READING

# 8

# Probability Concepts

**by Richard A. DeFusco, PhD, CFA, Dennis W. McLeavey, DBA, 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, DBA, 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 and interpret a scatterplot; |
| ☐ | **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. It also discusses scatter plots, a graphical depiction of the relatedness between two random variables. 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: $0 \leq P(E) \leq 1$.
  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 August 2018, analysts' fiscal year 2019 EPS forecasts for JetBlue Airways ranged from \$1.50 to \$2.20. Suppose one analyst asserts that the odds for the company beating the highest estimate, \$2.20, 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 FY2019 beating \$2.20* 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 FY2019 beating $2.20* 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. 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).

$$P(A \mid B) = P(AB)/P(B),\ P(B) \neq 0 \qquad (1)$$

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

$$P(AB) = P(A \mid B)P(B) \qquad (2)$$

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 over a 13-year period. Funds were classified into nine investment styles based on combinations of investment focus (growth, blend, and value) and funds' 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 over a 13-Year Period**

| | 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. Let A represent the event that a *fund is a Year 2 loser*, and let B represent the event that *the fund is a Year 1 loser*. Therefore, the event 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.

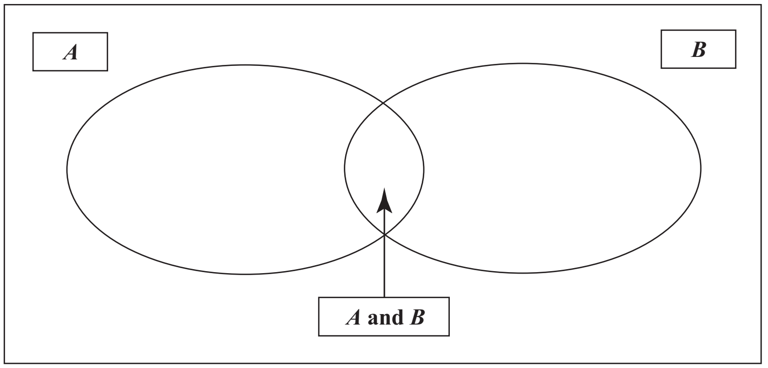
When we have two events, $A$ and $B$, that we are interested in, we often want to know the probability that either $A$ or $B$ occurs. Here the word "or" is inclusive, meaning that either $A$ or $B$ occurs or that both $A$ and $B$ occur. Put another way, the probability of $A$ or $B$ is the probability 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.

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

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:Year 1 (that is, EPS for the second quarter of Year 1) were larger than EPS for 1Q:Year 1.

1 What is the probability that 3Q:Year 1 EPS will be larger than 2Q:Year 1 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:Year 1 EPS will be larger than 2Q:Year 1 EPS is the unconditional probability of positive change, 0.55. The fact that 2Q:Year 1 EPS was larger than 1Q:Year 1 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

$$
\begin{aligned}
P(ABCDE) &= P(A)P(B)P(C)P(D)P(E) = (0.50)(0.50)(0.25)(0.55)(0.67) \\
&= 0.023031
\end{aligned}
$$

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:Year 1) is announced, and the change is a positive sequential change (the event $S$). You are interested in forecasting EPS for 3Q:Year 1.

**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:Year 1, given the positive change in sequential EPS for 2Q:Year 1, as shown below.

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.

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

The expected value of EPS is $2.34.

An equation that summarizes your calculation in Example 8 is

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

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.

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

Standard deviation is the positive square root of 0.038785:

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

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

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

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.

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

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.**

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

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

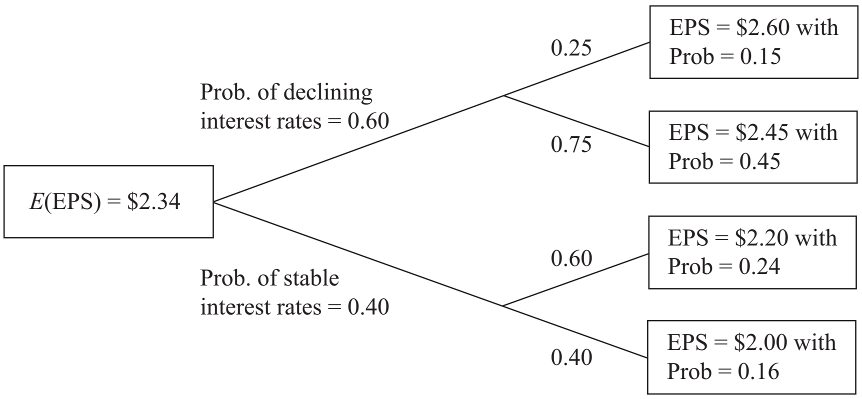
where $S_1, S_2, \ldots, S_n$ are mutually exclusive and exhaustive scenarios or events.

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

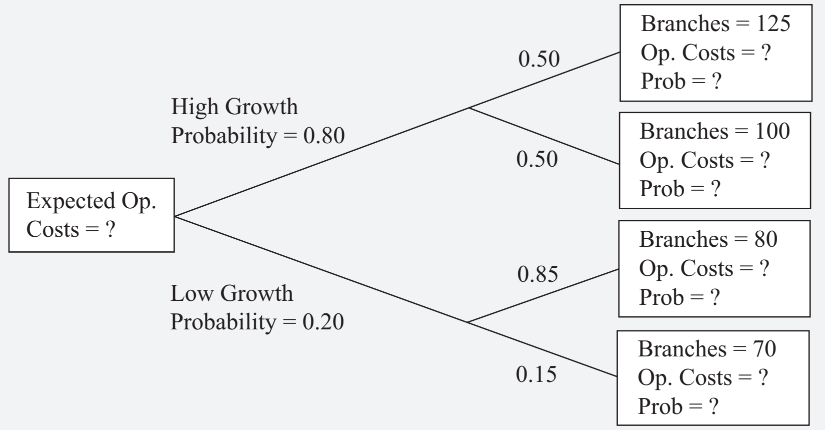
$$\hat{Y} = a + bX$$

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.

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

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 $P(\textit{the bond defaults}) = 0.06$. 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.

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

Solving for the promised return on the bond, you find $R = \{(1 + R_F)/[1 - P(\textit{the bond defaults})]\} - 1$. Substituting 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, $w_1, w_2, \ldots, w_n$, 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:

$$\begin{aligned} E\left(R_p\right) &= E\left(w_1 R_1 + w_2 R_2 + \ldots + w_n R_n\right) \\ &= w_1 E\left(R_1\right) + w_2 E\left(R_2\right) + \ldots + w_n E\left(R_n\right) \end{aligned}$$

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:

$$\begin{aligned} E\left(R_p\right) &= w_1 E\left(R_1\right) + w_2 E\left(R_2\right) + w_3 E\left(R_3\right) \\ &= 0.50(13\%) + 0.25(6\%) + 0.25(15\%) = 11.75\% \end{aligned}$$

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

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

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

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. The above measure is the population covariance and it is forward-looking. Sometimes analysts look at historical covariance for guidance on developing expectations for the future. For this purpose, the sample covariance, which is computed using a sample of historical data about the two variables, is appropriate. The sample covariance between two random variables $R_i$ and $R_j$, based on a sample of past data of size n is

$$\text{Cov}\left(R_i, R_j\right) = \sum_{i=1}^{n}\left(R_{i,t} - \bar{R}_i\right)\left(R_{j,t} - \bar{R}_j\right)\Big/(n-1) \quad (15)$$

The sample covariance is the average value of the product of the deviations of observations on two random variables from their sample means.[11] If the random variables are returns, the units of both forward-looking covariance and historical variance would be returns squared. In this reading, we will consider covariance in a forward-looking sense, unless mentioned otherwise.

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

11 The use of $n - 1$ in the denominator is a technical point; it ensures that the sample covariance is an unbiased estimate of population covariance.

$$= w_1^2\sigma^2(R_1) + w_1w_2\text{Cov}(R_1,R_2) + w_1w_3\text{Cov}(R_1,R_3)$$
$$+ w_1w_2\mathit{Cov}(R_1,R_2) + w_2^2\sigma^2(R_2) + w_2w_3\text{Cov}(R_2,R_3) \quad (16)$$
$$+ w_1w_3\mathit{Cov}(R_1,R_3) + w_2w_3\mathit{Cov}(R_2,R_3) + w_3^2\sigma^2(R_3)$$

The last step follows from the definitions of variance and covariance.[12] For the italicized covariance terms in Equation 16, 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 state Equation 16 is $\sigma^2(R_p) = \sum_{i=1}^{3}\sum_{j=1}^{3} w_iw_j\text{Cov}(R_i,R_j)$. 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

$$\sigma^2(R_p) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_iw_j\text{Cov}(R_i,R_j) \quad (17)$$

We see from Equation 16 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[13] 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: $\text{Cov}(R,R) = E\{[R - E(R)][R - E(R)]\} = E\{[R - E(R)]^2\} = \sigma^2(R)$.

---

**12** 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.
**13** 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.

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 |
|---|---|---|---|
| | $E(R_A)$ | $E(R_B)$ | $E(R_C)$ |

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

| Asset | A | B | C |
|---|---|---|---|
| A | $\mathbf{Cov(R_A,R_A)}$ | $\mathrm{Cov}(R_A,R_B)$ | $\mathrm{Cov}(R_A,R_C)$ |
| B | $\mathrm{Cov}(R_B,R_A)$ | $\mathbf{Cov(R_B,R_B)}$ | $\mathrm{Cov}(R_B,R_C)$ |
| C | $\mathrm{Cov}(R_C,R_A)$ | $\mathrm{Cov}(R_C,R_B)$ | $\mathbf{Cov(R_C,R_C)}$ |

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 $\mathrm{Cov}(R_B,R_A) = \mathrm{Cov}(R_A,R_B)$, $\mathrm{Cov}(R_C,R_A) = \mathrm{Cov}(R_A,R_C)$, and $\mathrm{Cov}(R_C,R_B) = \mathrm{Cov}(R_B,R_C)$. 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 16 and grouping variance terms together produces the following:

$$\begin{aligned}\sigma^2\left(R_p\right) &= w_1^2\sigma^2\left(R_1\right) + w_2^2\sigma^2\left(R_2\right) + w_3^2\sigma^2\left(R_3\right) + 2w_1w_2\text{Cov}\left(R_1,R_2\right)\\ &\quad + 2w_1w_3\text{Cov}\left(R_1,R_3\right) + 2w_2w_3\text{Cov}\left(R_2,R_3\right)\\ &= \left(0.50\right)^2\left(400\right) + \left(0.25\right)^2\left(81\right) + \left(0.25\right)^2\left(441\right)\\ &\quad + 2\left(0.50\right)\left(0.25\right)\left(45\right) + 2\left(0.50\right)\left(0.25\right)\left(189\right)\\ &\quad + 2\left(0.25\right)\left(0.25\right)\left(38\right)\\ &= 100 + 5.0625 + 27.5625 + 11.25 + 47.25 + 4.75 = 195.875\end{aligned} \tag{18}$$

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 $\rho_{ij}$.

The above definition of correlation is forward-looking because it involves dividing the forward-looking covariance by the product of forward-looking standard deviations. We can similarly compute an historical or sample correlation by dividing historical or sample covariance between two variables by the product of sample standard deviations of the two variables.[14]

Frequently, covariance is substituted out using the relationship $\text{Cov}(R_i,R_j) = \rho(R_i,R_j)\sigma(R_i)\sigma(R_j)$. Like covariance, the correlation coefficient is a measure of linear association. However, the division indicated in the definition of correlation 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.

14 Sample covariance is discussed earlier in this reading. Sample standard deviation is discussed in the reading on statistical concepts and market returns.

**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 may range from −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.[15] 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).[16]

We will make use of scatter plots to illustrate correlation. A **scatter plot** is a graph that shows the relationship between the observations for two data series in two dimensions. In contrast to correlation analysis, which expresses the relationship between two data series using a single number, a scatter plot depicts this same relationship graphically. Each observation in a scatter plot is represented as a point, and the points are not connected. A scatter plot shows only the actual observations of both data series plotted as pairs and does not include specifics about the observations.

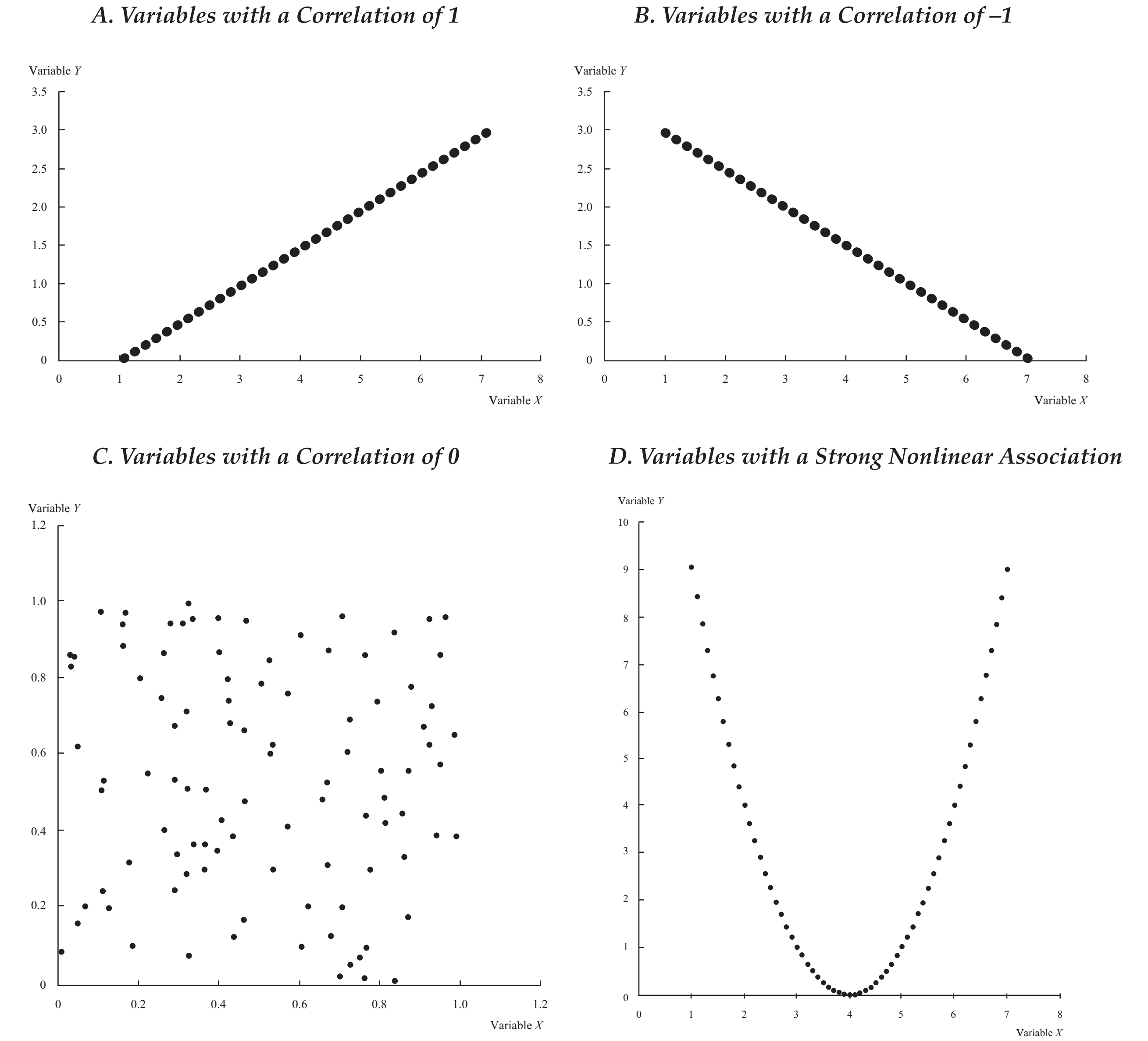
Figure 4 shows some scatter plots. Part A of the figure shows the scatter plot of two variables with a correlation of 1. Note that all the points on the scatter plot in Part A lie on a straight line with a positive slope. Whenever variable $X$ increases by one unit, variable $Y$ increases by half a unit. Because all of the points in the graph lie on a straight line, an increase of one unit in $X$ is associated with exactly the same half-unit increase in $Y$, regardless of the level of $X$. Even if the slope of the line in the figure were different (but positive), the correlation between the two variables would be 1 as long as all the points lie on that straight line.

Part B shows a scatter plot for two variables with a correlation coefficient of −1. Once again, the plotted observations fall on a straight line. In this graph, however, the line has a negative slope. As $X$ increases by one unit, $Y$ decreases by half a unit, regardless of the initial value of $X$.

Part C shows a scatter plot of two variables with a correlation of 0; they have no linear relation. This graph shows that the value of variable $X$ tells us absolutely nothing about the value of variable $Y$.

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

**Figure 4 Scatter Plots**




### *Limitations of Correlation Analysis.*

Part D of Figure 4 illustrates that correlation measures the linear association between two variables, but it may not always be reliable. Two variables can have a strong nonlinear relation and still have a very low correlation. For example, the relation $Y = (X - 4)^2$ is a nonlinear relation contrasted to the linear relation $Y = 2X - 4$. The nonlinear relation between variables $X$ and $Y$ is shown in Part D of Figure 4. Below a level of 4 for $X$, variable $Y$ decreases with increasing values of $X$. When $X$ is 4 or greater, however, $Y$ increases whenever $X$ increases. Even though these two variables are perfectly associated, there is no linear association between them.[17]

**17** The perfect association is the quadratic relationship $Y = (X - 4)^2$.

Correlation also may be an unreliable measure when outliers are present in one or both of the series. Outliers are small numbers of observations at either extreme (small or large) of a sample. The correlation may be quite sensitive to excluding outliers. In such a situation, we should consider whether it makes sense to exclude those outlier observations, and whether they are noise or news. As a general rule, we must determine whether a computed sample correlation changes greatly by removing a few outliers. But we must also use judgment to determine whether those outliers contain information about the two variables' relationship (and should thus be included in the correlation analysis) or contain no information (and should thus be excluded).

Keep in mind that correlation does not imply causation. Even if two variables are highly correlated, one does not necessarily cause the other in the sense that certain values of one variable bring about the occurrence of certain values of the other. Furthermore, correlations can be spurious in the sense of misleadingly pointing towards associations between variables.

The term **spurious correlation** has been used to refer to 1) correlation between two variables that reflects chance relationships in a particular data set, 2) correlation induced by a calculation that mixes each of two variables with a third, and 3) correlation between two variables arising not from a direct relation between them but from their relation to a third variable. As an example of the second kind of spurious correlation, two variables that are uncorrelated may be correlated if divided by a third variable. As an example of the third kind of spurious correlation, height may be positively correlated with the extent of a person's vocabulary, but the underlying relationships are between age and height and between age and vocabulary. Investment professionals must be cautious in basing investment strategies on high correlations. Spurious correlation may suggest investment strategies that appear profitable but actually would not be so, if implemented.

**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<br>$E(R_A)$ = 20% | B<br>$E(R_B)$ = 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:

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

$$\sigma(R_p) = 408.8125^{1/2} = 20.22 \text{ percent}$$

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.[18] 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 |

18 See any of the textbooks mentioned in Footnote 10.

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 |
| | | | | | $\text{Cov}(R_A,R_B) = 16$ |

*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 $\text{Cov}(R_A,R_B) = 16$.

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

$$\text{Cov}(R_A,R_B) = \sum_i \sum_j P(R_{A,i},R_{B,j})(R_{A,i} - ER_A)(R_{B,j} - ER_B) \quad (19)$$

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:

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

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.[19]

## TOPICS IN PROBABILITY

4

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.[20] Bayes' formula and related concepts have been applied in many business and investment decision-making contexts, including the evaluation of mutual fund performance.[21]

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

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

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

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*(*EPS exceeded consensus*) = 0.45
- *P*(*EPS met consensus*) = 0.30
- *P*(*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*(*DriveMed expands* | *EPS exceeded consensus*) = 0.75

*P*(*DriveMed expands* | *EPS met consensus*) = 0.20

*P*(*DriveMed expands* | *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:

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

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

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

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.

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

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

$$\begin{aligned}
&P(\textit{EPS exceeded consensus} \mid \textit{DriveMed expands}) + P(\textit{EPS met consensus} \mid \\
&\quad \textit{DriveMed expands}) + P(\textit{EPS fell short of consensus} \mid \textit{DriveMed expands}) \\
&\quad = 0.8232 + 0.1463 + 0.0305 = 1.0000
\end{aligned}$$

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,

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

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.[22] 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.

**22** 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

$$n! = n(n - 1)(n - 2)(n - 3)\ldots1$$

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.[23]

The next type of counting problem can be called labeling problems.[24] 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 last year. 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

23 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).

24 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

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

Here ${}_nC_r$ 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"):

$$\begin{aligned} {}_5C_3 &= 5!/[(5-3)!3!] \\ &= [(5)(4)(3)(2)(1)]/[(2)(1)(3)(2)(1)] = 120/12 = 10 \text{ ways} \end{aligned}$$

A useful fact can be illustrated as follows: ${}_5C_3 = 5!/(2!3!)$ equals ${}_5C_2 = 5!/(3!2!)$, as $3 + 2 = 5$; ${}_5C_4 = 5!/(1!4!)$ equals ${}_5C_1 = 5!/(4!1!)$, 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.[25]

- **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.[26]

---

**25** A more formal definition states that a permutation is an ordered subset of $n$ distinct objects.
**26** 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 $R_j$ in a forward-looking sense 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.
- The historical or sample covariance between two random variables $R_i$ and $R_j$ based on a sample of past data of size $n$ is the average value of the product of the deviations of observations on two random variables from their sample means:

$$\text{Cov}\left(R_i,R_j\right) = \sum_{i=1}^{n}\left(R_{i,t} - \bar{R}_i\right)\left(R_{j,t} - \bar{R}_j\right)\Big/\left(n-1\right)$$

- Correlation is a number between −1 and +1 that measures the co-movement (linear association) between two random variables: $\rho(R_i,R_j) = \text{Cov}(R_i,R_j)/[\sigma(R_i)\,\sigma(R_j)]$.
- If two variables have a very strong linear relation, then the absolute value of their correlation will be close to 1. If two variables have a weak linear relation, then the absolute value of their correlation will be close to 0.
- If the correlation coefficient is positive, the two variables are directly related; if the correlation coefficient is negative, the two variables are inversely related.
- A scatter plot shows graphically the relationship between two variables. If the points on the scatter plot cluster together in a straight line, the two variables have a strong linear relation.
- Even one outlier can greatly affect the correlation between two variables. Analysts should examine a scatter plot for the variables to determine whether outliers might affect a particular correlation.
- Correlations can be spurious in the sense of misleadingly pointing toward associations between variables.
- 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: $P(\textit{nonsurvivor}) = 0.40$.
- Fifty-five percent of the companies to which the test is administered pass it: $P(\textit{pass test}) = 0.55$.
- The probability that a company will pass the test given that it will subsequently survive 12 months, is 0.85: $P(\textit{pass test} \mid \textit{survivor}) = 0.85$.

- **A** What is $P(\textit{pass test} \mid \textit{nonsurvivor})$?

**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.

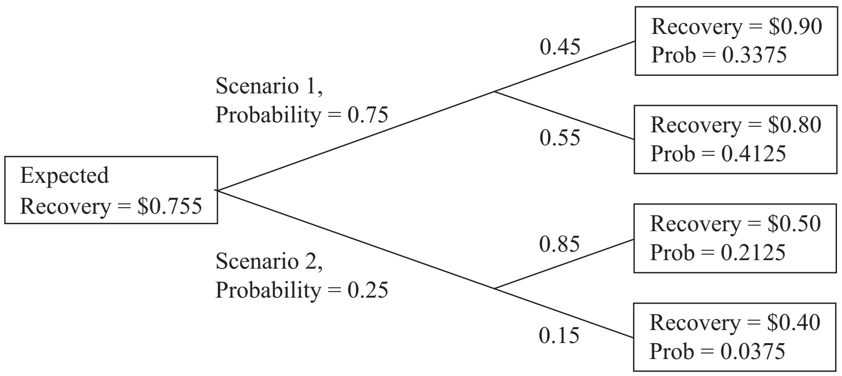
*Outcomes associated with Scenario 2:* With a 0.85 probability of a \$0.50 recovery per \$1 principal value, given Scenario 2, and with the probability of Scenario 2 equal to 0.25, the probability of recovering \$0.50 is 0.85(0.25) = 0.2125. By a similar calculation, the probability of recovering \$0.40 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) = $\sigma$(US bonds) × $\sigma$ (Spanish bonds) × $\rho$(US bond returns, Spanish bond returns) = $0.64 \times 0.56 \times 0.24 = 0.086$.

**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

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

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.