1.00	0.34
B	0.34	1.00

Diagonal terms are always equal to 1 in a correlation matrix.

Solution to 3:

$$\begin{aligned}\sigma^2(R_p) &= w_A^2\sigma^2(R_A) + w_B^2\sigma^2(R_B) + 2w_A w_B \text{Cov}(R_A,R_B)\\ &= (0.75)^2(625) + (0.25)^2(196) + 2(0.75)(0.25)(120)\\ &= 351.5625 + 12.25 + 45 = 408.8125\\ \sigma(R_p) &= 408.8125^{1/2} = 20.22 \text{ percent}\end{aligned}$$

How do we estimate return covariance and correlation? Frequently, we make forecasts on the basis of historical covariance or use other methods based on historical return data, such as a market model regression.[15] We can also calculate covariance using the **joint probability function** of the random variables, if that can be estimated. The joint probability function of two random variables X and Y, denoted $P(X,Y)$, gives the probability of joint occurrences of values of X and Y. For example, $P(3, 2)$, is the probability that X equals 3 and Y equals 2.

Suppose that the joint probability function of the returns on BankCorp stock (R_A) and the returns on NewBank stock (R_B) has the simple structure given in Table 12.

Table 12 Joint Probability Function of BankCorp and NewBank Returns (Entries Are Joint Probabilities)

	R_B = 20%	R_B = 16%	R_B = 10%
R_A = 25%	0.20	0	0
R_A = 12%	0	0.50	0
R_A = 10%	0	0	0.30

The expected return on BankCorp stock is 0.20(25%) + 0.50(12%) + 0.30(10%) = 14%. The expected return on NewBank stock is 0.20(20%) + 0.50(16%) + 0.30(10%) = 15%. The joint probability function above might reflect an analysis based on whether banking industry conditions are good, average, or poor. Table 13 presents the calculation of covariance.

Table 13 Covariance Calculations

Banking Industry Condition	Deviations BankCorp	Deviations NewBank	Product of Deviations	Probability of Condition	Probability-Weighted Product
Good	25–14	20–15	55	0.20	11
Average	12–14	16–15	–2	0.50	–1
Poor	10–14	10–15	20	0.30	6
					Cov(R_A,R_B) = 16

Note: Expected return for BankCorp is 14% and for NewBank, 15%.

The first and second columns of numbers show, respectively, the deviations of BankCorp and NewBank returns from their mean or expected value. The next column shows the product of the deviations. For example, for good industry conditions, (25 – 14)(20 – 15) = 11(5) = 55. Then 55 is multiplied or weighted by 0.20, the probability that banking industry conditions are good: 55(0.20) = 11. The calculations for average and poor banking conditions follow the same pattern. Summing up these probability-weighted products, we find that Cov(R_A,R_B) = 16.

A formula for computing the covariance between random variables R_A and R_B is

$$\mathrm{Cov}\left(R_A, R_B\right) = \sum_i \sum_j P\left(R_{A,i}, R_{B,j}\right)\left(R_{A,i} - ER_A\right)\left(R_{B,j} - ER_B\right) \quad \textbf{(18)}$$

15 See any of the textbooks mentioned in Footnote 10.

The formula tells us to sum all possible deviation cross-products weighted by the appropriate joint probability. In the example we just worked, as Table 12 shows, only three joint probabilities are nonzero. Therefore, in computing the covariance of returns in this case, we need to consider only three cross-products:

$$\begin{aligned}\text{Cov}(R_A,R_B) &= P(25,20)\big[(25-14)(20-15)\big] + P(12,16)\big[(12-14)\\ &\quad (16-15)\big] + P(10,10)\big[(10-14)(10-15)\big]\\ &= 0.20(11)(5) + 0.50(-2)(1) + 0.30(-4)(-5)\\ &= 11 - 1 + 6 = 16\end{aligned}$$

One theme of this reading has been independence. Two random variables are independent when every possible pair of events—one event corresponding to a value of X and another event corresponding to a value of Y—are independent events. When two random variables are independent, their joint probability function simplifies.

- **Definition of Independence for Random Variables.** Two random variables X and Y are independent if and only if $P(X,Y) = P(X)P(Y)$.

For example, given independence, $P(3,2) = P(3)P(2)$. We multiply the individual probabilities to get the joint probabilities. *Independence* is a stronger property than *uncorrelatedness* because correlation addresses only linear relationships. The following condition holds for independent random variables and, therefore, also holds for uncorrelated random variables.

- **Multiplication Rule for Expected Value of the Product of Uncorrelated Random Variables.** The expected value of the product of uncorrelated random variables is the product of their expected values.

  $E(XY) = E(X)E(Y)$ if X and Y are uncorrelated.

Many financial variables, such as revenue (price times quantity), are the product of random quantities. When applicable, the above rule simplifies calculating expected value of a product of random variables.[16]

4 TOPICS IN PROBABILITY

In the remainder of the reading we discuss two topics that can be important in solving investment problems. We start with Bayes' formula: what probability theory has to say about learning from experience. Then we move to a discussion of shortcuts and principles for counting.

4.1 Bayes' Formula

When we make decisions involving investments, we often start with viewpoints based on our experience and knowledge. These viewpoints may be changed or confirmed by new knowledge and observations. Bayes' formula is a rational method for adjusting our viewpoints as we confront new information.[17] Bayes' formula and related concepts have been applied in many business and investment decision-making contexts, including the evaluation of mutual fund performance.[18]

16 Otherwise, the calculation depends on conditional expected value; the calculation can be expressed as $E(XY) = E(X)\,E(Y \mid X)$.
17 Named after the Reverend Thomas Bayes (1702–61).
18 See Huij and Verbeek (2007).

Bayes' formula makes use of Equation 6, the total probability rule. To review, that rule expressed the probability of an event as a weighted average of the probabilities of the event, given a set of scenarios. Bayes' formula works in reverse; more precisely, it reverses the "given that" information. Bayes' formula uses the occurrence of the event to infer the probability of the scenario generating it. For that reason, Bayes' formula is sometimes called an inverse probability. In many applications, including the one illustrating its use in this section, an individual is updating his beliefs concerning the causes that may have produced a new observation.

- **Bayes' Formula.** Given a set of prior probabilities for an event of interest, if you receive new information, the rule for updating your probability of the event is

  $$\begin{aligned}&\text{Updated probability of event given the new information}\\&= \frac{\text{Probability of the new information given event}}{\text{Unconditional probability of the new information}} \times \text{Prior probability of event}\end{aligned}$$

  In probability notation, this formula can be written concisely as:

  $$P(\text{Event} \mid \text{Information}) = \frac{P(\text{Information} \mid \text{Event})}{P(\text{Information})} P(\text{Event})$$

To illustrate Bayes' formula, we work through an investment example that can be adapted to any actual problem. Suppose you are an investor in the stock of DriveMed, Inc. Positive earnings surprises relative to consensus EPS estimates often result in positive stock returns, and negative surprises often have the opposite effect. DriveMed is preparing to release last quarter's EPS result, and you are interested in which of these three events happened: *last quarter's EPS exceeded the consensus EPS estimate,* or *last quarter's EPS exactly met the consensus EPS estimate,* or *last quarter's EPS fell short of the consensus EPS estimate.* This list of the alternatives is mutually exclusive and exhaustive.

On the basis of your own research, you write down the following **prior probabilities** (or priors, for short) concerning these three events:

- $P(\textit{EPS exceeded consensus}) = 0.45$
- $P(\textit{EPS met consensus}) = 0.30$
- $P(\textit{EPS fell short of consensus}) = 0.25$

These probabilities are "prior" in the sense that they reflect only what you know now, before the arrival of any new information.

The next day, DriveMed announces that it is expanding factory capacity in Singapore and Ireland to meet increased sales demand. You assess this new information. The decision to expand capacity relates not only to current demand but probably also to the prior quarter's sales demand. You know that sales demand is positively related to EPS. So now it appears more likely that last quarter's EPS will exceed the consensus.

The question you have is, "In light of the new information, what is the updated probability that the prior quarter's EPS exceeded the consensus estimate?"

Bayes' formula provides a rational method for accomplishing this updating. We can abbreviate the new information as *DriveMed expands.* The first step in applying Bayes' formula is to calculate the probability of the new information (here: *DriveMed expands*), given a list of events or scenarios that may have generated it. The list of events should cover all possibilities, as it does here. Formulating these conditional probabilities is the key step in the updating process. Suppose your view is

$P(\textit{DriveMed expands} \mid \textit{EPS exceeded consensus}) = 0.75$

$P(\textit{DriveMed expands} \mid \textit{EPS met consensus}) = 0.20$

$P(\textit{DriveMed expands} \mid \textit{EPS fell short of consensus}) = 0.05$

Conditional probabilities of an observation (here: *DriveMed expands*) are sometimes referred to as **likelihoods**. Again, likelihoods are required for updating the probability.

Next, you combine these conditional probabilities or likelihoods with your prior probabilities to get the unconditional probability for DriveMed expanding, *P*(*DriveMed expands*), as follows:

$$
\begin{aligned}
&P(DriveMed\ expands)\\
&= P(DriveMed\ expands \mid EPS\ exceeded\ consensus)\\
&\quad \times P(EPS\ exceeded\ consensus)\\
&+ P(DriveMed\ expands \mid EPS\ met\ consensus)\\
&\quad \times P(EPS\ met\ consensus)\\
&+ P(DriveMed\ expands \mid EPS\ fell\ short\ of\ consensus)\\
&\quad \times P(EPS\ fell\ short\ of\ consensus)\\
&= 0.75(0.45) + 0.20(0.30) + 0.05(0.25) = 0.41\text{, or }41\%
\end{aligned}
$$

This is Equation 6, the total probability rule, in action. Now you can answer your question by applying Bayes' formula:

$$
\begin{aligned}
&P(EPS\ exceeded\ consensus \mid DriveMed\ expands)\\
&= \frac{P(DriveMed\ expands \mid EPS\ exceeded\ consensus)}{P(DriveMed\ expands)} P(EPS\ exceeded\ consensus)\\
&= (0.75/0.41)(0.45) = 1.829268(0.45) = 0.823171
\end{aligned}
$$

Prior to DriveMed's announcement, you thought the probability that DriveMed would beat consensus expectations was 45 percent. On the basis of your interpretation of the announcement, you update that probability to 82.3 percent. This updated probability is called your **posterior probability** because it reflects or comes after the new information.

The Bayes' calculation takes the prior probability, which was 45 percent, and multiplies it by a ratio—the first term on the right-hand side of the equal sign. The denominator of the ratio is the probability that DriveMed expands, as you view it without considering (conditioning on) anything else. Therefore, this probability is unconditional. The numerator is the probability that DriveMed expands, if last quarter's EPS actually exceeded the consensus estimate. This last probability is larger than unconditional probability in the denominator, so the ratio (1.83 roughly) is greater than 1. As a result, your updated or posterior probability is larger than your prior probability. Thus, the ratio reflects the impact of the new information on your prior beliefs.

EXAMPLE 13

Inferring whether DriveMed's EPS Met Consensus EPS

You are still an investor in DriveMed stock. To review the givens, your prior probabilities are *P*(*EPS exceeded consensus*) = 0.45, *P*(*EPS met consensus*) = 0.30, and *P*(*EPS fell short of consensus*) = 0.25. You also have the following conditional probabilities:

$$
\begin{aligned}
&P(DriveMed\ expands \mid EPS\ exceeded\ consensus) = 0.75\\
&P(DriveMed\ expands \mid EPS\ met\ consensus) = 0.20\\
&P(DriveMed\ expands \mid EPS\ fell\ short\ of\ consensus) = 0.05
\end{aligned}
$$

Recall that you updated your probability that last quarter's EPS exceeded the consensus estimate from 45 percent to 82.3 percent after DriveMed announced it would expand. Now you want to update your other priors.

1 Update your prior probability that DriveMed's EPS met consensus.

2 Update your prior probability that DriveMed's EPS fell short of consensus.

3 Show that the three updated probabilities sum to 1. (Carry each probability to four decimal places.)

4 Suppose, because of lack of prior beliefs about whether DriveMed would meet consensus, you updated on the basis of prior probabilities that all three possibilities were equally likely: $P(\textit{EPS exceeded consensus}) = P(\textit{EPS met consensus}) = P(\textit{EPS fell short of consensus}) = 1/3$. What is your estimate of the probability $P(\textit{EPS exceeded consensus} \mid \textit{DriveMed expands})$?

Solution to 1:

The probability is $P(\textit{EPS met consensus} \mid \textit{DriveMed expands}) =$

$$\frac{P(\textit{DriveMed expands} \mid \textit{EPS met consensus})}{P(\textit{DriveMed expands})} P(\textit{EPS met consensus})$$

The probability $P(\textit{DriveMed expands})$ is found by taking each of the three conditional probabilities in the statement of the problem, such as $P(\textit{DriveMed expands} \mid \textit{EPS exceeded consensus})$; multiplying each one by the prior probability of the conditioning event, such as $P(\textit{EPS exceeded consensus})$; then adding the three products. The calculation is unchanged from the problem in the text above: $P(\textit{DriveMed expands}) = 0.75(0.45) + 0.20(0.30) + 0.05(0.25) = 0.41$, or 41 percent. The other probabilities needed, $P(\textit{DriveMed expands} \mid \textit{EPS met consensus}) = 0.20$ and $P(\textit{EPS met consensus}) = 0.30$, are givens. So

$$\begin{aligned} &P(\textit{EPS met consensus} \mid \textit{DriveMed expands}) \\ &\quad = [P(\textit{DriveMed expands} \mid \textit{EPS met consensus})/P(\textit{DriveMed expands})] \\ &\qquad P(\textit{EPS met consensus}) \\ &\quad = (0.20/0.41)(0.30) = 0.487805(0.30) = 0.146341 \end{aligned}$$

After taking account of the announcement on expansion, your updated probability that last quarter's EPS for DriveMed just met consensus is 14.6 percent compared with your prior probability of 30 percent.

Solution to 2:

$P(\textit{DriveMed expands})$ was already calculated as 41 percent. Recall that $P(\textit{DriveMed expands} \mid \textit{EPS fell short of consensus}) = 0.05$ and $P(\textit{EPS fell short of consensus}) = 0.25$ are givens.

$$\begin{aligned} &P(\textit{EPS fell short of consensus} \mid \textit{DriveMed expands}) \\ &\quad = \left[P(\textit{DriveMed expands} \mid \textit{EPS fell short of consensus}) / \right. \\ &\qquad \left. P(\textit{DriveMed expands})\right] P(\textit{EPS fell short of consensus}) \\ &\quad = (0.05 / 0.41)(0.25) = 0.121951(0.25) = 0.030488 \end{aligned}$$

As a result of the announcement, you have revised your probability that DriveMed's EPS fell short of consensus from 25 percent (your prior probability) to 3 percent.

Solution to 3:

The sum of the three updated probabilities is

$$P(EPS\ exceeded\ consensus \mid DriveMed\ expands) + P(EPS\ met\ consensus \mid DriveMed\ expands) + P(EPS\ fell\ short\ of\ consensus \mid DriveMed\ expands) = 0.8232 + 0.1463 + 0.0305 = 1.0000$$

The three events (*EPS exceeded consensus*, *EPS met consensus*, *EPS fell short of consensus*) are mutually exclusive and exhaustive: One of these events or statements must be true, so the conditional probabilities must sum to 1. Whether we are talking about conditional or unconditional probabilities, whenever we have a complete set of the distinct possible events or outcomes, the probabilities must sum to 1. This calculation serves as a check on your work.

Solution to 4:

Using the probabilities given in the question,

$$\begin{aligned} &P(DriveMed\ expands) \\ &= P(DriveMed\ expands \mid EPS\ exceeded\ consensus) \\ &\quad P(EPS\ exceeded\ consensus) + P(DriveMed\ expands \mid \\ &\quad EPS\ met\ consensus) P(EPS\ met\ consensus) + P(DriveMed\ expands \mid \\ &\quad EPS\ fell\ short\ of\ consensus) P(EPS\ fell\ short\ of\ consensus) \\ &= 0.75(1/3) + 0.20(1/3) + 0.05(1/3) = 1/3 \end{aligned}$$

Not surprisingly, the probability of DriveMed expanding is 1/3 because the decision maker has no prior beliefs or views regarding how well EPS performed relative to the consensus estimate. Now we can use Bayes' formula to find *P*(*EPS exceeded consensus* | *DriveMed expands*) = [*P*(*DriveMed expands* | *EPS exceeded consensus*)/*P*(*DriveMed expands*)] *P*(*EPS exceeded consensus*) = [(0.75/(1/3)](1/3) = 0.75 or 75 percent. This probability is identical to your estimate of *P*(*DriveMed expands* | *EPS exceeded consensus*).

When the prior probabilities are equal, the probability of information given an event equals the probability of the event given the information. When a decision-maker has equal prior probabilities (called **diffuse priors**), the probability of an event is determined by the information.

4.2 Principles of Counting

The first step in addressing a question often involves determining the different logical possibilities. We may also want to know the number of ways that each of these possibilities can happen. In the back of our mind is often a question about probability. How likely is it that I will observe this particular possibility? Records of success and failure are an example. When we evaluate a market timer's record, one well-known evaluation method uses counting methods presented in this section.[19] An important investment model, the binomial option pricing model, incorporates the combination formula that we will cover shortly. We can also use the methods in this section to calculate what we called a priori probabilities in Section 2. When we can assume that the possible outcomes of a random variable are equally likely, the probability of an event equals the number of possible outcomes favorable for the event divided by the total number of outcomes.

19 Henriksson and Merton (1981).

In counting, enumeration (counting the outcomes one by one) is of course the most basic resource. What we discuss in this section are shortcuts and principles. Without these shortcuts and principles, counting the total number of outcomes can be very difficult and prone to error. The first and basic principle of counting is the multiplication rule.

- **Multiplication Rule of Counting.** If one task can be done in n_1 ways, and a second task, given the first, can be done in n_2 ways, and a third task, given the first two tasks, can be done in n_3 ways, and so on for k tasks, then the number of ways the k tasks can be done is $(n_1)(n_2)(n_3) \ldots (n_k)$.

Suppose we have three steps in an investment decision process. The first step can be done in two ways, the second in four ways, and the third in three ways. Following the multiplication rule, there are (2)(4)(3) = 24 ways in which we can carry out the three steps.

Another illustration is the assignment of members of a group to an equal number of positions. For example, suppose you want to assign three security analysts to cover three different industries. In how many ways can the assignments be made? The first analyst may be assigned in three different ways. Then two industries remain. The second analyst can be assigned in two different ways. Then one industry remains. The third and last analyst can be assigned in only one way. The total number of different assignments equals (3)(2)(1) = 6. The compact notation for the multiplication we have just performed is 3! (read: 3 factorial). If we had n analysts, the number of ways we could assign them to n tasks would be

$$n! = n(n-1)(n-2)(n-3)\ldots 1$$

or ***n* factorial**. (By convention, 0! = 1.) To review, in this application we repeatedly carry out an operation (here, job assignment) until we use up all members of a group (here, three analysts). With n members in the group, the multiplication formula reduces to n factorial.[20]

The next type of counting problem can be called labeling problems.[21] We want to give each object in a group a label, to place it in a category. The following example illustrates this type of problem.

A mutual fund guide ranked 18 bond mutual funds by total returns for the year 2014. The guide also assigned each fund one of five risk labels: *high risk* (four funds), *above-average risk* (four funds), *average risk* (three funds), *below-average risk* (four funds), and *low risk* (three funds); as 4 + 4 + 3 + 4 + 3 = 18, all the funds are accounted for. How many different ways can we take 18 mutual funds and label 4 of them high risk, 4 above-average risk, 3 average risk, 4 below-average risk, and 3 low risk, so that each fund is labeled?

The answer is close to 13 billion. We can label any of 18 funds *high risk* (the first slot), then any of 17 remaining funds, then any of 16 remaining funds, then any of 15 remaining funds (now we have 4 funds in the *high risk* group); then we can label any of 14 remaining funds *above-average risk*, then any of 13 remaining funds, and so forth. There are 18! possible sequences. However, order of assignment within a category does not matter. For example, whether a fund occupies the first or third slot of the four funds labeled *high risk*, the fund has the same label (*high risk*). Thus there are 4! ways to assign a given group of four funds to the four *high risk* slots. Making the same argument for the other categories, in total there are (4!)(4!)(3!)(4!)(3!) equivalent

20 The shortest explanation of n factorial is that it is the number of ways to order n objects in a row. In all the problems to which we apply this counting method, we must use up all the members of a group (sampling without replacement).

21 This discussion follows Kemeny, Schleifer, Snell, and Thompson (1972) in terminology and approach.

sequences. To eliminate such redundancies from the 18! total, we divide 18! by (4!)(4!)(3!)(4!)(3!). We have 18!/(4!)(4!)(3!)(4!)(3!) = 18!/(24)(24)(6)(24)(6) = 12,864,852,000. This procedure generalizes as follows.

- **Multinomial Formula (General Formula for Labeling Problems).** The number of ways that n objects can be labeled with k different labels, with n_1 of the first type, n_2 of the second type, and so on, with $n_1 + n_2 + \ldots + n_k = n$, is given by

$$\frac{n!}{n_1!n_2!\ldots n_k!}$$

The multinomial formula with two different labels ($k = 2$) is especially important. This special case is called the combination formula. A **combination** is a listing in which the order of the listed items does not matter. We state the combination formula in a traditional way, but no new concepts are involved. Using the notation in the formula below, the number of objects with the first label is $r = n_1$ and the number with the second label is $n - r = n_2$ (there are just two categories, so $n_1 + n_2 = n$). Here is the formula:

- **Combination Formula (Binomial Formula).** The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does not matter, is

$$_{n}C_{r} = \binom{n}{r} = \frac{n!}{(n-r)!r!}$$

Here $_{n}C_{r}$ and $\binom{n}{r}$ are shorthand notations for $n!/(n - r)!r!$ (read: n choose r, or n combination r).

If we label the r objects as *belongs to the group* and the remaining objects as *does not belong to the group*, whatever the group of interest, the combination formula tells us how many ways we can select a group of size r. We can illustrate this formula with the binomial option pricing model. This model describes the movement of the underlying asset as a series of moves, price up (U) or price down (D). For example, two sequences of five moves containing three up moves, such as UUUDD and UDUUD, result in the same final stock price. At least for an option with a payoff dependent on final stock price, the number but not the order of up moves in a sequence matters. How many sequences of five moves *belong to the group with three up moves?* The answer is 10, calculated using the combination formula ("5 choose 3"):

$$\begin{aligned} {}_{5}C_{3} &= 5!/(5-3)!3! \\ &= (5)(4)(3)(2)(1)/(2)(1)(3)(2)(1) = 120/12 = 10 \text{ ways} \end{aligned}$$

A useful fact can be illustrated as follows: $_{5}C_{3} = 5!/2!3!$ equals $_{5}C_{2} = 5!/3!2!$, as 3 + 2 = 5; $_{5}C_{4} = 5!/1!4!$ equals $_{5}C_{1} = 5!/4!1!$, as 4 + 1 = 5. This symmetrical relationship can save work when we need to calculate many possible combinations.

Suppose jurors want to select three companies out of a group of five to receive the first-, second-, and third-place awards for the best annual report. In how many ways can the jurors make the three awards? Order does matter if we want to distinguish among the three awards (the rank within the group of three); clearly the question makes order important. On the other hand, if the question were "In how many ways can the jurors choose three winners, without regard to place of finish?" we would use the combination formula.

To address the first question above, we need to count ordered listings such as *first place*, *New Company*; *second place*, *Fir Company*; *third place*, *Well Company*. An ordered listing is known as a **permutation**, and the formula that counts the number of permutations is known as the permutation formula.[22]

- **Permutation Formula.** The number of ways that we can choose r objects from a total of n objects, when the order in which the r objects are listed does matter, is

$$ {}_nP_r = \frac{n!}{(n-r)!} $$

So the jurors have ${}_5P_3 = 5!/(5-3)! = (5)(4)(3)(2)(1)/(2)(1) = 120/2 = 60$ ways in which they can make their awards. To see why this formula works, note that (5)(4)(3)(2)(1)/(2)(1) reduces to (5)(4)(3), after cancellation of terms. This calculation counts the number of ways to fill three slots choosing from a group of five people, according to the multiplication rule of counting. This number is naturally larger than it would be if order did not matter (compare 60 to the value of 10 for "5 choose 3" that we calculated above). For example, *first place*, *Well Company*; *second place*, *Fir Company*; *third place*, *New Company* contains the same three companies as *first place*, *New Company*; *second place*, *Fir Company*; *third place*, *Well Company*. If we were concerned only with award winners (without regard to place of finish), the two listings would count as one combination. But when we are concerned with the order of finish, the listings count as two permutations.

Answering the following questions may help you apply the counting methods we have presented in this section.

1. Does the task that I want to measure have a finite number of possible outcomes? If the answer is yes, you may be able to use a tool in this section, and you can go to the second question. If the answer is no, the number of outcomes is infinite, and the tools in this section do not apply.
2. Do I want to assign every member of a group of size n to one of n slots (or tasks)? If the answer is yes, use n factorial. If the answer is no, go to the third question.
3. Do I want to count the number of ways to apply one of three or more labels to each member of a group? If the answer is yes, use the multinomial formula. If the answer is no, go to the fourth question.
4. Do I want to count the number of ways that I can choose r objects from a total of n, when the order in which I list the r objects does not matter (can I give the r objects a label)? If the answer to these questions is yes, the combination formula applies. If the answer is no, go to the fifth question.
5. Do I want to count the number of ways I can choose r objects from a total of n, when the order in which I list the r objects is important? If the answer is yes, the permutation formula applies. If the answer is no, go to question 6.
6. Can the multiplication rule of counting be used? If it cannot, you may have to count the possibilities one by one, or use more advanced techniques than those presented here.[23]

22 A more formal definition states that a permutation is an ordered subset of n distinct objects.
23 Feller (1957) contains a very full treatment of counting problems and solution methods.

SUMMARY

In this reading, we have discussed the essential concepts and tools of probability. We have applied probability, expected value, and variance to a range of investment problems.

- A random variable is a quantity whose outcome is uncertain.
- Probability is a number between 0 and 1 that describes the chance that a stated event will occur.
- An event is a specified set of outcomes of a random variable.
- Mutually exclusive events can occur only one at a time. Exhaustive events cover or contain all possible outcomes.
- The two defining properties of a probability are, first, that $0 \le P(E) \le 1$ (where $P(E)$ denotes the probability of an event E), and second, that the sum of the probabilities of any set of mutually exclusive and exhaustive events equals 1.
- A probability estimated from data as a relative frequency of occurrence is an empirical probability. A probability drawing on personal or subjective judgment is a subjective probability. A probability obtained based on logical analysis is an a priori probability.
- A probability of an event E, $P(E)$, can be stated as odds for $E = P(E)/[1 - P(E)]$ or odds against $E = [1 - P(E)]/P(E)$.
- Probabilities that are inconsistent create profit opportunities, according to the Dutch Book Theorem.
- A probability of an event *not* conditioned on another event is an unconditional probability. The unconditional probability of an event A is denoted $P(A)$. Unconditional probabilities are also called marginal probabilities.
- A probability of an event given (conditioned on) another event is a conditional probability. The probability of an event A given an event B is denoted $P(A \mid B)$.
- The probability of both A and B occurring is the joint probability of A and B, denoted $P(AB)$.
- $P(A \mid B) = P(AB)/P(B)$, $P(B) \neq 0$.
- The multiplication rule for probabilities is $P(AB) = P(A \mid B)P(B)$.
- The probability that A or B occurs, or both occur, is denoted by $P(A \text{ or } B)$.
- The addition rule for probabilities is $P(A \text{ or } B) = P(A) + P(B) - P(AB)$.
- When events are independent, the occurrence of one event does not affect the probability of occurrence of the other event. Otherwise, the events are dependent.
- The multiplication rule for independent events states that if A and B are independent events, $P(AB) = P(A)P(B)$. The rule generalizes in similar fashion to more than two events.
- According to the total probability rule, if $S_1, S_2, \ldots, S_n$ are mutually exclusive and exhaustive scenarios or events, then $P(A) = P(A \mid S_1)P(S_1) + P(A \mid S_2)P(S_2) + \ldots + P(A \mid S_n)P(S_n)$.
- The expected value of a random variable is a probability-weighted average of the possible outcomes of the random variable. For a random variable X, the expected value of X is denoted $E(X)$.
- The total probability rule for expected value states that $E(X) = E(X \mid S_1)P(S_1) + E(X \mid S_2)P(S_2) + \ldots + E(X \mid S_n)P(S_n)$, where $S_1, S_2, \ldots, S_n$ are mutually exclusive and exhaustive scenarios or events.

- The variance of a random variable is the expected value (the probability-weighted average) of squared deviations from the random variable's expected value $E(X)$: $\sigma^2(X) = E\{[X - E(X)]^2\}$, where $\sigma^2(X)$ stands for the variance of X.
- Variance is a measure of dispersion about the mean. Increasing variance indicates increasing dispersion. Variance is measured in squared units of the original variable.
- Standard deviation is the positive square root of variance. Standard deviation measures dispersion (as does variance), but it is measured in the same units as the variable.
- Covariance is a measure of the co-movement between random variables.
- The covariance between two random variables R_i and R_j is the expected value of the cross-product of the deviations of the two random variables from their respective means: $\text{Cov}(R_i,R_j) = E\{[R_i - E(R_i)][R_j - E(R_j)]\}$. The covariance of a random variable with itself is its own variance.
- Correlation is a number between −1 and +1 that measures the co-movement (linear association) between two random variables: $\rho(R_i,R_j) = \text{Cov}(R_i,R_j)/[\sigma(R_i)\sigma(R_j)]$.
- To calculate the variance of return on a portfolio of n assets, the inputs needed are the n expected returns on the individual assets, n variances of return on the individual assets, and $n(n - 1)/2$ distinct covariances.
- Portfolio variance of return is $\sigma^2\left(R_p\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \text{Cov}\left(R_i, R_j\right)$.
- The calculation of covariance in a forward-looking sense requires the specification of a joint probability function, which gives the probability of joint occurrences of values of the two random variables.
- When two random variables are independent, the joint probability function is the product of the individual probability functions of the random variables.
- Bayes' formula is a method for updating probabilities based on new information.
- Bayes' formula is expressed as follows: Updated probability of event given the new information = [(Probability of the new information given event)/(Unconditional probability of the new information)] × Prior probability of event.
- The multiplication rule of counting says, for example, that if the first step in a process can be done in 10 ways, the second step, given the first, can be done in 5 ways, and the third step, given the first two, can be done in 7 ways, then the steps can be carried out in (10)(5)(7) = 350 ways.
- The number of ways to assign every member of a group of size n to n slots is $n! = n(n - 1)(n - 2)(n - 3) \ldots 1$. (By convention, $0! = 1$.)
- The number of ways that n objects can be labeled with k different labels, with n_1 of the first type, n_2 of the second type, and so on, with $n_1 + n_2 + \ldots + n_k = n$, is given by $n!/(n_1!n_2! \ldots n_k!)$. This expression is the multinomial formula.
- A special case of the multinomial formula is the combination formula. The number of ways to choose r objects from a total of n objects, when the order in which the r objects are listed does not matter, is

$$_nC_r = \binom{n}{r} = \frac{n!}{(n-r)!r!}$$

- The number of ways to choose r objects from a total of n objects, when the order in which the r objects are listed does matter, is

$$ {}_nP_r = \frac{n!}{(n-r)!} $$

This expression is the permutation formula.

REFERENCES

Bodie, Zvi, Alex Kane, and Alan J. Marcus. 2017. *Essentials of Investments*, 11th edition. New York: McGraw-Hill Irwin.

Elton, Edwin J., Martin J. Gruber, Stephen J. Brown, and William N. Goetzmann. 2013. *Modern Portfolio Theory and Investment Analysis*, 9th edition. Hoboken, NJ: Wiley.

Feller, William. 1957. *An Introduction to Probability Theory and Its Applications, Vol. I*, 2nd edition. New York: Wiley.

Henriksson, Roy D., and Robert C. Merton. 1981. "On Market Timing and Investment Performance, II. Statistical Procedures for Evaluating Forecasting Skills." *Journal of Business*, vol. 54, no. 4:513–533.

Huij, Joop, and Marno Verbeek. 2007. "Cross-sectional learning and short-run persistence in mutual fund performance." *Journal of Banking & Finance*, vol. 31:973–997.

Kemeny, John G., Arthur Schleifer, Jr, J. Laurie Snell, and Gerald L. Thompson. 1972. *Finite Mathematics with Business Applications*, 2nd edition. Englewood Cliffs, NJ: Prentice-Hall.

Lo, Andrew W. 1999. "The Three P's of Total Risk Management." *Financial Analysts Journal*, vol. 55, no. 1:13–26.

Ramsey, Frank P. 1931. "Truth and Probability." In *The Foundations of Mathematics and Other Logical Essays*, edited by R.B. Braithwaite. London: Routledge and Keegan Paul.

Reilly, Frank K., and Keith C. Brown. 2018. *Investment Analysis and Portfolio Management*, 11th edition. Mason, OH: Cengage South-Western.

Thanatawee, Yordying. 2013. "Ownership Structure with Dividend Policy: Evidence from Thailand." *International Journal of Economics and Finance*, vol. 5, no. 1:121–132.

Vidal-Garcia, Javier. 2013. "The Persistence of European Mutual Fund Performance." *Research in International Business and Finance*, vol. 28:45–67.

PRACTICE PROBLEMS

1 Suppose that 5 percent of the stocks meeting your stock-selection criteria are in the telecommunications (telecom) industry. Also, dividend-paying telecom stocks are 1 percent of the total number of stocks meeting your selection criteria. What is the probability that a stock is dividend paying, given that it is a telecom stock that has met your stock selection criteria?

2 You are using the following three criteria to screen potential acquisition targets from a list of 500 companies:

Criterion	Fraction of the 500 Companies Meeting the Criterion
Product lines compatible	0.20
Company will increase combined sales growth rate	0.45
Balance sheet impact manageable	0.78

If the criteria are independent, how many companies will pass the screen?

3 You apply both valuation criteria and financial strength criteria in choosing stocks. The probability that a randomly selected stock (from your investment universe) meets your valuation criteria is 0.25. Given that a stock meets your valuation criteria, the probability that the stock meets your financial strength criteria is 0.40. What is the probability that a stock meets both your valuation and financial strength criteria?

4 Suppose the prospects for recovering principal for a defaulted bond issue depend on which of two economic scenarios prevails. Scenario 1 has probability 0.75 and will result in recovery of \$0.90 per \$1 principal value with probability 0.45, or in recovery of \$0.80 per \$1 principal value with probability 0.55. Scenario 2 has probability 0.25 and will result in recovery of \$0.50 per \$1 principal value with probability 0.85, or in recovery of \$0.40 per \$1 principal value with probability 0.15.

A Compute the probability of each of the four possible recovery amounts: \$0.90, \$0.80, \$0.50, and \$0.40.

B Compute the expected recovery, given the first scenario.

C Compute the expected recovery, given the second scenario.

D Compute the expected recovery.

E Graph the information in a tree diagram.

5 You have developed a set of criteria for evaluating distressed credits. Companies that do not receive a passing score are classed as likely to go bankrupt within 12 months. You gathered the following information when validating the criteria:

- Forty percent of the companies to which the test is administered will go bankrupt within 12 months: $P(nonsurvivor) = 0.40$.
- Fifty-five percent of the companies to which the test is administered pass it: $P(pass\ test) = 0.55$.
- The probability that a company will pass the test given that it will subsequently survive 12 months, is 0.85: $P(pass\ test \mid survivor) = 0.85$.

A What is $P(pass\ test \mid nonsurvivor)$?

B Using Bayes' formula, calculate the probability that a company is a survivor, given that it passes the test; that is, calculate P(survivor | pass test).

C What is the probability that a company is a *nonsurvivor*, given that it fails the test?

D Is the test effective?

6 In probability theory, exhaustive events are *best* described as events:

A with a probability of zero.

B that are mutually exclusive.

C that include all potential outcomes.

7 Which probability estimate *most likely* varies greatly between people?

A An *a priori* probability

B An empirical probability

C A subjective probability

8 If the probability that Zolaf Company sales exceed last year's sales is 0.167, the odds for exceeding sales are *closest* to:

A 1 to 5.

B 1 to 6.

C 5 to 1.

9 The probability of an event given that another event has occurred is a:

A joint probability.

B marginal probability.

C conditional probability.

10 After estimating the probability that an investment manager will exceed his benchmark return in each of the next two quarters, an analyst wants to forecast the probability that the investment manager will exceed his benchmark return over the two-quarter period in total. Assuming that each quarter's performance is independent of the other, which probability rule should the analyst select?

A Addition rule

B Multiplication rule

C Total probability rule

11 Which of the following is a property of two dependent events?

A The two events must occur simultaneously.

B The probability of one event influences the probability of the other event.

C The probability of the two events occurring is the product of each event's probability.

12 Which of the following *best* describes how an analyst would estimate the expected value of a firm under the scenarios of bankruptcy and survivorship? The analyst would use:

A the addition rule.

B conditional expected values.

C the total probability rule for expected value.

13 An analyst developed two scenarios with respect to the recovery of $100,000 principal from defaulted loans:

Scenario	Probability of Scenario (%)	Amount Recovered ($)	Probability of Amount (%)
1	40	50,000	60
		30,000	40
2	60	80,000	90
		60,000	10

The amount of the expected recovery is *closest* to:

A \$36,400.

B \$63,600.

C \$81,600.

14 US and Spanish bonds have return standard deviations of 0.64 and 0.56, respectively. If the correlation between the two bonds is 0.24, the covariance of returns is *closest* to:

A 0.086.

B 0.670.

C 0.781.

15 The covariance of returns is positive when the returns on two assets tend to:

A have the same expected values.

B be above their expected value at different times.

C be on the same side of their expected value at the same time.

16 Which of the following correlation coefficients indicates the weakest linear relationship between two variables?

A –0.67

B –0.24

C 0.33

17 An analyst develops the following covariance matrix of returns:

	Hedge Fund	Market Index
Hedge fund	256	110
Market index	110	81

The correlation of returns between the hedge fund and the market index is *closest* to:

A 0.005.

B 0.073.

C 0.764.

18 All else being equal, as the correlation between two assets approaches +1.0, the diversification benefits:

A decrease.

B stay the same.

C increase.

19 Given a portfolio of five stocks, how many unique covariance terms, excluding variances, are required to calculate the portfolio return variance?

A 10

B 20

C 25

20 The probability distribution for a company's sales is:

Probability	Sales ($ millions)
0.05	70
0.70	40
0.25	25

The standard deviation of sales is *closest* to:

A $9.81 million.

B $12.20 million.

C $32.40 million.

21 Which of the following statements is *most* accurate? If the covariance of returns between two assets is 0.0023, then:

A the assets' risk is near zero.

B the asset returns are unrelated.

C the asset returns have a positive relationship.

22 An analyst produces the following joint probability function for a foreign index (FI) and a domestic index (DI).

	R_{DI} = 30%	R_{DI} = 25%	R_{DI} = 15%
R_{FI} = 25%	0.25		
R_{FI} = 15%		0.50	
R_{FI} = 10%			0.25

The covariance of returns on the foreign index and the returns on the domestic index is *closest* to:

A 26.39.

B 26.56.

C 28.12.

23 A manager will select 20 bonds out of his universe of 100 bonds to construct a portfolio. Which formula provides the number of possible portfolios?

A Permutation formula

B Multinomial formula

C Combination formula

24 A firm will select two of four vice presidents to be added to the investment committee. How many different groups of two are possible?

A 6

B 12

C 24

25 From an approved list of 25 funds, a portfolio manager wants to rank 4 mutual funds from most recommended to least recommended. Which formula is *most* appropriate to calculate the number of possible ways the funds could be ranked?

A Permutation formula

B Multinomial formula

C Combination formula

SOLUTIONS

1 Use Equation 1 to find this conditional probability: P(*stock is dividend paying* | *telecom stock that meets criteria*) = P(*stock is dividend paying and telecom stock that meets criteria*)/P(*telecom stock that meets criteria*) = 0.01/0.05 = 0.20.

2 According to the multiplication rule for independent events, the probability of a company meeting all three criteria is the product of the three probabilities. Labeling the event that a company passes the first, second, and third criteria, A, B, and C, respectively $P(ABC) = P(A)P(B)P(C) = (0.20)(0.45)(0.78) = 0.0702$. As a consequence, (0.0702)(500) = 35.10, so 35 companies pass the screen.

3 Use Equation 2, the multiplication rule for probabilities $P(AB) = P(A \mid B)P(B)$, defining A as the event that *a stock meets the financial strength criteria* and defining B as the event that *a stock meets the valuation criteria*. Then $P(AB) = P(A \mid B)P(B) = 0.40 \times 0.25 = 0.10$. The probability that a stock meets both the financial and valuation criteria is 0.10.

4 **A** *Outcomes associated with Scenario 1:* With a 0.45 probability of a \$0.90 recovery per \$1 principal value, given Scenario 1, and with the probability of Scenario 1 equal to 0.75, the probability of recovering \$0.90 is 0.45 (0.75) = 0.3375. By a similar calculation, the probability of recovering \$0.80 is 0.55(0.75) = 0.4125.

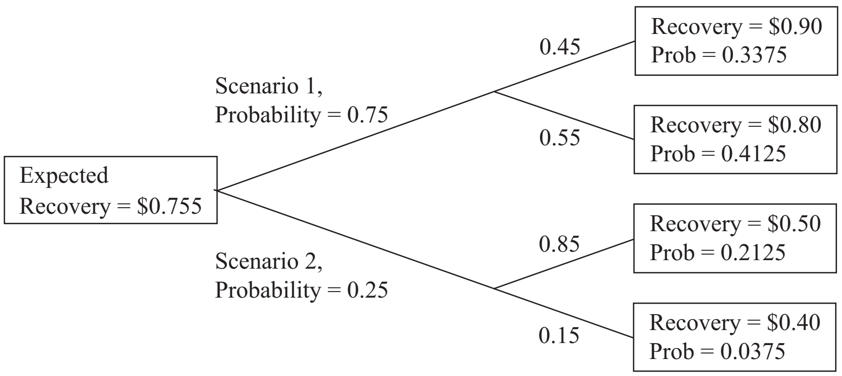
Outcomes associated with Scenario 2: With a 0.85 probability of a \$0.50 recovery per \$1 principal value, given Scenario 2, and with the probability of Scenario 2 equal to 0.25, the probability of recovering \$0.50 is 0.85(0.25) = 0.2125. By a similar calculation, the probability of recovering \$0.40 is 0.15(0.25) = 0.0375.

B E(*recovery* | *Scenario 1*) = 0.45(\$0.90) + 0.55(\$0.80) = \$0.845

C E(*recovery* | *Scenario 2*) = 0.85(\$0.50) + 0.15(\$0.40) = \$0.485

D E(*recovery*) = 0.75(\$0.845) + 0.25(\$0.485) = \$0.755

E



5 **A** We can set up the equation using the total probability rule:

$$P(pass\ test) = P(pass\ test \mid survivor)P(survivor) + P(pass\ test \mid nonsurvivor)P(nonsurvivor)$$

We know that P(*survivor*) = 1 – P(*nonsurvivor*) = 1 – 0.40 = 0.60. Therefore, P(*pass test*) = 0.55 = 0.85(0.60) + P(*pass test* | *nonsurvivor*)(0.40). Thus P(*pass test* | *nonsurvivor*) = [0.55 – 0.85(0.60)]/0.40 = 0.10.

B $$P(survivor \mid pass\ test) = \left[P(pass\ test \mid survivor)/P(pass\ test)\right]P(survivor)$$
$$= (0.85/0.55)0.60 = 0.927273$$

The information that a company passes the test causes you to update your probability that it is a survivor from 0.60 to approximately 0.927.

C According to Bayes' formula, $P(nonsurvivor \mid fail\ test) = [P(fail\ test \mid nonsurvivor)/ P(fail\ test)]P(nonsurvivor) = [P(fail\ test \mid nonsurvivor)/0.45]0.40$.

We can set up the following equation to obtain $P(fail\ test \mid nonsurvivor)$:

$$P(fail\ test) = P(fail\ test \mid nonsurvivor)P(nonsurvivor)$$
$$+P(fail\ test \mid survivor)P(survivor)$$
$$0.45 = P(fail\ test \mid nonsurvivor)0.40 + 0.15(0.60)$$

where $P(fail\ test \mid survivor) = 1 - P(pass\ test \mid survivor) = 1 - 0.85 = 0.15$. So $P(fail\ test \mid nonsurvivor) = [0.45 - 0.15(0.60)]/0.40 = 0.90$. Using this result with the formula above, we find $P(nonsurvivor \mid fail\ test) = (0.90/0.45)0.40 = 0.80$. Seeing that a company fails the test causes us to update the probability that it is a nonsurvivor from 0.40 to 0.80.

D A company passing the test greatly increases our confidence that it is a survivor. A company failing the test doubles the probability that it is a nonsurvivor. Therefore, the test appears to be useful.

6 C is correct. The term "exhaustive" means that the events cover all possible outcomes.

7 C is correct. A subjective probability draws on personal or subjective judgment that may be without reference to any particular data.

8 A is correct. Given odds for E of a to b, the implied probability of $E = a/(a + b)$. Stated in terms of odds a to b with $a = 1$, $b = 5$, the probability of $E = 1/(1 + 5) = 1/6 = 0.167$. This result confirms that a probability of 0.167 for beating sales is odds of 1 to 5.

9 C is correct. A conditional probability is the probability of an event given that another event has occurred.

10 B is correct. Because the events are independent, the multiplication rule is most appropriate for forecasting their joint probability. The multiplication rule for independent events states that the joint probability of both A and B occurring is P(AB) = P(A)P(B).

11 B is correct. The probability of the occurrence of one is related to the occurrence of the other. If we are trying to forecast one event, information about a dependent event may be useful.

12 C is correct. The total probability rule for expected value is used to estimate an expected value based on mutually exclusive and exhaustive scenarios.

13 B is correct. If Scenario 1 occurs, the expected recovery is 60% ($50,000) + 40% ($30,000) = $42,000, and if Scenario 2 occurs, the expected recovery is 90% ($80,000) + 10%($60,000) = $78,000. Weighting by the probability of each scenario, the expected recovery is 40%($42,000) + 60%($78,000) = $63,600. Alternatively, first calculating the probability of each amount occurring, the expected recovery is (40%)(60%)($50,000) + (40%)(40%)($30,000) + (60%)(90%)($80,000) + (60%)(10%)($60,000) = $63,600.

14 A is correct. The covariance is the product of the standard deviations and correlation using the formula Cov(US bond returns, Spanish bond returns) = σ(US bonds) × σ (Spanish bonds) × ρ(US bond returns, Spanish bond returns) = 0.64 × 0.56 × 0.24 = 0.086.

15 C is correct. The covariance of returns is positive when the returns on both assets tend to be on the same side (above or below) their expected values at the same time, indicating an average positive relationship between returns.

16 B is correct. Correlations near +1 exhibit strong positive linearity, whereas correlations near −1 exhibit strong negative linearity. A correlation of 0 indicates an absence of any linear relationship between the variables. The closer the correlation is to 0, the weaker the linear relationship.

17 C is correct. The correlation between two random variables R_i and R_j is defined as $\rho(R_i,R_j) = \text{Cov}(R_i,R_j)/\sigma(R_i)\sigma(R_j)$. Using the subscript i to represent hedge funds and the subscript j to represent the market index, the standard deviations are $\sigma(R_i) = 256^{1/2} = 16$ and $\sigma(R_j) = 81^{1/2} = 9$. Thus, $\rho(R_i,R_j) = \text{Cov}(R_i,R_j)/\sigma(R_i)\sigma(R_j) = 110/(16 \times 9) = 0.764$.

18 A is correct. As the correlation between two assets approaches +1, diversification benefits decrease. In other words, an increasingly positive correlation indicates an increasingly strong positive linear relationship and fewer diversification benefits.

19 A is correct. A covariance matrix for five stocks has 5 × 5 = 25 entries. Subtracting the 5 diagonal variance terms results in 20 off-diagonal entries. Because a covariance matrix is symmetrical, only 10 entries are unique (20/2 = 10).

20 A is correct. The analyst must first calculate expected sales as 0.05 × \$70 + 0.70 × \$40 + 0.25 × \$25 = \$3.50 million + \$28.00 million + \$6.25 million = \$37.75 million.

After calculating expected sales, we can calculate the variance of sales:

$$
\begin{aligned}
&= \sigma^2\ (\text{Sales}) \\
&= P(\$70)[\$70 - E(\text{Sales})]^2 + P(\$40)[\$40 - E(\text{Sales})]^2 + P(\$25) \\
&\quad [\$25 - E(\text{Sales})]^2 \\
&= 0.05(\$70 - 37.75)^2 + 0.70(\$40 - 37.75)^2 + 0.25(\$25 - 37.75)^2 \\
&= \$52.00 \text{ million} + \$3.54 \text{ million} + \$40.64 \text{ million} = \$96.18 \text{ million.}
\end{aligned}
$$

The standard deviation of sales is thus $\sigma = (\$96.18)^{1/2} = \$9.81$ million.

21 C is correct. The covariance of returns is positive when the returns on both assets tend to be on the same side (above or below) their expected values at the same time.

22 B is correct. The covariance is 26.56, calculated as follows. First, expected returns are

$$
\begin{aligned}
E(R_{FI}) &= (0.25 \times 25) + (0.50 \times 15) + (0.25 \times 10) \\
&= 6.25 + 7.50 + 2.50 = 16.25 \text{ and} \\
E(R_{DI}) &= (0.25 \times 30) + (0.50 \times 25) + (0.25 \times 15) \\
&= 7.50 + 12.50 + 3.75 = 23.75.
\end{aligned}
$$

Covariance is

$$
\begin{aligned}
\text{Cov}(R_{FI},R_{DI}) = &\sum_i \sum_j P\left(R_{FI,i}, R_{DI,j}\right)\left(R_{FI,i} - ER_{FI}\right)\left(R_{DI,j} - ER_{DI}\right) \\
&= 0.25[(25 - 16.25)(30 - 23.75)] + 0.50[(15 - 16.25) \\
&\quad (25 - 23.75)] + 0.25[(10 - 16.25)(15 - 23.75)] \\
&= 13.67 + (-0.78) + 13.67 = 26.56.
\end{aligned}
$$

23 C is correct. The combination formula provides the number of ways that r objects can be chosen from a total of n objects, when the order in which the r objects are listed does not matter. The order of the bonds within the portfolio does not matter.

24 A is correct. The answer is found using the combination formula

$$_nC_r = \binom{n}{r} = \frac{n!}{(n-r)!r!}$$

Here, $n = 4$ and $r = 2$, so the answer is $4!/[(4 - 2)!2!] = 24/[(2) \times (2)] = 6$. This result can be verified by assuming there are four vice presidents, VP1–VP4. The six possible additions to the investment committee are VP1 and VP2, VP1 and VP3, VP1 and VP4, VP2 and VP3, VP2 and VP4, and VP3 and VP4.

25 A is correct. The permutation formula is used to choose r objects from a total of n objects when order matters. Because the portfolio manager is trying to rank the four funds from most recommended to least recommended, the order of the funds matters; therefore, the permutation formula is most appropriate.

PRACTICE PROBLEMS

1 A European put option on stock conveys the right to sell the stock at a prespecified price, called the exercise price, at the maturity date of the option. The value of this put at maturity is (exercise price – stock price) or \$0, whichever is greater. Suppose the exercise price is \$100 and the underlying stock trades in ticks of \$0.01. At any time before maturity, the terminal value of the put is a random variable.

  A Describe the distinct possible outcomes for terminal put value. (Think of the put's maximum and minimum values and its minimum price increments.)

  B Is terminal put value, at a time before maturity, a discrete or continuous random variable?

  C Letting Y stand for terminal put value, express in standard notation the probability that terminal put value is less than or equal to \$24. No calculations or formulas are necessary.

2 Define the term "binomial random variable." Describe the types of problems for which the binomial distribution is used.

3 The value of the cumulative distribution function $F(x)$, where x is a particular outcome, for a discrete uniform distribution:

  A sums to 1.

  B lies between 0 and 1.

  C decreases as x increases.

4 For a binomial random variable with five trials, and a probability of success on each trial of 0.50, the distribution will be:

  A skewed.

  B uniform.

  C symmetric.

5 In a discrete uniform distribution with 20 potential outcomes of integers 1 to 20, the probability that X is greater than or equal to 3 but less than 6, $P(3 \leq X < 6)$, is:

  A 0.10.

  B 0.15.

  C 0.20.

6 Over the last 10 years, a company's annual earnings increased year over year seven times and decreased year over year three times. You decide to model the number of earnings increases for the next decade as a binomial random variable.

  A What is your estimate of the probability of success, defined as an increase in annual earnings?

  For Parts B, C, and D of this problem, assume the estimated probability is the actual probability for the next decade.

  B What is the probability that earnings will increase in exactly 5 of the next 10 years?

  C Calculate the expected number of yearly earnings increases during the next 10 years.

D Calculate the variance and standard deviation of the number of yearly earnings increases during the next 10 years.

E The expression for the probability function of a binomial random variable depends on two major assumptions. In the context of this problem, what must you assume about annual earnings increases to apply the binomial distribution in Part B? What reservations might you have about the validity of these assumptions?

7 A portfolio manager annually outperforms her benchmark 60% of the time. Assuming independent annual trials, what is the probability that she will outperform her benchmark four or more times over the next five years?

A 0.26

B 0.34

C 0.48

8 You are examining the record of an investment newsletter writer who claims a 70 percent success rate in making investment recommendations that are profitable over a one-year time horizon. You have the one-year record of the newsletter's seven most recent recommendations. Four of those recommendations were profitable. If all the recommendations are independent and the newsletter writer's skill is as claimed, what is the probability of observing four or fewer profitable recommendations out of seven in total?

9 You are forecasting sales for a company in the fourth quarter of its fiscal year. Your low-end estimate of sales is €14 million, and your high-end estimate is €15 million. You decide to treat all outcomes for sales between these two values as equally likely, using a continuous uniform distribution.

A What is the expected value of sales for the fourth quarter?

B What is the probability that fourth-quarter sales will be less than or equal to €14,125,000?

10 State the approximate probability that a normal random variable will fall within the following intervals:

A Mean plus or minus one standard deviation.

B Mean plus or minus two standard deviations.

C Mean plus or minus three standard deviations.

11 Find the area under the normal curve up to $z = 0.36$; that is, find $P(Z \leq 0.36)$. Interpret this value.

12 If an analyst expects a portfolio to outperform its benchmark with a 75% success rate in any measurement period, and the portfolio meets that objective in three of four quarters, what is the probability that the realized portfolio performance over the year is at or below this expectation?

A 0.26

B 0.42

C 0.68

13 In futures markets, profits or losses on contracts are settled at the end of each trading day. This procedure is called marking to market or daily resettlement. By preventing a trader's losses from accumulating over many days, marking to market reduces the risk that traders will default on their obligations. A futures markets trader needs a liquidity pool to meet the daily mark to market. If liquidity is exhausted, the trader may be forced to unwind his position at an unfavorable time.

Suppose you are using financial futures contracts to hedge a risk in your portfolio. You have a liquidity pool (cash and cash equivalents) of λ dollars per contract and a time horizon of T trading days. For a given size liquidity pool, λ, Kolb, Gay, and Hunter (1985) developed an expression for the probability stating that you will exhaust your liquidity pool within a T-day horizon as a result of the daily mark to market. Kolb et al. assumed that the expected change in futures price is 0 and that futures price changes are normally distributed. With σ representing the standard deviation of daily futures price changes, the standard deviation of price changes over a time horizon to day T is $\sigma\sqrt{T}$, given continuous compounding. With that background, the Kolb et al. expression is

$$\text{Probability of exhausting liquidity pool} = 2[1 - N(x)]$$

where $x = \lambda / \left(\sigma\sqrt{T}\right)$. Here x is a standardized value of λ. $N(x)$ is the standard normal cumulative distribution function. For some intuition about $1 - N(x)$ in the expression, note that the liquidity pool is exhausted if losses exceed the size of the liquidity pool at any time up to and including T; the probability of that event happening can be shown to be proportional to an area in the right tail of a standard normal distribution, $1 - N(x)$.

Using the Kolb et al. expression, answer the following questions:

A Your hedging horizon is five days, and your liquidity pool is $2,000 per contract. You estimate that the standard deviation of daily price changes for the contract is $450. What is the probability that you will exhaust your liquidity pool in the five-day period?

B Suppose your hedging horizon is 20 days, but all the other facts given in Part A remain the same. What is the probability that you will exhaust your liquidity pool in the 20-day period?

14 Which of the following is characteristic of the normal distribution?

A Asymmetry

B Kurtosis of 3

C Definitive limits or boundaries

15 Which of the following assets *most likely* requires the use of a multivariate distribution for modeling returns?

A A call option on a bond

B A portfolio of technology stocks

C A stock in a market index

16 The total number of parameters that fully characterizes a multivariate normal distribution for the returns on two stocks is:

A 3.

B 4.

C 5.

17 A client has a portfolio of common stocks and fixed-income instruments with a current value of £1,350,000. She intends to liquidate £50,000 from the portfolio at the end of the year to purchase a partnership share in a business. Furthermore, the client would like to be able to withdraw the £50,000 without reducing the initial capital of £1,350,000. The following table shows four alternative asset allocations.

Mean and Standard Deviation for Four Allocations (in Percent)				
	A	B	C	D
Expected annual return	16	12	10	9
Standard deviation of return	24	17	12	11

Address the following questions (assume normality for Parts B and C):

A Given the client's desire not to invade the £1,350,000 principal, what is the shortfall level, R_L? Use this shortfall level to answer Part B.

B According to the safety-first criterion, which of the allocations is the best?

C What is the probability that the return on the safety-first optimal portfolio will be less than the shortfall level, R_L?

Please refer to Exhibit 1 for Questions 18 and 19

Exhibit 1 *Z*-Table Values, $P(Z \le z) = N(z)$ for $z \ge 0$

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224

18 A portfolio has an expected mean return of 8 percent and standard deviation of 14 percent. The probability that its return falls between 8 and 11 percent is *closest* to:

A 8.3%

B 14.8%.

C 58.3%.

19 A portfolio has an expected return of 7% with a standard deviation of 13%. For an investor with a minimum annual return target of 4%, the probability that the portfolio return will fail to meet the target is *closest* to:

A 33%.

B 41%.

C 59%.

20 **A** Define Monte Carlo simulation and explain its use in finance.

B Compared with analytical methods, what are the strengths and weaknesses of Monte Carlo simulation for use in valuing securities?

21 A standard lookback call option on stock has a value at maturity equal to (Value of the stock at maturity – Minimum value of stock during the life of the option prior to maturity) or $0, whichever is greater. If the minimum value reached prior to maturity was $20.11 and the value of the stock at maturity is $23, for example, the call is worth $23 – $20.11 = $2.89. Briefly discuss how you might use Monte Carlo simulation in valuing a lookback call option.

22 Which of the following is a continuous random variable?

A The value of a futures contract quoted in increments of $0.05

B The total number of heads recorded in 1 million tosses of a coin

C The rate of return on a diversified portfolio of stocks over a three-month period

23 X is a discrete random variable with possible outcomes $X = \{1,2,3,4\}$. Three functions $f(x)$, $g(x)$, and $h(x)$ are proposed to describe the probabilities of the outcomes in X.

	Probability Function		
$X = x$	**$f(x) = P(X = x)$**	**$g(x) = P(X = x)$**	**$h(x) = P(X = x)$**
1	–0.25	0.20	0.20
2	0.25	0.25	0.25
3	0.50	0.50	0.30
4	0.25	0.05	0.35

The conditions for a probability function are satisfied by:

A $f(x)$.

B $g(x)$.

C $h(x)$.

24 The cumulative distribution function for a discrete random variable is shown in the following table.

$X = x$	**Cumulative Distribution Function $F(x) = P(X \le x)$**
1	0.15
2	0.25
3	0.50
4	0.60
5	0.95
6	1.00

The probability that X will take on a value of either 2 or 4 is *closest* to:

A 0.20.

B 0.35.

C 0.85.

25 Which of the following events can be represented as a Bernoulli trial?

A The flip of a coin

B The closing price of a stock

C The picking of a random integer between 1 and 10

26 The weekly closing prices of Mordice Corporation shares are as follows:

Date	Closing Price (€)
1 August	112
8 August	160
15 August	120

The continuously compounded return of Mordice Corporation shares for the period August 1 to August 15 is *closest to*:

A 6.90%

B 7.14%

C 8.95%

27 A stock is priced at $100.00 and follows a one-period binomial process with an up move that equals 1.05 and a down move that equals 0.97. If 1 million Bernoulli trials are conducted, and the average terminal stock price is $102.00, the probability of an up move (p) is *closest* to:

A 0.375.

B 0.500.

C 0.625.

28 A call option on a stock index is valued using a three-step binomial tree with an up move that equals 1.05 and a down move that equals 0.95. The current level of the index is $190, and the option exercise price is $200. If the option value is positive when the stock price exceeds the exercise price at expiration and $0 otherwise, the number of terminal nodes with a positive payoff is:

A one.

B two.

C three.

29 A random number between zero and one is generated according to a continuous uniform distribution. What is the probability that the first number generated will have a value of exactly 0.30?

A 0%

B 30%

C 70%

30 A Monte Carlo simulation can be used to:

A directly provide precise valuations of call options.

B simulate a process from historical records of returns.

C test the sensitivity of a model to changes in assumptions.

31 A limitation of Monte Carlo simulation is:

A its failure to do “what if” analysis.

B that it requires historical records of returns

C its inability to independently specify cause-and-effect relationships.

32 Which parameter equals zero in a normal distribution?

A Kurtosis

B Skewness

C Standard deviation

33 An analyst develops the following capital market projections.

	Stocks	Bonds
Mean Return	10%	2%
Standard Deviation	15%	5%

Assuming the returns of the asset classes are described by normal distributions, which of the following statements is correct?

A Bonds have a higher probability of a negative return than stocks.

B On average, 99% of stock returns will fall within two standard deviations of the mean.

C The probability of a bond return less than or equal to 3% is determined using a Z-score of 0.25.

34 A client holding a £2,000,000 portfolio wants to withdraw £90,000 in one year without invading the principal. According to Roy's safety-first criterion, which of the following portfolio allocations is optimal?

	Allocation A	Allocation B	Allocation C
Expected annual return	6.5%	7.5%	8.5%
Standard deviation of returns	8.35%	10.21%	14.34%

A Allocation A

B Allocation B

C Allocation C

35 In contrast to normal distributions, lognormal distributions:

A are skewed to the left.

B have outcomes that cannot be negative.

C are more suitable for describing asset returns than asset prices.

36 The lognormal distribution is a more accurate model for the distribution of stock prices than the normal distribution because stock prices are:

A symmetrical.

B unbounded.

C non-negative.

37 The price of a stock at $t = 0$ is $208.25 and at $t = 1$ is $186.75. The continuously compounded rate of return for the stock from $t = 0$ to $t = 1$ is *closest* to:

A –10.90%.

B –10.32%.

C 11.51%.

SOLUTIONS

1 **A** The put's minimum value is \$0. The put's value is \$0 when the stock price is at or above \$100 at the maturity date of the option. The put's maximum value is \$100 = \$100 (the exercise price) − \$0 (the lowest possible stock price). The put's value is \$100 when the stock is worthless at the option's maturity date. The put's minimum price increments are \$0.01. The possible outcomes of terminal put value are thus \$0.00, \$0.01, \$0.02, …, \$100.

B The price of the underlying has minimum price fluctuations of \$0.01: These are the minimum price fluctuations for terminal put value. For example, if the stock finishes at \$98.20, the payoff on the put is \$100 − \$98.20 = \$1.80. We can specify that the nearest values to \$1.80 are \$1.79 and \$1.81. With a continuous random variable, we cannot specify the nearest values. So, we must characterize terminal put value as a discrete random variable.

C The probability that terminal put value is less than or equal to \$24 is $P(Y \leq 24)$ or $F(24)$, in standard notation, where F is the cumulative distribution function for terminal put value.

2 A binomial random variable is defined as the number of successes in n Bernoulli trials (a trial that produces one of two outcomes). The binomial distribution is used to make probability statements about a record of successes and failures or about anything with binary (twofold) outcomes.

3 B is correct. The value of the cumulative distribution function lies between 0 and 1 for any x: $0 \leq F(x) \leq 1$.

4 C is correct. The binomial distribution is symmetric when the probability of success on a trial is 0.50, but it is asymmetric or skewed otherwise. Here it is given that $p = 0.50$

5 B is correct. The probability of any outcome is 0.05, $P(1) = 1/20 = 0.05$. The probability that X is greater than or equal to 3 but less than 6, which is expressed as $P(3 \leq X < 6) = P(3) + P(4) + P(5) = 0.05 + 0.05 + 0.05 = 0.15$

6 **A** The probability of an earnings increase (success) in a year is estimated as 7/10 = 0.70 or 70 percent, based on the record of the past 10 years.

B The probability that earnings will increase in 5 out of the next 10 years is about 10.3 percent. Define a binomial random variable X, counting the number of earnings increases over the next 10 years. From Part A, the probability of an earnings increase in a given year is $p = 0.70$ and the number of trials (years) is $n = 10$. Equation 1 gives the probability that a binomial random variable has x successes in n trials, with the probability of success on a trial equal to p.

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x}$$

For this example,

$$\binom{10}{5} 0.7^5 0.3^{10-5} = \frac{10!}{(10-5)!5!} 0.7^5 0.3^{10-5}$$
$$= 252 \times 0.16807 \times 0.00243 = 0.102919$$

We conclude that the probability that earnings will increase in exactly 5 of the next 10 years is 0.1029, or approximately 10.3 percent.

C The expected number of yearly increases is $E(X) = np = 10 \times 0.70 = 7$.

D The variance of the number of yearly increases over the next 10 years is $\sigma^2 = np(1 - p) = 10 \times 0.70 \times 0.30 = 2.1$. The standard deviation is 1.449 (the positive square root of 2.1).

E You must assume that 1) the probability of an earnings increase (success) is constant from year to year and 2) earnings increases are independent trials. If current and past earnings help forecast next year's earnings, Assumption 2 is violated. If the company's business is subject to economic or industry cycles, neither assumption is likely to hold.

7 B is correct. To calculate the probability of 4 years of outperformance, use the formula:

$$p(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x}$$

Using this formula to calculate the probability in 4 of 5 years, $n = 5$, $x = 4$ and $p = 0.60$.

Therefore,

$$p(4) = \frac{5!}{(5-4)!4!} 0.6^4 (1-0.6)^{5-4} = [120/24](0.1296)(0.40) = 0.2592$$

$$p(5) = \frac{5!}{(5-5)!5!} 0.6^5 (1-0.6)^{5-5} = [120/120](0.0778)(1) = 0.0778$$

The probability of outperforming 4 or more times is $p(4) + p(5) = 0.2592 + 0.0778 = 0.3370$

8 The observed success rate is 4/7 = 0.571, or 57.1 percent. The probability of four or fewer successes is $F(4) = p(4) + p(3) + p(2) + p(1) + p(0)$, where $p(4)$, $p(3)$, $p(2)$, $p(1)$, and $p(0)$ are respectively the probabilities of 4, 3, 2, 1, and 0 successes, according to the binomial distribution with $n = 7$ and $p = 0.70$. We have

$p(4) = (7!/4!3!)(0.70^4)(0.30^3) = 35(0.006483) = 0.226895$

$p(3) = (7!/3!4!)(0.70^3)(0.30^4) = 35(0.002778) = 0.097241$

$p(2) = (7!/2!5!)(0.70^2)(0.30^5) = 21(0.001191) = 0.025005$

$p(1) = (7!/1!6!)(0.70^1)(0.30^6) = 7(0.000510) = 0.003572$

$p(0) = (7!/0!7!)(0.70^0)(0.30^7) = 1(0.000219) = 0.000219$

Summing all these probabilities, you conclude that $F(4) = 0.226895 + 0.097241 + 0.025005 + 0.003572 + 0.000219 = 0.352931$, or 35.3 percent.

9 A The expected value of fourth-quarter sales is €14,500,000, calculated as (€14,000,000 + €15,000,000)/2. With a continuous uniform random variable, the mean or expected value is the midpoint between the smallest and largest values. (See Example 7.)

B The probability that fourth-quarter sales will be less than €14,125,000 is 0.125 or 12.5 percent, calculated as (€14,125,000 – €14,000,000)/(€15,000,000 – €14,000,000).

10 A Approximately 68 percent of all outcomes of a normal random variable fall within plus or minus one standard deviation of the mean.

B Approximately 95 percent of all outcomes of a normal random variable fall within plus or minus two standard deviations of the mean.

C Approximately 99 percent of all outcomes of a normal random variable fall within plus or minus three standard deviations of the mean.

11 The area under the normal curve for $z = 0.36$ is 0.6406 or 64.06 percent. The following table presents an excerpt from the tables of the standard normal cumulative distribution function in the back of this volume. To locate $z = 0.36$, find 0.30 in the fourth row of numbers, then look at the column for 0.06 (the second decimal place of 0.36). The entry is 0.6406.

$P(Z \le x) = N(x)$ for $x \ge 0$ or $P(Z \le z) = N(z)$ for $z \ge 0$

x or *z*	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.10	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.20	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.30	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	**0.6406**	0.6443	0.6480	0.6517
0.40	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.50	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224

The interpretation of 64.06 percent for $z = 0.36$ is that 64.06 percent of observations on a standard normal random variable are smaller than or equal to the value 0.36. (So 100% – 64.06% = 35.94% of the values are greater than 0.36.)

12 C is correct. The probability that the performance is at or below the expectation is calculated by finding $F(3) = p(3) + p(2) + p(1)$ using the formula:

$$p(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x}$$

Using this formula,

$$p(3) = \frac{4!}{(4-3)!3!} 0.75^3 (1-0.75)^{4-3} = [24/6](0.42)(0.25) = 0.42$$

$$p(2) = \frac{4!}{(4-2)!2!} 0.75^2 (1-0.75)^{4-2} = [24/4](0.56)(0.06) = 0.20$$

$$p(1) = \frac{4!}{(4-1)!1!} 0.75^1 (1-0.75)^{4-1} = [24/6](0.75)(0.02) = 0.06$$

Therefore,

$F(3) = p(3) + p(2) + p(1) = 0.42 + 0.20 + 0.06 = 0.68$ or 68.0

13 **A** The probability of exhausting the liquidity pool is 4.7 percent. First calculate $x = \lambda / \left(\sigma\sqrt{T}\right) = \$2{,}000 / \left(\$450\sqrt{5}\right) = 1.987616$. We can round this value to 1.99 to use the standard normal tables in the back of this book. Using those tables, we find that $N(1.99) = 0.9767$. Thus, the probability of exhausting the liquidity pool is $2[1 - N(1.99)] = 2(1 - 0.9767) = 0.0466$ or about 4.7 percent.

B The probability of exhausting the liquidity pool is now 32.2 percent. The calculation follows the same steps as those in Part A. We calculate $x = \lambda / \left(\sigma\sqrt{T}\right) = \$2{,}000 / \left(\$450\sqrt{20}\right) = 0.993808$. We can round this value to 0.99 to use the standard normal tables in the back of this book. Using those tables, we find that $N(0.99) = 0.8389$. Thus, the probability of exhausting the liquidity pool is $2[1 - N(0.99)] = 2(1 - 0.8389) = 0.3222$ or about 32.2 percent. This is a substantial probability that you will run out of funds to meet mark to market.

In their paper, Kolb et al. call the probability of exhausting the liquidity pool the probability of ruin, a traditional name for this type of calculation.

14 B is correct. The normal distribution has a skewness of 0, a kurtosis of 3, and a mean, median and mode that are all equal.

15 B is correct. Multivariate distributions specify the probabilities for a group of related random variables. A portfolio of technology stocks represents a group of related assets. Accordingly, statistical interrelationships must be considered, resulting in the need to use a multivariate normal distribution.

16 C is correct. A bivariate normal distribution (two stocks) will have two means, two variances and one correlation. A multivariate normal distribution for the returns on n stocks will have n means, n variances and $n(n - 1)/2$ distinct correlations.

17 **A** Because £50,000/£1,350,000 is 3.7 percent, for any return less than 3.7 percent the client will need to invade principal if she takes out £50,000. So R_L = 3.7 percent.

B To decide which of the allocations is safety-first optimal, select the alternative with the highest ratio $[E(R_P) - R_L]/\sigma_P$:

Allocation A: $0.5125 = (16 - 3.7)/24$

Allocation B: $0.488235 = (12 - 3.7)/17$

Allocation C: $0.525 = (10 - 3.7)/12$

Allocation D: $0.481818 = (9 - 3.7)/11$

Allocation C, with the largest ratio (0.525), is the best alternative according to the safety-first criterion.

C To answer this question, note that $P(R_C < 3.7) = N(-0.525)$. We can round 0.525 to 0.53 for use with tables of the standard normal cdf. First, we calculate $N(-0.53) = 1 - N(0.53) = 1 - 0.7019 = 0.2981$ or about 30 percent. The safety-first optimal portfolio has a roughly 30 percent chance of not meeting a 3.7 percent return threshold.

18 A is correct. $P(8\% \leq \text{Portfolio return} \leq 11\%) = N(Z \text{ corresponding to } 11\%) - N(Z \text{ corresponding to } 8\%)$. For the first term, $Z = (11\% - 8\%)/14\% = 0.21$ approximately, and using the table of cumulative normal distribution given in the problem, $N(0.21) = 0.5832$. To get the second term immediately, note that 8 percent is the mean, and for the normal distribution 50 percent of the probability lies on either side of the mean. Therefore, $N(Z$ corresponding to 8%) must equal 50 percent. So $P(8\% \leq \text{Portfolio return} \leq 11\%) = 0.5832 - 0.50 = 0.0832$ or approximately 8.3 percent.

19 B is correct. There are three steps, which involve standardizing the portfolio return: First, subtract the portfolio mean return from each side of the inequality: $P(\text{Portfolio return} - 7\%) \leq 4\% - 7\%)$. Second, divide each side of the inequality by the standard deviation of portfolio return: $P[(\text{Portfolio return} - 7\%)/13\% \leq (4\% - 7\%)/13\%] = P(Z \leq -0.2308) = N(-0.2308)$. Third, recognize that on the left-hand side we have a standard normal variable, denoted by Z and $N(-x) = 1 - N(x)$. Rounding -0.2308 to -0.23 for use with the cumulative distribution function (cdf) table, we have $N(-0.23) = 1 - N(0.23) = 1 - 0.5910 = 0.409$, approximately 41 percent. The probability that the portfolio will underperform the target is about 41 percent.

20 A Elements that should appear in a definition of Monte Carlo simulation are that it makes use of a computer; that it is used to represent the operation of a complex system, or in some applications, to find an approximate solution to a problem; and that it involves the generation of a large number of random samples from a specified probability distribution. The exact wording can vary, but one definition follows:

Monte Carlo simulation in finance involves the use of a computer to represent the operation of a complex financial system. In some important applications, Monte Carlo simulation is used to find an approximate solution to a complex financial problem. An integral part of Monte Carlo simulation is the generation of a large number of random samples from a probability distribution.

B *Strengths.* Monte Carlo simulation can be used to price complex securities for which no analytic expression is available, particularly European-style options.

Weaknesses. Monte Carlo simulation provides only statistical estimates, not exact results. Analytic methods, when available, provide more insight into cause-and-effect relationships than does Monte Carlo simulation.

21 In the text, we described how we could use Monte Carlo simulation to value an Asian option, a complex European-style option. Just as we can calculate the average value of the stock over a simulation trial to value an Asian option, we can also calculate the minimum value of the stock over a simulation trial. Then, for a given simulation trial, we can calculate the terminal value of the call, given the minimum value of the stock for the simulation trial. We can then discount back this terminal value to the present to get the value of the call today ($t = 0$). The average of these $t = 0$ values over all simulation trials is the Monte Carlo simulated value of the lookback call option.

22 C is correct. The rate of return is a random variable because the future outcomes are uncertain, and it is continuous because it can take on an unlimited number of outcomes.

23 B is correct. The function $g(x)$ satisfies the conditions of a probability function. All of the values of $g(x)$ are between 0 and 1, and the values of $g(x)$ all sum to 1.

24 A is correct. The probability that X will take on a value of 4 or less is: $F(4) = P(X \leq 4) = p(1) + p(2) + p(3) + p(4) = 0.60$. The probability that X will take on a value of 3 or less is: $F(3) = P(X \leq 3) = p(1) + p(2) + \text{p}(3) = 0.50$. So, the probability that X will take on a value of 4 is: $F(4) - F(3) = p(4) = 0.10$. The probability of $X = 2$ can be found using the same logic: $F(2) - F(1) = p(2) = 0.25 - 0.15 = 0.10$. The probability of X taking on a value of 2 or 4 is: $p(2) + p(4) = 0.10 + 0.10 = 0.20$.

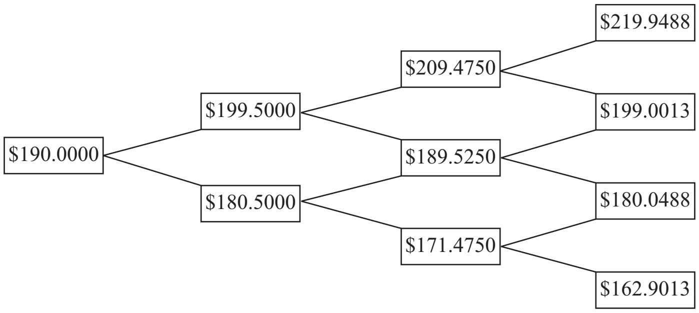
25 A is correct. A trial, such as a coin flip, will produce one of two outcomes. Such a trial is a Bernoulli trial.

26 A is correct. The continuously compounded return of an asset over a period is equal to the natural log of period's change. In this case:

$$\ln(120/112) = 6.90\%$$

27 C is correct. The probability of an up move (p) can be found by solving the equation: $(p)uS + (1 - p)dS = (p)105 + (1 - p)97 = 102$. Solving for p gives $8p = 5$, so that $p = 0.625$.

28 A is correct. Only the top node value of \$219.9488 exceeds \$200.



29 A is correct. The probability of generating a random number equal to any fixed point under a continuous uniform distribution is zero.

30 C is correct. A characteristic feature of Monte Carlo simulation is the generation of a large number of random samples from a specified probability distribution or distributions to represent the role of risk in the system.

31 C is correct. Monte Carlo simulation is a complement to analytical methods. Monte Carlo simulation provides statistical estimates and not exact results. Analytical methods, when available, provide more insight into cause-and-effect relationships

32 B is correct. A normal distribution has a skewness of zero (it is symmetrical around the mean). A non-zero skewness implies asymmetry in a distribution.

33 A is correct. The chance of a negative return falls in the area to the left of 0% under a standard normal curve. By standardizing the returns and standard deviations of the two assets, the likelihood of either asset experiencing a negative return may be determined: Z-score (standardized value) = $(X - \mu)/\sigma$

Z-score for a bond return of 0% = $(0 - 2)/5 = -0.40$.

Z-score for a stock return of 0% = $(0 - 10)/15 = -0.67$.

For bonds, a 0% return falls 0.40 standard deviations below the mean return of 2%. In contrast, for stocks, a 0% return falls 0.67 standard deviations below the mean return of 10%. A standard deviation of 0.40 is less than a standard deviation of 0.67. Negative returns thus occupy more of the left tail of the bond distribution than the stock distribution. Thus, bonds are more likely than stocks to experience a negative return.

34 B is correct. Allocation B has the highest safety-first ratio. The threshold return level R_L for the portfolio is £90,000/£2,000,000 = 4.5%, thus any return less than R_L = 4.5% will invade the portfolio principal. To compute the allocation that is safety-first optimal, select the alternative with the highest ratio:

$$\frac{\left[E\left(R_P - R_L\right)\right]}{\sigma_P}$$

$$\text{Allocation A} = \frac{6.5 - 4.5}{8.35} = 0.240$$

$$\text{Allocation B} = \frac{7.5 - 4.5}{10.21} = 0.294$$

$$\text{Allocation C} = \frac{8.5 - 4.5}{14.34} = 0.279$$

35 B is correct. By definition, lognormal random variables cannot have negative values.

36 C is correct. A lognormal distributed variable has a lower bound of zero. The lognormal distribution is also right skewed, which is a useful property in describing asset prices.

37 A is correct. The continuously compounded return from $t = 0$ to $t = 1$ is $r_{0,1} = \ln(S_1/S_0) = \ln(186.75/208.25) = -0.10897 = -10.90\%$.

PRACTICE PROBLEMS

1 Peter Biggs wants to know how growth managers performed last year. Biggs assumes that the population cross-sectional standard deviation of growth manager returns is 6 percent and that the returns are independent across managers.

A How large a random sample does Biggs need if he wants the standard deviation of the sample means to be 1 percent?

B How large a random sample does Biggs need if he wants the standard deviation of the sample means to be 0.25 percent?

2 Petra Munzi wants to know how value managers performed last year. Munzi estimates that the population cross-sectional standard deviation of value manager returns is 4 percent and assumes that the returns are independent across managers.

A Munzi wants to build a 95 percent confidence interval for the mean return. How large a random sample does Munzi need if she wants the 95 percent confidence interval to have a total width of 1 percent?

B Munzi expects a cost of about $10 to collect each observation. If she has a $1,000 budget, will she be able to construct the confidence interval she wants?

3 Assume that the equity risk premium is normally distributed with a population mean of 6 percent and a population standard deviation of 18 percent. Over the last four years, equity returns (relative to the risk-free rate) have averaged −2.0 percent. You have a large client who is very upset and claims that results this poor should *never* occur. Evaluate your client's concerns.

A Construct a 95 percent confidence interval around the population mean for a sample of four-year returns.

B What is the probability of a −2.0 percent or lower average return over a four-year period?

4 Compare the standard normal distribution and Student's t-distribution.

5 Find the reliability factors based on the t-distribution for the following confidence intervals for the population mean (df = degrees of freedom, n = sample size):

A A 99 percent confidence interval, df = 20.

B A 90 percent confidence interval, df = 20.

C A 95 percent confidence interval, n = 25.

D A 95 percent confidence interval, n = 16.

6 Assume that monthly returns are normally distributed with a mean of 1 percent and a sample standard deviation of 4 percent. The population standard deviation is unknown. Construct a 95 percent confidence interval for the sample mean of monthly returns if the sample size is 24.

7 Ten analysts have given the following fiscal year earnings forecasts for a stock:

Forecast (X_i)	Number of Analysts (n_i)
1.40	1
1.43	1
1.44	3

Forecast (X_i)	Number of Analysts (n_i)
1.45	2
1.47	1
1.48	1
1.50	1

Because the sample is a small fraction of the number of analysts who follow this stock, assume that we can ignore the finite population correction factor. Assume that the analyst forecasts are normally distributed.

A What are the mean forecast and standard deviation of forecasts?

B Provide a 95 percent confidence interval for the population mean of the forecasts.

8 Thirteen analysts have given the following fiscal-year earnings forecasts for a stock:

Forecast (X_i)	Number of Analysts (n_i)
0.70	2
0.72	4
0.74	1
0.75	3
0.76	1
0.77	1
0.82	1

Because the sample is a small fraction of the number of analysts who follow this stock, assume that we can ignore the finite population correction factor.

A What are the mean forecast and standard deviation of forecasts?

B What aspect of the data makes us uncomfortable about using t-tables to construct confidence intervals for the population mean forecast?

9 Explain the differences between constructing a confidence interval when sampling from a normal population with a known population variance and sampling from a normal population with an unknown variance.

10 An exchange rate has a given expected future value and standard deviation.

A Assuming that the exchange rate is normally distributed, what are the probabilities that the exchange rate will be at least 2 or 3 standard deviations away from its mean?

B Assume that you do not know the distribution of exchange rates. Use Chebyshev's inequality (that at least $1 - 1/k^2$ proportion of the observations will be within k standard deviations of the mean for any positive integer k greater than 1) to calculate the maximum probabilities that the exchange rate will be at least 2 or 3 standard deviations away from its mean.

11 Although he knows security returns are not independent, a colleague makes the claim that because of the central limit theorem, if we diversify across a large number of investments, the portfolio standard deviation will eventually approach zero as n becomes large. Is he correct?

12 Why is the central limit theorem important?

13 What is wrong with the following statement of the central limit theorem?

Central Limit Theorem. "If the random variables $X_1, X_2, X_3, \ldots, X_n$ are a random sample of size n from any distribution with finite mean μ and variance σ^2, then the distribution of $\overline{X}$ will be approximately normal, with a standard deviation of $\sigma/\sqrt{n}$."

14 Suppose we take a random sample of 30 companies in an industry with 200 companies. We calculate the sample mean of the ratio of cash flow to total debt for the prior year. We find that this ratio is 23 percent. Subsequently, we learn that the population cash flow to total debt ratio (taking account of all 200 companies) is 26 percent. What is the explanation for the discrepancy between the sample mean of 23 percent and the population mean of 26 percent?

- **A** Sampling error.
- **B** Bias.
- **C** A lack of consistency.

15 Alcorn Mutual Funds is placing large advertisements in several financial publications. The advertisements prominently display the returns of 5 of Alcorn's 30 funds for the past 1-, 3-, 5-, and 10-year periods. The results are indeed impressive, with all of the funds beating the major market indexes and a few beating them by a large margin. Is the Alcorn family of funds superior to its competitors?

16 Julius Spence has tested several predictive models in order to identify undervalued stocks. Spence used about 30 company-specific variables and 10 market-related variables to predict returns for about 5,000 North American and European stocks. He found that a final model using eight variables applied to telecommunications and computer stocks yields spectacular results. Spence wants you to use the model to select investments. Should you? What steps would you take to evaluate the model?

17 The *best* approach for creating a stratified random sample of a population involves:

- **A** drawing an equal number of simple random samples from each subpopulation.
- **B** selecting every *k*th member of the population until the desired sample size is reached.
- **C** drawing simple random samples from each subpopulation in sizes proportional to the relative size of each subpopulation.

18 A population has a non-normal distribution with mean μ and variance σ^2. The sampling distribution of the sample mean computed from samples of large size from that population will have:

- **A** the same distribution as the population distribution.
- **B** its mean approximately equal to the population mean.
- **C** its variance approximately equal to the population variance.

19 A sample mean is computed from a population with a variance of 2.45. The sample size is 40. The standard error of the sample mean is *closest* to:

- **A** 0.039.
- **B** 0.247.
- **C** 0.387.

20 An estimator with an expected value equal to the parameter that it is intended to estimate is described as:

A efficient.
B unbiased.
C consistent.

21 If an estimator is consistent, an increase in sample size will increase the:
A accuracy of estimates.
B efficiency of the estimator.
C unbiasedness of the estimator.

22 For a two-sided confidence interval, an increase in the degree of confidence will result in:
A a wider confidence interval.
B a narrower confidence interval.
C no change in the width of the confidence interval.

23 As the t-distribution's degrees of freedom decrease, the t-distribution *most likely*:
A exhibits tails that become fatter.
B approaches a standard normal distribution.
C becomes asymmetrically distributed around its mean value.

24 For a sample size of 17, with a mean of 116.23 and a variance of 245.55, the width of a 90% confidence interval using the appropriate t-distribution is *closest to*:
A 13.23.
B 13.27.
C 13.68.

25 For a sample size of 65 with a mean of 31 taken from a normally distributed population with a variance of 529, a 99% confidence interval for the population mean will have a lower limit *closest* to:
A 23.64.
B 25.41.
C 30.09.

26 An increase in sample size is *most likely* to result in a:
A wider confidence interval.
B decrease in the standard error of the sample mean.
C lower likelihood of sampling from more than one population.

27 A report on long-term stock returns focused exclusively on all currently publicly traded firms in an industry is *most likely* susceptible to:
A look-ahead bias.
B survivorship bias.
C intergenerational data mining.

28 Which sampling bias is *most likely* investigated with an out-of-sample test?
A Look-ahead bias
B Data-mining bias
C Sample selection bias

29 Which of the following characteristics of an investment study *most likely* indicates time-period bias?
A The study is based on a short time-series.

B Information not available on the test date is used.

C A structural change occurred prior to the start of the study's time series.

SOLUTIONS

1 **A** The standard deviation or standard error of the sample mean is $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. Substituting in the values for $\sigma_{\bar{X}}$ and σ, we have 1% = $6\%/\sqrt{n}$, or $\sqrt{n}$ = 6. Squaring this value, we get a random sample of n = 36.

B As in Part A, the standard deviation of sample mean is $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. Substituting in the values for $\sigma_{\bar{X}}$ and σ, we have 0.25% = $6\%/\sqrt{n}$, or $\sqrt{n}$ = 24. Squaring this value, we get a random sample of n = 576, which is substantially larger than for Part A of this question.

2 **A** Assume the sample size will be large and thus the 95 percent confidence interval for the mean of a sample of manager returns is $\bar{X} \pm 1.96 s_{\bar{X}}$, where $s_{\bar{X}} = s/\sqrt{n}$. Munzi wants the distance between the upper limit and lower limit in the confidence interval to be 1 percent, which is

$$\left(\bar{X} + 1.96 s_{\bar{X}}\right) - \left(\bar{X} - 1.96 s_{\bar{X}}\right) = 1\%$$

Simplifying this equation, we get $2\left(1.96 s_{\bar{X}}\right)$ = 1%. Finally, we have $3.92 s_{\bar{X}}$ = 1%, which gives us the standard deviation of the sample mean, $s_{\bar{X}}$ = 0.255%. The distribution of sample means is $s_{\bar{X}} = s/\sqrt{n}$. Substituting in the values for $s_{\bar{X}}$ and s, we have 0.255% = $4\%/\sqrt{n}$, or $\sqrt{n}$ = 15.69. Squaring this value, we get a random sample of n = 246.

B With her budget, Munzi can pay for a sample of up to 100 observations, which is far short of the 246 observations needed. Munzi can either proceed with her current budget and settle for a wider confidence interval or she can raise her budget (to around $2,460) to get the sample size for a 1 percent width in her confidence interval.

3 **A** This is a small-sample problem in which the sample comes from a normal population with a known standard deviation; thus we use the z-distribution in the solution. For a 95 percent confidence interval (and 2.5 percent in each tail), the critical z-value is 1.96. For returns that are normally distributed, a 95 percent confidence interval is of the form

$$\mu + 1.96\frac{\sigma}{\sqrt{n}}$$

The lower limit is $X_1 = \mu - 1.96\frac{\sigma}{\sqrt{n}} = 6\% - 1.96\frac{18\%}{\sqrt{4}}$ = 6% − 1.96(9%) = −11.64%.

The upper limit is $X_u = \mu + 1.96\frac{\sigma}{\sqrt{n}} = 6\% + 1.96\frac{18\%}{\sqrt{4}}$ = 6% + 1.96(9%) = 23.64%.

There is a 95 percent probability that four-year average returns will be between −11.64 percent and +23.64 percent.

B The critical z-value associated with the −2.0 percent return is

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{-2\% - 6\%}{18\%/\sqrt{4}} = \frac{-8\%}{9\%} = -0.89$$

Using a normal table, the probability of a z-value less than −0.89 is $P(Z < -0.89) = 0.1867$. Unfortunately, although your client is unhappy with the investment result, a four-year average return of −2.0 percent or lower should occur 18.67 percent of the time.

4 (Refer to Figure 1 to help visualize the answer to this question.) Basically, only one standard normal distribution exists, but many t-distributions exist—one for every different number of degrees of freedom. The normal distribution and the t-distribution for a large number of degrees of freedom are practically the same. The lower the degrees of freedom, the flatter the t-distribution becomes. The t-distribution has less mass (lower probabilities) in the center of the distribution and more mass (higher probabilities) out in both tails. Therefore, the confidence intervals based on t-values will be wider than those based on the normal distribution. Stated differently, the probability of being within a given number of standard deviations (such as within ±1 standard deviation or ±2 standard deviations) is lower for the t-distribution than for the normal distribution.

5 **A** For a 99 percent confidence interval, the reliability factor we use is $t_{0.005}$; for df = 20, this factor is 2.845.

B For a 90 percent confidence interval, the reliability factor we use is $t_{0.05}$; for df = 20, this factor is 1.725.

C Degrees of freedom equals $n - 1$, or in this case $25 - 1 = 24$. For a 95 percent confidence interval, the reliability factor we use is $t_{0.025}$; for df = 24, this factor is 2.064.

D Degrees of freedom equals $16 - 1 = 15$. For a 95 percent confidence interval, the reliability factor we use is $t_{0.025}$; for df = 15, this factor is 2.131.

6 Because this is a small sample from a normal population and we have only the sample standard deviation, we use the following model to solve for the confidence interval of the population mean:

$$\bar{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$$

where we find $t_{0.025}$ (for a 95 percent confidence interval) for df = $n - 1 = 24 - 1 = 23$; this value is 2.069. Our solution is $1\% \pm 2.069(4\%)/\sqrt{24} = 1\% \pm 2.069(0.8165) = 1\% \pm 1.69$. The 95 percent confidence interval spans the range from −0.69 percent to +2.69 percent.

7 The following table summarizes the calculations used in the answers.

Forecast (X_i)	Number of Analysts (n_i)	$X_i n_i$	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$(X_i - \bar{X})^2 n_i$
1.40	1	1.40	−0.05	0.0025	0.0025
1.43	1	1.43	−0.02	0.0004	0.0004
1.44	3	4.32	−0.01	0.0001	0.0003
1.45	2	2.90	0.00	0.0000	0.0000
1.47	1	1.47	0.02	0.0004	0.0004
1.48	1	1.48	0.03	0.0009	0.0009
1.50	1	1.50	0.05	0.0025	0.0025
Sums	10	14.50			0.0070

A With $n = 10, \bar{X} = \sum_{i=1}^{10} X_i / n = 14.50/10 = 1.45$. The variance is $s^2 = \left[\sum_{i=1}^{10}\left(X_i - \bar{X}\right)^2\right] \Big/ (n-1) = 0.0070/9 = 0.0007778$. The sample standard deviation is s = $\sqrt{0.0007778} = 0.02789$.

B The confidence interval for the mean can be estimated by using $\bar{X} \pm t_{\alpha/2}\left(s/\sqrt{n}\right)$. For 9 degrees of freedom, the reliability factor, $t_{0.025}$, equals 2.262 and the confidence interval is

$$1.45 \pm 2.262 \times 0.02789/\sqrt{10} = 1.45 \pm 2.262(0.00882)$$
$$= 1.45 \pm 0.02$$

The confidence interval for the population mean ranges from 1.43 to 1.47.

8 The following table summarizes the calculations used in the answers.

Forecast (X_i)	**Number of Analysts (n_i)**	$X_i n_i$	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$(X_i - \bar{X})^2 n_i$
0.70	2	1.40	–0.04	0.0016	0.0032
0.72	4	2.88	–0.02	0.0004	0.0016
0.74	1	0.74	0.00	0.0000	0.0000
0.75	3	2.25	0.01	0.0001	0.0003
0.76	1	0.76	0.02	0.0004	0.0004
0.77	1	0.77	0.03	0.0009	0.0009
0.82	1	0.82	0.08	0.0064	0.0064
Sums	13	9.62			0.0128

A With $n = 13, \bar{X} = \sum_{i=1}^{13} X_i / n = 9.62/13 = 0.74$. The variance is $s^2 = \left[\sum_{i=1}^{13}\left(X_i - \bar{X}\right)^2\right] \Big/ (n-1) = 0.0128/12 = 0.001067$. The sample standard deviation is $s = \sqrt{0.001067} = 0.03266$.

B The sample is small, and the distribution appears to be bimodal. We cannot compute a confidence interval for the population mean because we have probably sampled from a distribution that is not normal.

9 If the population variance is known, the confidence interval is

$$\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

The confidence interval for the population mean is centered at the sample mean, $\bar{X}$. The population standard deviation is σ, and the sample size is *n*. The population standard deviation divided by the square root of *n* is the standard error of the estimate of the mean. The value of *z* depends on the desired degree of confidence. For a 95 percent confidence interval, $z_{0.025} = 1.96$ and the confidence interval estimate is

$$\bar{X} \pm 1.96\frac{\sigma}{\sqrt{n}}$$

If the population variance is not known, we make two changes to the technique used when the population variance is known. First, we must use the sample standard deviation instead of the population standard deviation. Second, we use the t-distribution instead of the normal distribution. The critical t-value will depend on degrees of freedom $n - 1$. If the sample size is large, we have the alternative of using the z-distribution with the sample standard deviation.

10 **A** The probabilities can be taken from a normal table, in which the critical z-values are 2.00 or 3.00 and we are including the probabilities in both tails. The probabilities that the exchange rate will be at least 2 or 3 standard deviations away from the mean are

$$P\left(\left|X - \mu\right| \geq 2\sigma\right) = 0.0456$$

$$P\left(\left|X - \mu\right| \geq 3\sigma\right) = 0.0026$$

B With Chebyshev's inequality, the maximum probability of the exchange rate being at least k standard deviations from the mean is $P(| X - \mu | \geq k\sigma) \leq (1/k)^2$. The maximum probabilities of the rate being at least 2 or 3 standard deviations away from the mean are

$$P\left(\left|X - \mu\right| \geq 2\sigma\right) \leq \left(1/2\right)^2 = 0.2500$$

$$P\left(\left|X - \mu\right| \geq 3\sigma\right) \leq \left(1/3\right)^2 = 0.1111$$

The probability of the rate being outside 2 or 3 standard deviations of the mean is much smaller with a known normal distribution than when the distribution is unknown and we are relying on Chebyshev's inequality.

11 No. First the conclusion on the limit of zero is wrong; second, the support cited for drawing the conclusion (i.e., the central limit theorem) is not relevant in this context.

12 In many instances, the distribution that describes the underlying population is not normal or the distribution is not known. The central limit theorem states that if the sample size is large, regardless of the shape of the underlying population, the distribution of the sample mean is approximately normal. Therefore, even in these instances, we can still construct confidence intervals (and conduct tests of inference) as long as the sample size is large (generally $n \geq 30$).

13 The statement makes the following mistakes:

- Given the conditions in the statement, the distribution of $\bar{X}$ will be approximately normal only for large sample sizes.
- The statement omits the important element of the central limit theorem that the distribution of $\bar{X}$ will have mean μ.

14 A is correct. The discrepancy arises from sampling error. Sampling error exists whenever one fails to observe every element of the population, because a sample statistic can vary from sample to sample. As stated in the reading, the sample mean is an unbiased estimator, a consistent estimator, and an efficient estimator of the population mean. Although the sample mean is an unbiased estimator of the population mean—the expected value of the sample mean equals the population mean—because of sampling error, we do not expect the sample mean to exactly equal the population mean in any one sample we may take.

15 No, we cannot say that Alcorn Mutual Funds as a group is superior to competitors. Alcorn Mutual Funds' advertisement may easily mislead readers because the advertisement does not show the performance of all its funds. In particular, Alcorn Mutual Funds is engaging in sample selection bias by presenting the investment results from its best-performing funds only.

16 Spence may be guilty of data mining. He has used so many possible combinations of variables on so many stocks, it is not surprising that he found some instances in which a model worked. In fact, it would have been more surprising if he had not found any. To decide whether to use his model, you should do two things: First, ask that the model be tested on out-of-sample data—that is, data that were not used in building the model. The model may not be successful with out-of-sample data. Second, examine his model to make sure that the relationships in the model make economic sense, have a story, and have a future.

17 C is correct. Stratified random sampling involves dividing a population into subpopulations based on one or more classification criteria. Then, simple random samples are drawn from each subpopulation in sizes proportional to the relative size of each subpopulation. These samples are then pooled to form a stratified random sample.

18 B is correct. Given a population described by any probability distribution (normal or non-normal) with finite variance, the central limit theorem states that the sampling distribution of the sample mean will be approximately normal, with the mean approximately equal to the population mean, when the sample size is large.

19 B is correct. Taking the square root of the known population variance to determine the population standard deviation (σ) results in:

$$\sigma = \sqrt{2.45} = 1.565$$

The formula for the standard error of the sample mean (σ_X), based on a known sample size (n), is:

$$\sigma_X = \frac{\sigma}{\sqrt{n}}$$

Therefore,

$$\sigma_X = \frac{1.565}{\sqrt{40}} = 0.247$$

20 B is correct. An unbiased estimator is one for which the expected value equals the parameter it is intended to estimate.

21 A is correct. A consistent estimator is one for which the probability of estimates close to the value of the population parameter increases as sample size increases. More specifically, a consistent estimator's sampling distribution becomes concentrated on the value of the parameter it is intended to estimate as the sample size approaches infinity.

22 A is correct. As the degree of confidence increases (e.g., from 95% to 99%), a given confidence interval will become wider. A confidence interval is a range for which one can assert with a given probability $1 - \alpha$, called the degree of confidence, that it will contain the parameter it is intended to estimate.

23 A is correct. A standard normal distribution has tails that approach zero faster than the t-distribution. As degrees of freedom increase, the tails of the t-distribution become less fat and the t-distribution begins to look more like a standard normal distribution. But as degrees of freedom decrease, the tails of the t-distribution become fatter.

24 B is correct. The confidence interval is calculated using the following equation:

$$\bar{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$$

Sample standard deviation (s) = $\sqrt{245.55}$ = 15.670.

For a sample size of 17, degrees of freedom equal 16, so $t_{0.05}$ = 1.746.

The confidence interval is calculated as

$$116.23 \pm 1.746\frac{15.67}{\sqrt{17}} = 116.23 \pm 6.6357$$

Therefore, the interval spans 109.5943 to 122.8656, meaning its width is equal to approximately 13.271. (This interval can be alternatively calculated as 6.6357 × 2).

25 A is correct. To solve, use the structure of Confidence interval = Point estimate ± Reliability factor × Standard error, which, for a normally distributed population with known variance, is represented by the following formula:

$$\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

For a 99% confidence interval, use $z_{0.005}$ = 2.58.

Also, $\sigma = \sqrt{529} = 23$.

Therefore, the lower limit $= 31 - 2.58\frac{23}{\sqrt{65}} = 23.6398$.

26 B is correct. All else being equal, as the sample size increases, the standard error of the sample mean decreases and the width of the confidence interval also decreases.

27 B is correct. A report that uses a current list of stocks does not account for firms that failed, merged, or otherwise disappeared from the public equity market in previous years. As a consequence, the report is biased. This type of bias is known as survivorship bias.

28 B is correct. An out-of-sample test is used to investigate the presence of data-mining bias. Such a test uses a sample that does not overlap the time period of the sample on which a variable, strategy, or model was developed.

29 A is correct. A short time series is likely to give period-specific results that may not reflect a longer time period.

PRACTICE PROBLEMS

1 Which of the following statements about hypothesis testing is correct?

A The null hypothesis is the condition a researcher hopes to support.

B The alternative hypothesis is the proposition considered true without conclusive evidence to the contrary.

C The alternative hypothesis exhausts all potential parameter values not accounted for by the null hypothesis.

2 Identify the appropriate test statistic or statistics for conducting the following hypothesis tests. (Clearly identify the test statistic and, if applicable, the number of degrees of freedom. For example, "We conduct the test using an x-statistic with y degrees of freedom.")

A H_0: $\mu = 0$ versus H_a: $\mu \neq 0$, where μ is the mean of a normally distributed population with unknown variance. The test is based on a sample of 15 observations.

B H_0: $\mu = 0$ versus H_a: $\mu \neq 0$, where μ is the mean of a normally distributed population with unknown variance. The test is based on a sample of 40 observations.

C H_0: $\mu \leq 0$ versus H_a: $\mu > 0$, where μ is the mean of a normally distributed population with known variance σ^2. The sample size is 45.

D H_0: $\sigma^2 = 200$ versus H_a: $\sigma^2 \neq 200$, where σ^2 is the variance of a normally distributed population. The sample size is 50.

E $H_0: \sigma_1^2 = \sigma_2^2$ versus $H_a: \sigma_1^2 \neq \sigma_2^2$, where σ_1^2 is the variance of one normally distributed population and σ_2^2 is the variance of a second normally distributed population. The test is based on two independent random samples.

F H_0: (Population mean 1) – (Population mean 2) = 0 versus H_a: (Population mean 1) – (Population mean 2) ≠ 0, where the samples are drawn from normally distributed populations with unknown variances. The observations in the two samples are correlated.

G H_0: (Population mean 1) – (Population mean 2) = 0 versus H_a: (Population mean 1) – (Population mean 2) ≠ 0, where the samples are drawn from normally distributed populations with unknown but assumed equal variances. The observations in the two samples (of size 25 and 30, respectively) are independent.

3 For each of the following hypothesis tests concerning the population mean, μ, state the rejection point condition or conditions for the test statistic (e.g., $t > 1.25$); n denotes sample size.

A H_0: $\mu = 10$ versus H_a: $\mu \neq 10$, using a t-test with $n = 26$ and $\alpha = 0.05$

B H_0: $\mu = 10$ versus H_a: $\mu \neq 10$, using a t-test with $n = 40$ and $\alpha = 0.01$

C H_0: $\mu \leq 10$ versus H_a: $\mu > 10$, using a t-test with $n = 40$ and $\alpha = 0.01$

D H_0: $\mu \leq 10$ versus H_a: $\mu > 10$, using a t-test with $n = 21$ and $\alpha = 0.05$

E H_0: $\mu \geq 10$ versus H_a: $\mu < 10$, using a t-test with $n = 19$ and $\alpha = 0.10$

F H_0: $\mu \geq 10$ versus H_a: $\mu < 10$, using a t-test with $n = 50$ and $\alpha = 0.05$

4 For each of the following hypothesis tests concerning the population mean, μ, state the rejection point condition or conditions for the test statistic (e.g., $z > 1.25$); n denotes sample size.

- **A** H_0: μ = 10 versus H_a: μ ≠ 10, using a z-test with n = 50 and α = 0.01
- **B** H_0: μ = 10 versus H_a: μ ≠ 10, using a z-test with n = 50 and α = 0.05
- **C** H_0: μ = 10 versus H_a: μ ≠ 10, using a z-test with n = 50 and α = 0.10
- **D** H_0: μ ≤ 10 versus H_a: μ > 10, using a z-test with n = 50 and α = 0.05

5 Willco is a manufacturer in a mature cyclical industry. During the most recent industry cycle, its net income averaged $30 million per year with a standard deviation of $10 million ($n$ = 6 observations). Management claims that Willco's performance during the most recent cycle results from new approaches and that we can dismiss profitability expectations based on its average or normalized earnings of $24 million per year in prior cycles.

- **A** With μ as the population value of mean annual net income, formulate null and alternative hypotheses consistent with testing Willco management's claim.
- **B** Assuming that Willco's net income is at least approximately normally distributed, identify the appropriate test statistic.
- **C** Identify the rejection point or points at the 0.05 level of significance for the hypothesis tested in Part A.
- **D** Determine whether or not to reject the null hypothesis at the 0.05 significance level.

The following information relates to Questions 6–7

Performance in Forecasting Quarterly Earnings per Share

	Number of Forecasts	Mean Forecast Error (Predicted – Actual)	Standard Deviations of Forecast Errors
Analyst A	101	0.05	0.10
Analyst B	121	0.02	0.09

6 Investment analysts often use earnings per share (EPS) forecasts. One test of forecasting quality is the zero-mean test, which states that optimal forecasts should have a mean forecasting error of 0. (Forecasting error = Predicted value of variable − Actual value of variable.)

You have collected data (shown in the table above) for two analysts who cover two different industries: Analyst A covers the telecom industry; Analyst B covers automotive parts and suppliers.

- **A** With μ as the population mean forecasting error, formulate null and alternative hypotheses for a zero-mean test of forecasting quality.
- **B** For Analyst A, using both a t-test and a z-test, determine whether to reject the null at the 0.05 and 0.01 levels of significance.
- **C** For Analyst B, using both a t-test and a z-test, determine whether to reject the null at the 0.05 and 0.01 levels of significance.

7 Reviewing the EPS forecasting performance data for Analysts A and B, you want to investigate whether the larger average forecast errors of Analyst A are due to chance or to a higher underlying mean value for Analyst A. Assume that the forecast errors of both analysts are normally distributed and that the samples are independent.

A Formulate null and alternative hypotheses consistent with determining whether the population mean value of Analyst A's forecast errors (μ_1) are larger than Analyst B's (μ_2).

B Identify the test statistic for conducting a test of the null hypothesis formulated in Part A.

C Identify the rejection point or points for the hypothesis tested in Part A, at the 0.05 level of significance.

D Determine whether or not to reject the null hypothesis at the 0.05 level of significance.

---

8 The table below gives data on the monthly returns on the S&P 500 and small-cap stocks for the period January 1960 through December 1999 and provides statistics relating to their mean differences.

Measure	S&P 500 Return (%)	Small-Cap Stock Return (%)	Differences (S&P 500–Small-Cap Stock)
January 1960–December 1999, 480 months			
Mean	1.0542	1.3117	–0.258
Standard deviation	4.2185	5.9570	3.752
January 1960–December 1979, 240 months			
Mean	0.6345	1.2741	–0.640
Standard deviation	4.0807	6.5829	4.096
January 1980–December 1999, 240 months			
Mean	1.4739	1.3492	0.125
Standard deviation	4.3197	5.2709	3.339

Let μ_d stand for the population mean value of difference between S&P 500 returns and small-cap stock returns. Use a significance level of 0.05 and suppose that mean differences are approximately normally distributed.

A Formulate null and alternative hypotheses consistent with testing whether any difference exists between the mean returns on the S&P 500 and small-cap stocks.

B Determine whether or not to reject the null hypothesis at the 0.05 significance level for the January 1960 to December 1999 period.

C Determine whether or not to reject the null hypothesis at the 0.05 significance level for the January 1960 to December 1979 subperiod.

D Determine whether or not to reject the null hypothesis at the 0.05 significance level for the January 1980 to December 1999 subperiod.

9 During a 10-year period, the standard deviation of annual returns on a portfolio you are analyzing was 15 percent a year. You want to see whether this record is sufficient evidence to support the conclusion that the portfolio's underlying variance of return was less than 400, the return variance of the portfolio's benchmark.

A Formulate null and alternative hypotheses consistent with the verbal description of your objective.

B Identify the test statistic for conducting a test of the hypotheses in Part A.

C Identify the rejection point or points at the 0.05 significance level for the hypothesis tested in Part A.

D Determine whether the null hypothesis is rejected or not rejected at the 0.05 level of significance.

10 You are investigating whether the population variance of returns on the S&P 500/BARRA Growth Index changed subsequent to the October 1987 market crash. You gather the following data for 120 months of returns before October 1987 and for 120 months of returns after October 1987. You have specified a 0.05 level of significance.

Time Period	*n*	Mean Monthly Return (%)	Variance of Returns
Before October 1987	120	1.416	22.367
After October 1987	120	1.436	15.795

A Formulate null and alternative hypotheses consistent with the verbal description of the research goal.

B Identify the test statistic for conducting a test of the hypotheses in Part A.

C Determine whether or not to reject the null hypothesis at the 0.05 level of significance. (Use the F-tables in the back of this volume.)

11 In the step "stating a decision rule" in testing a hypothesis, which of the following elements must be specified?

A Critical value

B Power of a test

C Value of a test statistic

12 Which of the following statements is correct with respect to the null hypothesis?

A It is considered to be true unless the sample provides evidence showing it is false.

B It can be stated as "not equal to" provided the alternative hypothesis is stated as "equal to."

C In a two-tailed test, it is rejected when evidence supports equality between the hypothesized value and population parameter.

13 An analyst is examining a large sample with an unknown population variance. To test the hypothesis that the historical average return on an index is less than or equal to 6%, which of the following is the *most* appropriate test?

A One-tailed z-test

B Two-tailed z-test

C One-tailed F-test

14 A hypothesis test for a normally-distributed population at a 0.05 significance level implies a:

A 95% probability of rejecting a true null hypothesis.

B 95% probability of a Type I error for a two-tailed test.

C 5% critical value rejection region in a tail of the distribution for a one-tailed test.

15 Which of the following statements regarding a one-tailed hypothesis test is correct?

A The rejection region increases in size as the level of significance becomes smaller.

B A one-tailed test more strongly reflects the beliefs of the researcher than a two-tailed test.

C The absolute value of the rejection point is larger than that of a two-tailed test at the same level of significance.

16 The value of a test statistic is *best* described as the basis for deciding whether to:

A reject the null hypothesis.

B accept the null hypothesis.

C reject the alternative hypothesis.

17 Which of the following is a Type I error?

A Rejecting a true null hypothesis

B Rejecting a false null hypothesis

C Failing to reject a false null hypothesis

18 A Type II error is *best* described as:

A rejecting a true null hypothesis.

B failing to reject a false null hypothesis.

C failing to reject a false alternative hypothesis.

19 The level of significance of a hypothesis test is *best* used to:

A calculate the test statistic.

B define the test's rejection points.

C specify the probability of a Type II error.

20 You are interested in whether excess risk-adjusted return (alpha) is correlated with mutual fund expense ratios for US large-cap growth funds. The following table presents the sample.

Mutual Fund	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Alpha (X)	−0.52	−0.13	−0.60	−1.01	−0.26	−0.89	−0.42	−0.23	−0.60
Expense Ratio (Y)	1.34	0.92	1.02	1.45	1.35	0.50	1.00	1.50	1.45

A Formulate null and alternative hypotheses consistent with the verbal description of the research goal.

B Identify the test statistic for conducting a test of the hypotheses in Part A.

C Justify your selection in Part B.

D Determine whether or not to reject the null hypothesis at the 0.05 level of significance.

21 All else equal, is specifying a smaller significance level in a hypothesis test likely to increase the probability of a:

	Type I error?	Type II error?
A	No	No
B	No	Yes
C	Yes	No

22 The probability of correctly rejecting the null hypothesis is the:

A p-value.

B power of a test.

C level of significance.

23 The power of a hypothesis test is:

A equivalent to the level of significance.

B the probability of not making a Type II error.

C unchanged by increasing a small sample size.

24 When making a decision in investments involving a statistically significant result, the:

A economic result should be presumed meaningful.

B statistical result should take priority over economic considerations.

C economic logic for the future relevance of the result should be further explored.

25 An analyst tests the profitability of a trading strategy with the null hypothesis being that the average abnormal return before trading costs equals zero. The calculated t-statistic is 2.802, with critical values of ± 2.756 at significance level $\alpha = 0.01$. After considering trading costs, the strategy's return is near zero. The results are *most likely*:

A statistically but not economically significant.

B economically but not statistically significant.

C neither statistically nor economically significant.

26 Which of the following statements is correct with respect to the p-value?

A It is a less precise measure of test evidence than rejection points.

B It is the largest level of significance at which the null hypothesis is rejected.

C It can be compared directly with the level of significance in reaching test conclusions.

27 Which of the following represents a correct statement about the p-value?

A The p-value offers less precise information than does the rejection points approach.

B A larger p-value provides stronger evidence in support of the alternative hypothesis.

C A p-value less than the specified level of significance leads to rejection of the null hypothesis.

28 Which of the following statements on p-value is correct?

A The p-value is the smallest level of significance at which H_0 can be rejected.

B The p-value indicates the probability of making a Type II error.

C The lower the p-value, the weaker the evidence for rejecting the H_0.

29 The following table shows the significance level (α) and the p-value for three hypothesis tests.

	α	*p*-value
Test 1	0.05	0.10
Test 2	0.10	0.08
Test 3	0.10	0.05

The evidence for rejecting H_0 is strongest for:

A Test 1.

B Test 2.

C Test 3.

30 Which of the following tests of a hypothesis concerning the population mean is *most* appropriate?

A A z-test if the population variance is unknown and the sample is small

B A z-test if the population is normally distributed with a known variance

C A t-test if the population is non-normally distributed with unknown variance and a small sample

31 For a small sample with unknown variance, which of the following tests of a hypothesis concerning the population mean is most appropriate?

A A t-test if the population is normally distributed

B A t-test if the population is non-normally distributed

C A z-test regardless of the normality of the population distribution

32 For a small sample from a normally distributed population with unknown variance, the *most* appropriate test statistic for the mean is the:

A z-statistic.

B t-statistic.

C χ^2 statistic.

33 An investment consultant conducts two independent random samples of 5-year performance data for US and European absolute return hedge funds. Noting a 50 basis point return advantage for US managers, the consultant decides to test whether the two means are statistically different from one another at a 0.05 level of significance. The two populations are assumed to be normally distributed with unknown but equal variances. Results of the hypothesis test are contained in the tables below.

	Sample Size	Mean Return %	Standard Deviation
US Managers	50	4.7	5.4
European Managers	50	4.2	4.8

Null and Alternative Hypotheses	H_0: $\mu_{US} - \mu_E = 0$; H_a: $\mu_{US} - \mu_E \neq 0$
Test Statistic	0.4893
Critical Value Rejection Points	±1.984

μ_{US} is the mean return for US funds and μ_E is the mean return for European funds.

The results of the hypothesis test indicate that the:

A null hypothesis is not rejected.

B alternative hypothesis is statistically confirmed.

C difference in mean returns is statistically different from zero.

34 A pooled estimator is used when testing a hypothesis concerning the:

A equality of the variances of two normally distributed populations.

B difference between the means of two at least approximately normally distributed populations with unknown but assumed equal variances.

C difference between the means of two at least approximately normally distributed populations with unknown and assumed unequal variances.

35 When evaluating mean differences between two dependent samples, the *most* appropriate test is a:

A chi-square test.

B paired comparisons test.

C z-test.

36 A fund manager reported a 2% mean quarterly return over the past ten years for its entire base of 250 client accounts that all follow the same investment strategy. A consultant employing the manager for 45 client accounts notes that their mean quarterly returns were 0.25% less over the same period. The consultant tests the hypothesis that the return disparity between the returns of his clients and the reported returns of the fund manager's 250 client accounts are significantly different from zero.

Assuming normally distributed populations with unknown population variances, the *most* appropriate test statistic is:

A a paired comparisons t-test.

B a t-test of the difference between the two population means.

C an approximate t-test of mean differences between the two populations.

37 A chi-square test is *most* appropriate for tests concerning:

A a single variance.

B differences between two population means with variances assumed to be equal.

C differences between two population means with variances assumed to not be equal.

38 Which of the following should be used to test the difference between the variances of two normally distributed populations?

A t-test

B F-test

C Paired comparisons test

39 In which of the following situations would a non-parametric test of a hypothesis *most likely* be used?

A The sample data are ranked according to magnitude.

B The sample data come from a normally distributed population.

C The test validity depends on many assumptions about the nature of the population.

40 An analyst is examining the monthly returns for two funds over one year. Both funds' returns are non-normally distributed. To test whether the mean return of one fund is greater than the mean return of the other fund, the analyst can use:

A a parametric test only.

B a nonparametric test only.

C both parametric and nonparametric tests.

SOLUTIONS

1 C is correct. Together, the null and alternative hypotheses account for all possible values of the parameter. Any possible values of the parameter not covered by the null must be covered by the alternative hypothesis (e.g., H_0: $\theta \leq 5$ versus H_a: $\theta > 5$).

2 **A** The appropriate test statistic is a t-statistic with $n - 1 = 15 - 1 = 14$ degrees of freedom. A t-statistic is theoretically correct when the sample comes from a normally distributed population with unknown variance. When the sample size is also small, there is no practical alternative.

B The appropriate test statistic is a t-statistic with $40 - 1 = 39$ degrees of freedom. A t-statistic is theoretically correct when the sample comes from a normally distributed population with unknown variance. When the sample size is large (generally, 30 or more is a "large" sample), it is also possible to use a z-statistic, whether the population is normally distributed or not. A test based on a t-statistic is more conservative than a z-statistic test.

C The appropriate test statistic is a z-statistic because the sample comes from a normally distributed population with known variance. (The known population standard deviation is used to compute the standard error of the mean using Equation 2 in the text.)

D The appropriate test statistic is chi-square (χ^2) with $50 - 1 = 49$ degrees of freedom.

E The appropriate test statistic is the F-statistic (the ratio of the sample variances).

F The appropriate test statistic is a t-statistic for a paired observations test (a paired comparisons test), because the samples are correlated.

G The appropriate test statistic is a t –statistic using a pooled estimate of the population variance. The t-statistic has $25 + 30 - 2 = 53$ degrees of freedom. This statistic is appropriate because the populations are normally distributed with unknown variances; because the variances are assumed equal, the observations can be pooled to estimate the common variance. The requirement of independent samples for using this statistic has been met.

3 **A** With degrees of freedom (df) $n - 1 = 26 - 1 = 25$, the rejection point conditions for this two-sided test are $t > 2.060$ and $t < -2.060$. Because the significance level is 0.05, $0.05/2 = 0.025$ of the probability is in each tail. The tables give one-sided (one-tailed) probabilities, so we used the 0.025 column. Read across df = 25 to the $\alpha = 0.025$ column to find 2.060, the rejection point for the right tail. By symmetry, −2.060 is the rejection point for the left tail.

B With df = 39, the rejection point conditions for this two-sided test are $t > 2.708$ and $t < -2.708$. This is a two-sided test, so we use the $0.01/2 = 0.005$ column. Read across df = 39 to the $\alpha = 0.005$ column to find 2.708, the rejection point for the right tail. By symmetry, −2.708 is the rejection point for the left tail.

C With df = 39, the rejection point condition for this one-sided test is $t > 2.426$. Read across df = 39 to the $\alpha = 0.01$ column to find 2.426, the rejection point for the right tail. Because we have a "greater than" alternative, we are concerned with only the right tail.

D With df = 20, the rejection point condition for this one-sided test is $t > 1.725$. Read across df = 20 to the $\alpha = 0.05$ column to find 1.725, the rejection point for the right tail. Because we have a "greater than" alternative, we are concerned with only the right tail.

E With df = 18, the rejection point condition for this one-sided test is $t < -1.330$. Read across df = 18 to the $\alpha = 0.10$ column to find 1.330, the rejection point for the right tail. By symmetry, the rejection point for the left tail is −1.330.

F With df = 49, the rejection point condition for this one-sided test is $t < -1.677$. Read across df = 49 to the $\alpha = 0.05$ column to find 1.677, the rejection point for the right tail. By symmetry, the rejection point for the left tail is −1.677.

4 Recall that with a z-test (in contrast to the t-test), we do not employ degrees of freedom. The standard normal distribution is a single distribution applicable to all z-tests. You should refer to "Rejection Points for a z-Test" in Section 3.1 to answer these questions.

A This is a two-sided test at a 0.01 significance level. In Part C of "Rejection Points for a z-Test," we find that the rejection point conditions are $z > 2.575$ and $z < -2.575$.

B This is a two-sided test at a 0.05 significance level. In Part B of "Rejection Points for a z-Test," we find that the rejection point conditions are $z > 1.96$ and $z < -1.96$.

C This is a two-sided test at a 0.10 significance level. In Part A of "Rejection Points for a z-Test," we find that the rejection point conditions are $z > 1.645$ and $z < -1.645$.

D This is a one-sided test at a 0.05 significance level. In Part B of "Rejection Points for a z-Test," we find that the rejection point condition for a test with a "greater than" alternative hypothesis is $z > 1.645$.

5 A As stated in the text, we often set up the "hoped for" or "suspected" condition as the alternative hypothesis. Here, that condition is that the population value of Willco's mean annual net income exceeds \$24 million. Thus we have H_0: $\mu \le 24$ versus H_a: $\mu > 24$.

B Given that net income is normally distributed with unknown variance, the appropriate test statistic is t with $n - 1 = 6 - 1 = 5$ degrees of freedom.

C In the t-distribution table in the back of the book, in the row for df = 5 under $\alpha = 0.05$, we read the rejection point (critical value) of 2.015. We will reject the null if $t > 2.015$.

D The t-test is given by Equation 4:

$$t_5 = \frac{\overline{X} - \mu_0}{s/\sqrt{n}} = \frac{30 - 24}{10/\sqrt{6}} = \frac{6}{4.082483} = 1.469694$$

or 1.47. Because 1.47 does not exceed 2.015, we do not reject the null hypothesis. The difference between the sample mean of \$30 million and the hypothesized value of \$24 million under the null is not statistically significant.

6 A H_0: $\mu = 0$ versus H_a: $\mu \ne 0$.

B The t-test is based on $t = \dfrac{\overline{X} - \mu_0}{s/\sqrt{n}}$ with $n - 1 = 101 - 1 = 100$ degrees of freedom. At the 0.05 significance level, we reject the null if $t > 1.984$ or if $t < -1.984$. At the 0.01 significance level, we reject the null if $t > 2.626$ or if $t <$

−2.626. For Analyst A, we have $t = (0.05 - 0)\big/\left(0.10/\sqrt{101}\right)$ = 0.05/0.00995 = 5.024938 or 5.025. We clearly reject the null hypothesis at both the 0.05 and 0.01 levels.

The calculation of the z-statistic with unknown variance, as in this case, is the same as the calculation of the t-statistic. The rejection point conditions for a two-tailed test are as follows: $z > 1.96$ and $z < -1.96$ at the 0.05 level; and $z > 2.575$ and $z < -2.575$ at the 0.01 level. Note that the z-test is a less conservative test than the t-test, so when the z-test is used, the null is easier to reject. Because $z = 5.025$ is greater than 2.575, we reject the null at the 0.01 level; we also reject the null at the 0.05 level.

In summary, Analyst A's EPS forecasts appear to be biased upward—they tend to be too high.

C For Analyst B, the t-test is based on t with 121 − 1 = 120 degrees of freedom. At the 0.05 significance level, we reject the null if $t > 1.980$ or if $t < -1.980$. At the 0.01 significance level, we reject the null if $t > 2.617$ or if $t < -2.617$. We calculate $t = (0.02 - 0)\big/\left(0.09/\sqrt{121}\right)$ = 0.02/0.008182 = 2.444444 or 2.44. Because 2.44 > 1.98, we reject the null at the 0.05 level. However, 2.44 is not larger than 2.617, so we do not reject the null at the 0.01 level.

For a z-test, the rejection point conditions are the same as given in Part B, and we come to the same conclusions as with the t-test. Because 2.44 > 1.96, we reject the null at the 0.05 significance level; however, because 2.44 is not greater than 2.575, we do not reject the null at the 0.01 level.

The mean forecast error of Analyst B is only $0.02; but because the test is based on a large number of observations, it is sufficient evidence to reject the null of mean zero forecast errors at the 0.05 level.

7 A Stating the suspected condition as the alternative hypothesis, we have

$$H_0\text{: } \mu_1 - \mu_2 \le 0 \text{ versus } H_a\text{: } \mu_1 - \mu_2 > 0$$

where

μ_1 = the population mean value of Analyst A's forecast errors
μ_2 = the population mean value of Analyst B's forecast errors

B We have two normally distributed populations with unknown variances. Based on the samples, it is reasonable to assume that the population variances are equal. The samples are assumed to be independent; this assumption is reasonable because the analysts cover quite different industries. The appropriate test statistic is t using a pooled estimate of the common variance. The number of degrees of freedom is

$$n_1 + n_2 - 2 = 101 + 121 - 2 = 222 - 2 = 220.$$

C For df = 200 (the closest value to 220), the rejection point for a one-sided test at the 0.05 significance level is 1.653.

D We first calculate the pooled estimate of variance:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(101 - 1)(0.10)^2 + (121 - 1)(0.09)^2}{101 + 121 - 2}$$

$$= \frac{1.972}{220} = 0.008964$$

Then

$$t = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\left(\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}\right)^{1/2}} = \frac{\left(0.05 - 0.02\right) - 0}{\left(\dfrac{0.008964}{101} + \dfrac{0.008964}{121}\right)^{1/2}}$$

$$= \frac{0.03}{0.01276} = 2.351018$$

or 2.35. Because 2.35 > 1.653, we reject the null hypothesis in favor of the alternative hypothesis that the population mean forecast error of Analyst A is greater than that of Analyst B.

8 **A** We test H_0: $\mu_d = 0$ versus H_a: $\mu_d \neq 0$.

B This is a paired comparisons t-test with $n - 1 = 480 - 1 = 479$ degrees of freedom. At the 0.05 significance level, we reject the null hypothesis if either $t > 1.96$ or $t < -1.96$. We use df = ∞ in the t-distribution table under α = 0.025 because we have a very large sample and a two-sided test.

$$t = \frac{\bar{d} - \mu_{d0}}{s_{\bar{d}}} = \frac{-0.258 - 0}{3.752/\sqrt{480}} = \frac{-0.258}{0.171255} = -1.506529 \text{ or } -1.51$$

At the 0.05 significance level, because neither rejection point condition is met, we do not reject the null hypothesis that the mean difference between the returns on the S&P 500 and small-cap stocks during the entire sample period was 0.

C This t-test now has $n - 1 = 240 - 1 = 239$ degrees of freedom. At the 0.05 significance level, we reject the null hypothesis if either $t > 1.972$ or $t < -1.972$, using df = 200 in the t-distribution tables.

$$t = \frac{\bar{d} - \mu_{d0}}{s_{\bar{d}}} = \frac{-0.640 - 0}{4.096/\sqrt{240}} = \frac{-0.640}{0.264396} = -2.420615 \text{ or } -2.42$$

Because −2.42 < −1.972, we reject the null hypothesis at the 0.05 significance level. During this subperiod, small-cap stocks significantly outperformed the S&P 500.

D This t-test has $n - 1 = 240 - 1 = 239$ degrees of freedom. At the 0.05 significance level, we reject the null hypothesis if either $t > 1.972$ or $t < -1.972$, using df = 200 in the t-distribution tables.

$$t = \frac{\bar{d} - \mu_{d0}}{s_{\bar{d}}} = \frac{0.125 - 0}{3.339/\sqrt{240}} = \frac{0.125}{0.215532} = 0.579962 \text{ or } 0.58$$

At the 0.05 significance level, because neither rejection point condition is met, we do not reject the null hypothesis that for the January 1980–December 1999 period, the mean difference between the returns on the S&P 500 and small-cap stocks was zero.

9 **A** We have a "less than" alternative hypothesis, where σ^2 is the variance of return on the portfolio. The hypotheses are H_0: $\sigma^2 \geq 400$ versus H_a: $\sigma^2 < 400$, where 400 is the hypothesized value of variance, σ_0^2.

B The test statistic is chi-square with $10 - 1 = 9$ degrees of freedom.

C The rejection point is found across degrees of freedom of 9, under the 0.95 column (95 percent of probability above the value). It is 3.325. We will reject the null hypothesis if we find that $\chi^2 < 3.325$.

D The test statistic is calculated as

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{9 \times 15^2}{400} = \frac{2{,}025}{400} = 5.0625 \text{ or } 5.06$$

Because 5.06 is not less than 3.325, we do not reject the null hypothesis.

10 **A** We have a "not equal to" alternative hypothesis:

$$H_0\colon \sigma^2_{\text{Before}} = \sigma^2_{\text{After}} \text{ versus } H_a\colon \sigma^2_{\text{Before}} \neq \sigma^2_{\text{After}}$$

B To test a null hypothesis of the equality of two variances, we use an F-test:

$$F = \frac{s_1^2}{s_2^2}$$

C The "before" sample variance is larger, so following a convention for calculating F-statistics, the "before" sample variance goes in the numerator. $F = 22.367/15.795 = 1.416$, with 120 − 1 = 119 numerator and denominator degrees of freedom. Because this is a two-tailed test, we use F-tables for the 0.025 level (df = 0.05/2). Using the tables in the back of the volume, the closest value to 119 is 120 degrees of freedom. At the 0.05 level, the rejection point is 1.43. (Using the Insert/Function/Statistical feature on a Microsoft Excel spreadsheet, we would find FINV(0.025, 119, 119) = 1.434859 as the critical F-value.) Because 1.416 is not greater than 1.43, we do not reject the null hypothesis that the "before" and "after" variances are equal.

11 A is correct. The critical value in a decision rule is the rejection point for the test. It is the point with which the test statistic is compared to determine whether to reject the null hypothesis, which is part of the fourth step in hypothesis testing.

12 A is correct. The null hypothesis is the hypothesis to be tested. The null hypothesis is considered to be true unless the evidence indicates that it is false, in which case the alternative hypothesis is accepted.

13 A is correct. If the population sampled has unknown variance and the sample is large, a z-test may be used. Hypotheses involving "greater than" or "less than" postulations are one-sided (one-tailed). In this situation, the null and alternative hypotheses are stated as H_0: $\mu \leq 6\%$ and H_a: $\mu > 6\%$, respectively. A one-tailed t-test is also acceptable in this case.

14 C is correct. For a one-tailed hypothesis test, there is a 5% critical value rejection region in one tail of the distribution.

15 B is correct. One-tailed tests in which the alternative is "greater than" or "less than" represent the beliefs of the researcher more firmly than a "not equal to" alternative hypothesis.

16 A is correct. Calculated using a sample, a test statistic is a quantity whose value is the basis for deciding whether to reject the null hypothesis.

17 A is correct. The definition of a Type I error is when a true null hypothesis is rejected.

18 B is correct. A Type II error occurs when a false null hypothesis is not rejected.

19 B is correct. The level of significance is used to establish the rejection points of the hypothesis test.

20 **A** We have a "not equal to" alternative hypothesis:

$$H_0\colon \rho = 0 \text{ versus } H_a\colon \rho \neq 0$$

B We would use the nonparametric Spearman rank correlation coefficient to conduct the test.

C Mutual fund expense ratios are bounded from above and below, and in practice there is at least a lower bound on alpha (as any return cannot be less than −100 percent). These variables are markedly non-normally distributed, and the assumptions of a parametric test are not likely to be fulfilled. Thus a nonparametric test appears to be appropriate.

D The calculation of the Spearman rank correlation coefficient is given in the following table.

Mutual Fund	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Alpha (X)	−0.52	−0.13	−0.60	−1.01	−0.26	−0.89	−0.42	−0.23	−0.60
Expense Ratio (Y)	1.34	0.92	1.02	1.45	1.35	0.50	1.00	1.50	1.45
X Rank	5	1	6.5	9	3	8	4	2	6.5
Y Rank	5	8	6	2.5	4	9	7	1	2.5
d_i	0	−7	0.5	6.5	−1	−1	−3	1	4
d_i^2	0	49	0.25	42.25	1	1	9	1	16

$$r_s = 1 - \frac{6\sum d_i^2}{n\left(n^2 - 1\right)} = 1 - \frac{6(119.50)}{9(81 - 1)} = 0.0042$$

We use Table 11 to tabulate the rejection points for a test on the Spearman rank correlation. Given a sample size of 9 in a two-tailed test at a 0.05 significance level, the upper-tail rejection point is 0.6833 (we use the 0.025 column). Thus we reject the null hypothesis if the Spearman rank correlation coefficient is less than −0.6833 or greater than 0.6833. Because r_S is equal to 0.0042, we do not reject the null hypothesis.

21 B is correct. Specifying a smaller significance level decreases the probability of a Type I error (rejecting a true null hypothesis), but increases the probability of a Type II error (not rejecting a false null hypothesis). As the level of significance decreases, the null hypothesis is less frequently rejected.

22 B is correct. The power of a test is the probability of rejecting the null hypothesis when it is false.

23 B is correct. The power of a hypothesis test is the probability of correctly rejecting the null when it is false. Failing to reject the null when it is false is a Type II error. Thus, the power of a hypothesis test is the probability of not committing a Type II error.

24 C is correct. When a statistically significant result is also economically meaningful, one should further explore the logic of why the result might work in the future.

25 A is correct. The hypothesis is a two-tailed formulation. The t-statistic of 2.802 falls outside the critical rejection points of less than −2.756 and greater than 2.756, therefore the null hypothesis is rejected; the result is statistically significant. However, despite the statistical results, trying to profit on the strategy is not likely to be economically meaningful because the return is near zero after transaction costs

26 C is correct. When directly comparing the *p*-value with the level of significance, it can be used as an alternative to using rejection points to reach conclusions on hypothesis tests. If the *p*-value is smaller than the specified level of significance, the null hypothesis is rejected. Otherwise, the null hypothesis is not rejected.

27 C is correct. The *p*-value is the smallest level of significance at which the null hypothesis can be rejected for a given value of the test statistic. The null hypothesis is rejected when the *p*-value is less than the specified significance level.

28 A is correct. The *p*-value is the smallest level of significance (α) at which the null hypothesis can be rejected.

29 C is correct. The *p*-value is the smallest level of significance (α) at which the null hypothesis can be rejected. If the *p*-value is less than α, the null can be rejected. The smaller the *p*-value, the stronger the evidence is against the null hypothesis and in favor of the alternative hypothesis. Thus, the evidence for rejecting the null is strongest for Test 3.

30 B is correct. The *z*-test is theoretically the correct test to use in those limited cases when testing the population mean of a normally distributed population with known variance.

31 A is correct. A *t*-test is used if the sample is small and drawn from a normally or approximately normally distributed population.

32 B is correct. A *t*-statistic is the most appropriate for hypothesis tests of the population mean when the variance is unknown and the sample is small but the population is normally distributed.

33 A is correct. The *t*-statistic value of 0.4893 does not fall into the critical value rejection regions (≤ -1.984 or > 1.984). Instead it falls well within the acceptance region. Thus, H_0 cannot be rejected; the result is not statistically significant at the 0.05 level.

34 B is correct. The assumption that the variances are equal allows for the combining of both samples to obtain a pooled estimate of the common variance.

35 B is correct. A paired comparisons test is appropriate to test the mean differences of two samples believed to be dependent.

36 A is correct. The sample sizes for both the fund manager and the consultant's accounts consists of forty quarterly periods of returns. However, the consultant's client accounts are a subset of the fund manager's entire account base. As such, they are not independent samples. When samples are dependent, a paired comparisons test is appropriate to conduct tests of the differences in dependent items.

37 A is correct. A chi-square test is used for tests concerning the variance of a single normally distributed population.

38 B is correct. An *F*-test is used to conduct tests concerning the difference between the variances of two normally distributed populations with random independent samples.

39 A is correct. A non-parametric test is used when the data are given in ranks.

40 B is correct. There are only 12 (monthly) observations over the one year of the sample and thus the samples are small. Additionally, the funds' returns are non-normally distributed. Therefore, the samples do not meet the distributional assumptions for a parametric test. The Mann–Whitney U test (a nonparametric test) could be used to test the differences between population means

PRACTICE PROBLEMS

1 Technical analysis relies most importantly on:
- A price and volume data.
- B accurate financial statements.
- C fundamental analysis to confirm conclusions.

2 Which of the following is *not* an assumption of technical analysis?
- A Security markets are efficient.
- B The security under analysis is freely traded.
- C Market trends and patterns tend to repeat themselves.

3 Drawbacks of technical analysis include which of the following?
- A It identifies changes in trends only after the fact.
- B Deviations from intrinsic value can persist for long periods.
- C It usually requires detailed knowledge of the financial instrument under analysis.

4 Why is technical analysis especially useful in the analysis of commodities and currencies?
- A Valuation models cannot be used to determine fundamental intrinsic value for these securities.
- B Government regulators are more likely to intervene in these markets.
- C These types of securities display clearer trends than equities and bonds do.

5 A daily bar chart provides:
- A a logarithmically scaled horizontal axis.
- B a horizontal axis that represents changes in price.
- C high and low prices during the day and the day's opening and closing prices.

6 A candlestick chart is similar to a bar chart *except* that the candlestick chart:
- A represents upward movements in price with X's.
- B also graphically shows the range of the period's highs and lows.
- C has a body that is light or dark depending on whether the security closed higher or lower than its open.

7 In analyzing a price chart, high or increasing volume *most likely* indicates which of the following?
- A Predicts a reversal in the price trend.
- B Predicts that a trendless period will follow.
- C Confirms a rising or declining trend in prices.

8 In constructing a chart, using a logarithmic scale on the vertical axis is likely to be *most useful* for which of the following applications?
- A The price of gold for the past 100 years.
- B The share price of a company over the past month.
- C Yields on 10-year US Treasuries for the past 5 years.

9 A downtrend line is constructed by drawing a line connecting:
- A the lows of the price chart.

B the highs of the price chart.

C the highest high to the lowest low of the price chart.

10 Exhibit 1 depicts GreatWall Information Industry Co., Ltd., ordinary shares, traded on the Shenzhen Stock Exchange, for late 2008 through late 2009 in renminbi (RMB).

Exhibit 1 Candlestick Chart: GreatWall Information Industry Co., Ltd. Price Data, November 2008–September 2009 (Price Measured in RMB × 10)

Based on Exhibit 1, the uptrend was *most likely* broken at a price level nearest to:

A 7 RMB.

B 8.5 RMB.

C 10 RMB.

11 The "change in polarity" principle states which of the following?

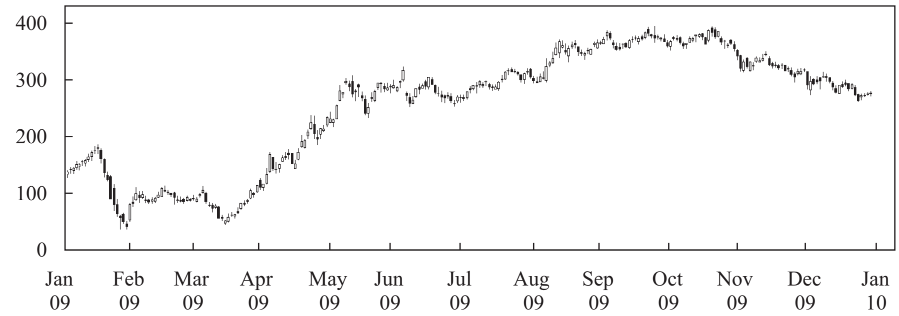
A Once an uptrend is broken, it becomes a downtrend.

B Once a resistance level is breached, it becomes a support level.

C The short-term moving average has crossed over the longer-term moving average.

12 Exhibit 2 depicts Barclays ordinary shares, traded on the London Stock Exchange, for 2009 in British pence.

Exhibit 2 Candlestick Chart: Barclays plc Price Data, January 2009–January 2010 (Price Measured in British Pence)



Based on Exhibit 2, Barclays appears to show resistance at a level nearest to:

A 50p.

B 275p.

C 390p.

13 Exhibit 3 depicts Archer Daniels Midland Company common shares, traded on the New York Stock Exchange, for 1996 to 2001 in US dollars.

Exhibit 3 Candlestick Chart: Archer Daniels Midland Company, February 1996–February 2001

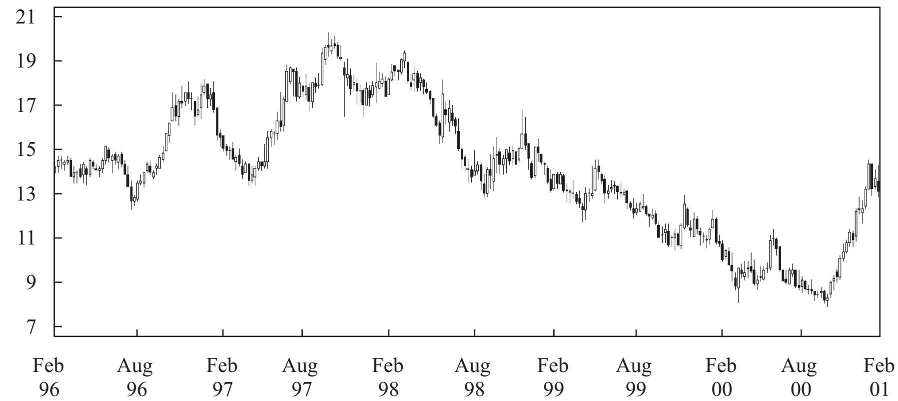


Exhibit 3 illustrates *most* clearly which type of pattern?

A Triangle.

B Triple top.

C Head and shoulders.

14 In an inverted head and shoulders pattern, if the neckline is at €100, the shoulders at €90, and the head at €75, the price target is *closest* to which of the following?

A €50.

B €110.

C €125.

15 Which flow-of-funds indicator is considered bearish for equities?

A A large increase in the number of IPOs.

B Higher-than-average cash balances in mutual funds.

C An upturn in margin debt but one that is still below the long-term average.

16 A TRIN with a value of less than 1.0 indicates:

A the market is in balance.

B there is more volume in rising shares.

C there is more volume in declining shares.

17 Bollinger Bands are constructed by plotting:

A a MACD line and a signal line.

B a moving-average line with an uptrend line above and downtrend line below.

C a moving-average line with upper and lower lines that are at a set number of standard deviations apart.

18 Which of the following is *not* a momentum oscillator?

A MACD.

B Stochastic oscillator.

C Bollinger Bands.

19 Which of the following is a continuation pattern?

A Triangle.

B Triple top.

C Head and shoulders.

20 Which of the following is a reversal pattern?

A Pennant.

B Rectangle.

C Double bottom.

21 Which of the following is generally true of the head and shoulders pattern?

A Volume is important in interpreting the data.

B The neckline, once breached, becomes a support level.

C Head and shoulders patterns are generally followed by an uptrend in the security's price.

22 Nikolai Kondratieff concluded in the 1920s that since the 1780s, Western economies have generally followed a cycle of how many years?

A 18.

B 54.

C 76.

23 Based on the decennial pattern of cycles, how would the return of the Dow Jones Industrial Average (DJIA) in the year 2015 compare with the return in 2020?

A The return would be better.

B The return would be worse.

C The answer cannot be determined because the theory does not apply to both of those years.

24 According to the US presidential cycle theory, the DJIA has the best performance during which year?

A The presidential election year itself.

B The first year following a presidential election.

C The third year following a presidential election.

25 What is a major problem with long-term cycle theories?

A The sample size is small.

B The data are usually hard to observe.

C They occur over such a long period that they are difficult to discern.

26 In 1938, R. N. Elliott proposed a theory that equity markets move:

A in stochastic waves.

B in cycles following Fibonacci ratios.

C in waves dependent on other securities.

27 All of the following are names of Elliott cycles *except*:

A presidential.

B supercycle.

C grand supercycle.

28 To identify intermarket relationships, technicians commonly use:

A stochastic oscillators.

B Fibonacci ratios.

C relative strength analysis.

SOLUTIONS

1 A is correct. Almost all technical analysis relies on these data inputs.

2 A is correct. Technical analysis works because markets are *not* efficient and rational and because human beings tend to behave similarly in similar circumstances. The result is market trends and patterns that repeat themselves and are somewhat predictable.

3 A is correct. Trends generally must be in place for some time before they are recognizable. Thus, some time may be needed for a change in trend to be identified.

4 A is correct. Commodities and currencies do not have underlying financial statements or an income stream; thus, fundamental analysis is useless in determining theoretical values for them or whether they are over- or undervalued.

5 C is correct. The top and bottom of the bars indicate the highs and lows for the day; the line on the left indicates the opening price and the line on the right indicates the closing price.

6 C is correct. Dark and light shading is a unique feature of candlestick charts.

7 C is correct. Rising volume shows conviction by many market participants, which is likely to lead to a continuation of the trend.

8 A is correct. The price of gold in nominal dollars was several orders of magnitude cheaper 100 years ago than it is today (roughly US$20 then versus US$1,100 today). Such a wide range of prices lends itself well to being graphically displayed on a logarithmic scale.

9 B is correct. A downtrend line is constructed by drawing a line connecting the highs of the price chart.

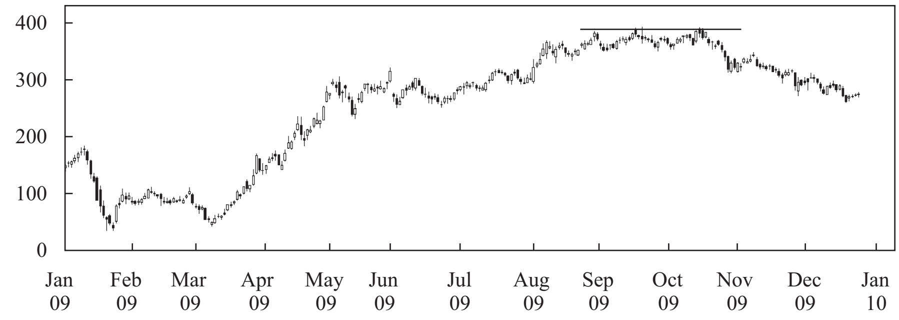
10 B is correct. It is demonstrated in the following chart:

Exhibit 1 Candlestick Chart: GreatWall Information Industry Co., Ltd. Price Data, November 2008–September 2009 (Price Measured in RMB × 10)

11 B is correct.

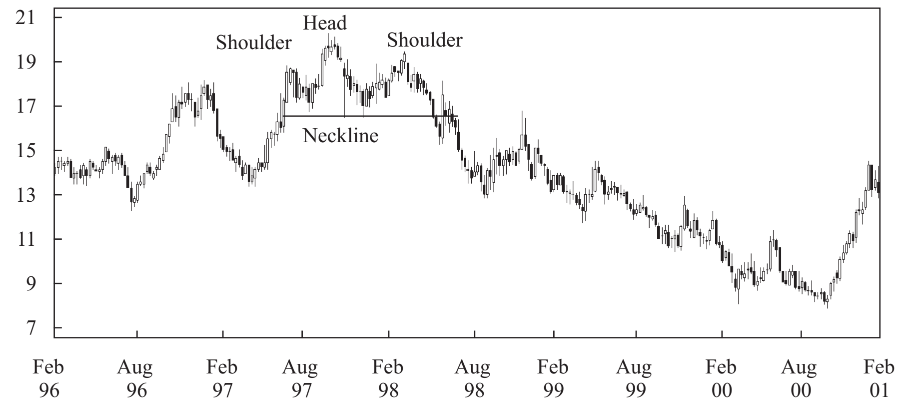
12 C is correct. As shown in the following chart, Barclays shares traded up to 390p on three occasions, each several weeks apart, and declined thereafter each time.

Exhibit 2 Candlestick Chart: Barclays plc Price Data, January 2009–January 2010 (Price Measured in British Pence)



13 C is correct. The left shoulder formed at around US$18.50, the head formed at around US$20.50, and the second shoulder formed at around US$19.

Exhibit 3 Candlestick Chart: Archer Daniels Midland Company, February 1996–February 2001



14 C is correct. Target = Neckline + (Neckline − Head): €100 + (€100 − €75) = €125

15 A is correct. A large increase in the number of IPOs increases the supply of equity and if overall demand remains the same, puts downward pressure on equities. Also, companies tend to issue shares of equity when the managers believe they will receive a premium price, which is also an indicator of a market top.

16 B is correct. A value below 1.0 is a bullish sign; it means more volume is in rising shares than in declining ones. The TRIN is calculated as: (Advancing issues/Declining issues)/(Volume of advancing issues/Volume of declining issues).

17 C is correct. Bollinger Bands consist of a moving average and a higher line representing the moving average plus a set number of standard deviations from average price (for the same number of periods as used to calculate the moving average) and a lower line that is a moving average minus the same number of standard deviations.

18 C is correct. Bollinger Bands are price-based indicators, *not* momentum oscillators, which are constructed so that they oscillate between a high and a low or around 0 or 100.

19 A is correct. Triangles are one of several continuation patterns.

20 C is correct. It is one of several reversal patterns.

21 A is correct. Volume is necessary to confirm the various market rallies and reversals during the formation of the head and shoulders pattern.

22 B is correct.

23 A is correct. The decennial pattern theory states that years ending with a 5 will have the best performance of any of the 10 years in a decade and that those ending with a zero will have the worst.

24 C is correct. A possible reason for the superior performance in the third year is that the US presidential election occurs, together with a number of other elections, in a four-year cycle, so the politicians desiring to be reelected inject money into the economy in the third year to improve their chances of winning the following year.

25 A is correct. Long-term cycles require many years to complete; thus, not many cycles are available to observe.

26 B is correct.

27 A is correct. This is the term for a separate cycle theory.

28 C is correct. Relative strength analysis is often used to compare two asset classes or two securities.