READING

9

Common Probability Distributions

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 probability distribution and distinguish between discrete and continuous random variables and their probability functions;
☐	**b.** describe the set of possible outcomes of a specified discrete random variable;
☐	**c.** interpret a cumulative distribution function;
☐	**d.** calculate and interpret probabilities for a random variable, given its cumulative distribution function;
☐	**e.** define a discrete uniform random variable, a Bernoulli random variable, and a binomial random variable;
☐	**f.** calculate and interpret probabilities given the discrete uniform and the binomial distribution functions;
☐	**g.** construct a binomial tree to describe stock price movement;
☐	**h.** define the continuous uniform distribution and calculate and interpret probabilities, given a continuous uniform distribution;
☐	**i.** explain the key properties of the normal distribution;
☐	**j.** distinguish between a univariate and a multivariate distribution and explain the role of correlation in the multivariate normal distribution;
☐	**k.** determine the probability that a normally distributed random variable lies inside a given interval;
☐	**l.** define the standard normal distribution, explain how to standardize a random variable, and calculate and interpret probabilities using the standard normal distribution;
☐	**m.** define shortfall risk, calculate the safety-first ratio, and select an optimal portfolio using Roy's safety-first criterion;

(continued)

LEARNING OUTCOMES

Mastery	The candidate should be able to:
☐	**n.** explain the relationship between normal and lognormal distributions and why the lognormal distribution is used to model asset prices;
☐	**o.** distinguish between discretely and continuously compounded rates of return and calculate and interpret a continuously compounded rate of return, given a specific holding period return;
☐	**p.** explain Monte Carlo simulation and describe its applications and limitations;
☐	**q.** compare Monte Carlo simulation and historical simulation.

1 INTRODUCTION TO COMMON PROBABILITY DISTRIBUTIONS

In nearly all investment decisions we work with random variables. The return on a stock and its earnings per share are familiar examples of random variables. To make probability statements about a random variable, we need to understand its probability distribution. A **probability distribution** specifies the probabilities of the possible outcomes of a random variable.

In this reading, we present important facts about four probability distributions and their investment uses. These four distributions—the uniform, binomial, normal, and lognormal—are used extensively in investment analysis. They are used in such basic valuation models as the Black–Scholes–Merton option pricing model, the binomial option pricing model, and the capital asset pricing model. With the working knowledge of probability distributions provided in this reading, you will also be better prepared to study and use other quantitative methods such as hypothesis testing, regression analysis, and time-series analysis.

After discussing probability distributions, we end the reading with an introduction to Monte Carlo simulation, a computer-based tool for obtaining information on complex problems. For example, an investment analyst may want to experiment with an investment idea without actually implementing it. Or she may need to price a complex option for which no simple pricing formula exists. In these cases and many others, Monte Carlo simulation is an important resource. To conduct a Monte Carlo simulation, the analyst must identify risk factors associated with the problem and specify probability distributions for them. Hence, Monte Carlo simulation is a tool that requires an understanding of probability distributions.

Before we discuss specific probability distributions, we define basic concepts and terms. We then illustrate the operation of these concepts through the simplest distribution, the uniform distribution. That done, we address probability distributions that have more applications in investment work but also greater complexity.

DISCRETE RANDOM VARIABLES

2

A **random variable** is a quantity whose future outcomes are uncertain. The two basic types of random variables are discrete random variables and continuous random variables. A **discrete random variable** can take on at most a countable number of possible values. For example, a discrete random variable X can take on a limited number of outcomes $x_1, x_2, \ldots, x_n$ (n possible outcomes), or a discrete random variable Y can take on an unlimited number of outcomes $y_1, y_2, \ldots$ (without end).[1] Because we can count all the possible outcomes of X and Y (even if we go on forever in the case of Y), both X and Y satisfy the definition of a discrete random variable. By contrast, we cannot count the outcomes of a **continuous random variable**. We cannot describe the possible outcomes of a continuous random variable Z with a list $z_1, z_2, \ldots$ because the outcome $(z_1 + z_2)/2$, not in the list, would always be possible. Rate of return is an example of a continuous random variable.

In working with a random variable, we need to understand its possible outcomes. For example, a majority of the stocks traded on the New Zealand Stock Exchange are quoted in ticks of NZ$0.01. Quoted stock price is thus a discrete random variable with possible values NZ$0, NZ$0.01, NZ$0.02, … But we can also model stock price as a continuous random variable (as a lognormal random variable, to look ahead). In many applications, we have a choice between using a discrete or a continuous distribution. We are usually guided by which distribution is most efficient for the task we face. This opportunity for choice is not surprising, as many discrete distributions can be approximated with a continuous distribution, and vice versa. In most practical cases, a probability distribution is only a mathematical idealization, or approximate model, of the relative frequencies of a random variable's possible outcomes.

EXAMPLE 1

The Distribution of Bond Price

You are researching a probability model for bond price, and you begin by thinking about the characteristics of bonds that affect price. What are the lowest and the highest possible values for bond price? Why? What are some other characteristics of bonds that may affect the distribution of bond price?

The lowest possible value of bond price is 0, when the bond is worthless. Identifying the highest possible value for bond price is more challenging. The promised payments on a coupon bond are the coupons (interest payments) plus the face amount (principal). The price of a bond is the present discounted value of these promised payments. Because investors require a return on their investments, 0 percent is the lower limit on the discount rate that investors would use to discount a bond's promised payments. At a discount rate of 0 percent, the price of a bond is the sum of the face value and the remaining coupons without any discounting. The discount rate thus places the upper limit on bond price. Suppose, for example, that face value is $1,000 and two $40 coupons remain; the interval $0 to $1,080 captures all possible values of the bond's price. This upper limit decreases through time as the number of remaining payments decreases.

1 We follow the convention that an uppercase letter represents a random variable and a lowercase letter represents an outcome or specific value of the random variable. Thus X refers to the random variable, and x refers to an outcome of X. We subscript outcomes, as in x_1 and x_2, when we need to distinguish among different outcomes in a list of outcomes of a random variable.

Other characteristics of a bond also affect its price distribution. Pull to par value is one such characteristic: As the maturity date approaches, the standard deviation of bond price tends to grow smaller as bond price converges to par value. Embedded options also affect bond price. For example, with bonds that are currently callable, the issuer may retire the bonds at a prespecified premium above par; this option of the issuer cuts off part of the bond's upside. Modeling bond price distribution is a challenging problem.

Every random variable is associated with a probability distribution that describes the variable completely. We can view a probability distribution in two ways. The basic view is the **probability function**, which specifies the probability that the random variable takes on a specific value: $P(X = x)$ is the probability that a random variable X takes on the value x. (Note that capital X represents the random variable and lowercase x represents a specific value that the random variable may take.) For a discrete random variable, the shorthand notation for the probability function is $p(x) = P(X = x)$. For continuous random variables, the probability function is denoted $f(x)$ and called the **probability density function** (pdf), or just the density.[2]

A probability function has two key properties (which we state, without loss of generality, using the notation for a discrete random variable):

- $0 \le p(x) \le 1$, because probability is a number between 0 and 1.
- The sum of the probabilities $p(x)$ over all values of X equals 1. If we add up the probabilities of all the distinct possible outcomes of a random variable, that sum must equal 1.

We are often interested in finding the probability of a range of outcomes rather than a specific outcome. In these cases, we take the second view of a probability distribution, the cumulative distribution function (cdf). The **cumulative distribution function**, or distribution function for short, gives the probability that a random variable X is less than or equal to a particular value x, $P(X \le x)$. For both discrete and continuous random variables, the shorthand notation is $F(x) = P(X \le x)$. How does the cumulative distribution function relate to the probability function? The word "cumulative" tells the story. To find $F(x)$, we sum up, or cumulate, values of the probability function for all outcomes less than or equal to x. The function of the cdf is parallel to that of cumulative relative frequency, which we discussed in the reading on statistical concepts and market returns.

Next, we illustrate these concepts with examples and show how we use discrete and continuous distributions. We start with the simplest distribution, the discrete uniform.

2.1 The Discrete Uniform Distribution

The simplest of all probability distributions is the discrete uniform distribution. Suppose that the possible outcomes are the integers (whole numbers) 1 to 8, inclusive, and the probability that the random variable takes on any of these possible values is the same for all outcomes (that is, it is uniform). With eight outcomes, $p(x) = 1/8$, or 0.125, for all values of X (X = 1, 2, 3, 4, 5, 6, 7, 8); the statement just made is a complete description of this discrete uniform random variable. The distribution has a finite number of specified outcomes, and each outcome is equally likely. Table 1 summarizes the two views of this random variable, the probability function and the cumulative distribution function.

2 The technical term for the probability function of a discrete random variable, probability mass function (pmf), is used less frequently.

Table 1 Probability Function and Cumulative Distribution Function for a Discrete Uniform Random Variable

$X = x$	Probability Function $p(x) = P(X = x)$	Cumulative Distribution Function $F(x) = P(X \le x)$
1	0.125	0.125
2	0.125	0.250
3	0.125	0.375
4	0.125	0.500
5	0.125	0.625
6	0.125	0.750
7	0.125	0.875
8	0.125	1.000

We can use Table 1 to find three probabilities: $P(X \le 7)$, $P(4 \le X \le 6)$, and $P(4 < X \le 6)$. The following examples illustrate how to use the cdf to find the probability that a random variable will fall in any interval (for any random variable, not only the uniform).

- The probability that X is less than or equal to 7, $P(X \le 7)$, is the next-to-last entry in the third column, 0.875 or 87.5 percent.
- To find $P(4 \le X \le 6)$, we need to find the sum of three probabilities: $p(4)$, $p(5)$, and $p(6)$. We can find this sum in two ways. We can add $p(4)$, $p(5)$, and $p(6)$ from the second column. Or we can calculate the probability as the difference between two values of the cumulative distribution function:

  $$F(6) = P(X \le 6) = p(6) + p(5) + p(4) + p(3) + p(2) + p(1)$$

  $$F(3) = P(X \le 3) = p(3) + p(2) + p(1)$$

  so

  $$P(4 \le X \le 6) = F(6) - F(3) = p(6) + p(5) + p(4) = 3/8$$

  So we calculate the second probability as $F(6) - F(3) = 3/8$.
- The third probability, $P(4 < X \le 6)$, the probability that X is less than or equal to 6 but greater than 4, is $p(5) + p(6)$. We compute it as follows, using the cdf:

  $$P(4 < X \le 6) = P(X \le 6) - P(X \le 4) = F(6) - F(4) = p(6) + p(5) = 2/8$$

  So we calculate the third probability as $F(6) - F(4) = 2/8$.

Suppose we want to check that the discrete uniform probability function satisfies the general properties of a probability function given earlier. The first property is $0 \le p(x) \le 1$. We see that $p(x) = 1/8$ for all x in the first column of the table. (Note that $p(x)$ equals 0 for numbers x such as −14 or 12.215 that are not in that column.) The first property is satisfied. The second property is that the probabilities sum to 1. The entries in the second column of Table 1 do sum to 1.

The cdf has two other characteristic properties:

- The cdf lies between 0 and 1 for any x: $0 \le F(x) \le 1$.
- As we increase x, the cdf either increases or remains constant.

Check these statements by looking at the third column in Table 1.

We now have some experience working with probability functions and cdfs for discrete random variables. Later in this reading, we will discuss Monte Carlo simulation, a methodology driven by random numbers. As we will see, the uniform distribution has an important technical use: It is the basis for generating random numbers, which in turn produce random observations for all other probability distributions.[3]

2.2 The Binomial Distribution

In many investment contexts, we view a result as either a success or a failure, or as binary (twofold) in some other way. When we make probability statements about a record of successes and failures, or about anything with binary outcomes, we often use the binomial distribution. What is a good model for how a stock price moves through time? Different models are appropriate for different uses. Cox, Ross, and Rubinstein (1979) developed an option pricing model based on binary moves, price up or price down, for the asset underlying the option. Their binomial option pricing model was the first of a class of related option pricing models that have played an important role in the development of the derivatives industry. That fact alone would be sufficient reason for studying the binomial distribution, but the binomial distribution has uses in decision-making as well.

The building block of the binomial distribution is the **Bernoulli random variable**, named after the Swiss probabilist Jakob Bernoulli (1654–1704). Suppose we have a trial (an event that may repeat) that produces one of two outcomes. Such a trial is a **Bernoulli trial**. If we let Y equal 1 when the outcome is success and Y equal 0 when the outcome is failure, then the probability function of the Bernoulli random variable Y is

$$p(1) = P(Y = 1) = p$$

$$p(0) = P(Y = 0) = 1 - p$$

where p is the probability that the trial is a success. Our next example is the very first step on the road to understanding the binomial option pricing model.

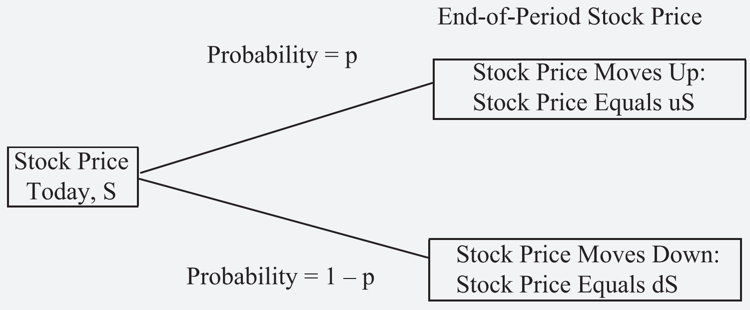
EXAMPLE 2

One-Period Stock Price Movement as a Bernoulli Random Variable

Suppose we describe stock price movement in the following way. Stock price today is S. Next period stock price can move up or down. The probability of an up move is p, and the probability of a down move is $1 - p$. Thus, stock price is a Bernoulli random variable with probability of success (an up move) equal to p. When the stock moves up, ending price is uS, with u equal to 1 plus the rate of return if the stock moves up. For example, if the stock earns 0.01 or 1 percent on an up move, $u = 1.01$. When the stock moves down, ending price is dS, with d equal to 1 plus the rate of return if the stock moves down. For example, if the stock earns −0.01 or −1 percent on a down move, $d = 0.99$. Figure 1 shows a diagram of this model of stock price dynamics.

3 See Hillier (2014). Random numbers initially generated by computers are usually random positive integer numbers that are converted to approximate continuous uniform random numbers between 0 and 1. Then the continuous uniform random numbers are used to produce random observations on other distributions, such as the normal, using various techniques. We will discuss random observation generation further in the section on Monte Carlo simulation.

Figure 1 One-Period Stock Price as a Bernoulli Random Variable



We will continue with the above example later. In the model of stock price movement in Example 2, success and failure at a given trial relate to up moves and down moves, respectively. In the following example, success is a profitable trade and failure is an unprofitable one.

EXAMPLE 3

A Trading Desk Evaluates Block Brokers (1)

You work in equities trading at an institutional money manager that regularly trades with several block brokers. Blocks are orders to sell or buy that are too large for the liquidity ordinarily available in dealer networks or stock exchanges. Your firm has known interests in certain kinds of stock. Block brokers call your trading desk when they want to sell blocks of stocks that they think your firm may be interested in buying. You know that these transactions have definite risks. For example, if the broker's client (the seller of the shares) has unfavorable information on the stock, or if the total amount he is selling through all channels is not truthfully communicated to you, you may see an immediate loss on the trade. From time to time, your firm audits the performance of block brokers. Your firm calculates the post-trade, market-risk-adjusted dollar returns on stocks purchased from block brokers. On that basis, you classify each trade as unprofitable or profitable. You have summarized the performance of the brokers in a spreadsheet, excerpted in Table 2 for November of last year. (The broker names are coded BB001 and BB002.)

Table 2 Block Trading Gains and Losses

	Profitable Trades	Losing Trades
BB001	3	9
BB002	5	3

View each trade as a Bernoulli trial. Calculate the percentage of profitable trades with the two block brokers for last November. These are estimates of p, the underlying probability of a successful (profitable) trade with each broker.

Your firm has logged 3 + 9 = 12 trades (the row total) with block broker BB001. Because 3 of the 12 trades were profitable, the percentage of profitable trades was 3/12 or 25 percent. With broker BB002, the percentage of profitable trades was 5/8 or 62.5 percent. A trade is a Bernoulli trial, and the above calculations provide estimates of the underlying probability of a profitable trade (success) with the two brokers. For broker BB001, your estimate is $\hat{p}$ = 0.25; for broker BB002, your estimate is $\hat{p}$ = 0.625.[4]

In n Bernoulli trials, we can have 0 to n successes. If the outcome of an individual trial is random, the total number of successes in n trials is also random. A **binomial random variable** X is defined as the number of successes in n Bernoulli trials. A binomial random variable is the sum of Bernoulli random variables Y_i, $i = 1, 2, \ldots, n$:

$$X = Y_1 + Y_2 + \ldots + Y_n$$

where Y_i is the outcome on the ith trial (1 if a success, 0 if a failure). We know that a Bernoulli random variable is defined by the parameter p. The number of trials, n, is the second parameter of a binomial random variable. The binomial distribution makes these assumptions:

- The probability, p, of success is constant for all trials.
- The trials are independent.

The second assumption has great simplifying force. If individual trials were correlated, calculating the probability of a given number of successes in n trials would be much more complicated.

Under the above two assumptions, a binomial random variable is completely described by two parameters, n and p. We write

$$X \sim B(n,p)$$

which we read as "X has a binomial distribution with parameters n and p." You can see that a Bernoulli random variable is a binomial random variable with $n = 1$: $Y \sim B(1, p)$.

Now we can find the general expression for the probability that a binomial random variable shows x successes in n trials. We can think in terms of a model of stock price dynamics that can be generalized to allow any possible stock price movements if the periods are made extremely small. Each period is a Bernoulli trial: With probability p, the stock price moves up; with probability $1 - p$, the price moves down. A success is an up move, and x is the number of up moves or successes in n periods (trials). With each period's moves independent and p constant, the number of up moves in n periods is a binomial random variable. We now develop an expression for $P(X = x)$, the probability function for a binomial random variable.

Any sequence of n periods that shows exactly x up moves must show $n - x$ down moves. We have many different ways to order the up moves and down moves to get a total of x up moves, but given independent trials, any sequence with x up moves must occur with probability $p^x(1 - p)^{n-x}$. Now we need to multiply this probability by the number of different ways we can get a sequence with x up moves. Using a basic result in counting from the reading on probability concepts, there are

$$\frac{n!}{(n - x)!x!}$$

4 The "hat" over p indicates that it is an estimate of p, the underlying probability of a profitable trade with the broker.

different sequences in n trials that result in x up moves (or successes) and $n - x$ down moves (or failures). Recall from the reading on probability concepts that n factorial ($n!$) is defined as $n(n - 1)(n - 2) \ldots 1$ (and $0! = 1$ by convention). For example, $5! = (5)(4)(3)(2)(1) = 120$. The combination formula $n!/[(n - x)!x!]$ is denoted by

$$\binom{n}{x}$$

(read "n combination x" or "n choose x"). For example, over three periods, exactly three different sequences have two up moves: UUD, UDU, and DUU. We confirm this by

$$\binom{3}{2} = \frac{3!}{(3-2)!2!} = \frac{(3)(2)(1)}{(1)(2)(1)} = 3$$

If, hypothetically, each sequence with two up moves had a probability of 0.15, then the total probability of two up moves in three periods would be $3 \times 0.15 = 0.45$. This example should persuade you that for X distributed $B(n, p)$, the probability of x successes in n trials is given by

$$p(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x} \tag{1}$$

Some distributions are always symmetric, such as the normal, and others are always asymmetric or skewed, such as the lognormal. The binomial distribution is symmetric when the probability of success on a trial is 0.50, but it is asymmetric or skewed otherwise.

We illustrate Equation 1 (the probability function) and the cdf through the symmetrical case. Consider a random variable distributed $B(n = 5, p = 0.50)$. Table 3 contains a complete description of this random variable. The fourth column of Table 3 is Column 2, n combination x, times Column 3, $p^x(1 - p)^{n-x}$; Column 4 gives the probability for each value of the number of up moves from the first column. The fifth column, cumulating the entries in the fourth column, is the cumulative distribution function.

Table 3 Binomial Probabilities, $p = 0.50$ and $n = 5$

Number of Up Moves, x (1)	Number of Possible Ways to Reach x Up Moves (2)	Probability for Each Way (3)	Probability for x, $p(x)$ (4) = (2) × (3)	$F(x) = P(X \le x)$ (5)
0	1	$0.50^0(1 - 0.50)^5 = 0.03125$	0.03125	0.03125
1	5	$0.50^1(1 - 0.50)^4 = 0.03125$	0.15625	0.18750
2	10	$0.50^2(1 - 0.50)^3 = 0.03125$	0.31250	0.50000
3	10	$0.50^3(1 - 0.50)^2 = 0.03125$	0.31250	0.81250
4	5	$0.50^4(1 - 0.50)^1 = 0.03125$	0.15625	0.96875
5	1	$0.50^5(1 - 0.50)^0 = 0.03125$	0.03125	1.00000

What would happen if we kept $n = 5$ but sharply lowered the probability of success on a trial to 10 percent? "Probability for Each Way" for $X = 0$ (no up moves) would then be about 59 percent: $0.10^0(1 - 0.10)^5 = 0.59049$. Because zero successes could still happen one way (Column 2), $p(0) = 59$ percent. You may want to check that given $p = 0.10$, $P(X \le 2) = 99.14$ percent: The probability of two or fewer up moves would be more than 99 percent. The random variable's probability would be massed on 0, 1, and 2 up moves, and the probability of larger outcomes would be minute. The outcomes of 3 and larger would be the long right tail, and the distribution would be right skewed. On the other hand, if we set $p = 0.90$, we would have the mirror image of the distribution with $p = 0.10$. The distribution would be left skewed.

With an understanding of the binomial probability function in hand, we can continue with our example of block brokers.

EXAMPLE 4

A Trading Desk Evaluates Block Brokers (2)

You now want to evaluate the performance of the block brokers in Example 3. You begin with two questions:

1 If you are paying a fair price on average in your trades with a broker, what should be the probability of a profitable trade?

2 Did each broker meet or miss that expectation on probability?

You also realize that the brokers' performance has to be evaluated in light of the sample's size, and for that you need to use the binomial probability function (Equation 1). You thus address the following (referring to the data in Example 3):

3 Under the assumption that the prices of trades were fair,

  A calculate the probability of three or fewer profitable trades with broker BB001.

  B calculate the probability of five or more profitable trades with broker BB002.

Solution to 1 and 2:

If the price you trade at is fair, 50 percent of the trades you do with a broker should be profitable.[5] The rate of profitable trades with broker BB001 was 25 percent. Therefore, broker BB001 missed your performance expectation. Broker BB002, at 62.5 percent profitable trades, exceeded your expectation.

Solution to 3:

A For broker BB001, the number of trades (the trials) was $n = 12$, and 3 were profitable. You are asked to calculate the probability of three or fewer profitable trades, $F(3) = p(3) + p(2) + p(1) + p(0)$.

Suppose the underlying probability of a profitable trade with BB001 is $p = 0.50$. With $n = 12$ and $p = 0.50$, according to Equation 1 the probability of three profitable trades is

$$p(3) = \binom{n}{x} p^x (1 - p)^{n-x} = \binom{12}{3}\left(0.50^3\right)\left(0.50^9\right)$$

$$= \frac{12!}{(12 - 3)!3!} 0.50^{12} = 220(0.000244) = 0.053711$$

5 Of course, you need to adjust for the direction of the overall market after the trade (any broker's record will be helped by a bull market) and perhaps make other risk adjustments. Assume that these adjustments have been made.

The probability of exactly 3 profitable trades out of 12 is 5.4 percent if broker BB001 were giving you fair prices. Now you need to calculate the other probabilities:

$$p(2) = [12!/(12 - 2)!2!](0.50^2)(0.50^{10}) = 66(0.000244) = 0.016113$$

$$p(1) = [12!/(12 - 1)!1!](0.50^1)(0.50^{11}) = 12(0.000244) = 0.00293$$

$$p(0) = [12!/(12 - 0)!0!](0.50^0)(0.50^{12}) = 1(0.000244) = 0.000244$$

Adding all the probabilities, $F(3) = 0.053711 + 0.016113 + 0.00293 + 0.000244 = 0.072998$ or 7.3 percent. The probability of doing 3 or fewer profitable trades out of 12 would be 7.3 percent if your trading desk were getting fair prices from broker BB001.

B For broker BB002, you are assessing the probability that the underlying probability of a profitable trade with this broker was 50 percent, despite the good results. The question was framed as the probability of doing five or more profitable trades if the underlying probability is 50 percent: $1 - F(4) = p(5) + p(6) + p(7) + p(8)$. You could calculate $F(4)$ and subtract it from 1, but you can also calculate $p(5) + p(6) + p(7) + p(8)$ directly.

You begin by calculating the probability that exactly 5 out of 8 trades would be profitable if BB002 were giving you fair prices:

$$\begin{aligned} p(5) &= \binom{8}{5}\left(0.50^5\right)\left(0.50^3\right) \\ &= 56\left(0.003906\right) = 0.21875 \end{aligned}$$

The probability is about 21.9 percent. The other probabilities are

$$p(6) = 28(0.003906) = 0.109375$$

$$p(7) = 8(0.003906) = 0.03125$$

$$p(8) = 1(0.003906) = 0.003906$$

So $p(5) + p(6) + p(7) + p(8) = 0.21875 + 0.109375 + 0.03125 + 0.003906 = 0.363281$ or 36.3 percent.[6] A 36.3 percent probability is substantial; the underlying probability of executing a fair trade with BB002 might well have been 0.50 despite your success with BB002 in November of last year. If one of the trades with BB002 had been reclassified from profitable to unprofitable, exactly half the trades would have been profitable. In summary, your trading desk is getting at least fair prices from BB002; you will probably want to accumulate additional evidence before concluding that you are trading at better-than-fair prices.

The magnitude of the profits and losses in these trades is another important consideration. If all profitable trades had small profits but all unprofitable trades had large losses, for example, you might lose money on your trades even if the majority of them were profitable.

In the next example, the binomial distribution helps in evaluating the performance of an investment manager.

6 In this example all calculations were worked through by hand, but binomial probability and cdf functions are also available in computer spreadsheet programs.

EXAMPLE 5

Meeting a Tracking Objective

You work for a pension fund sponsor. You have assigned a new money manager to manage a $500 million portfolio indexed on the MSCI EAFE (Europe, Australasia, and Far East) Index, which is designed to measure developed-market equity performance excluding the United States and Canada. After research, you believe it is reasonable to expect that the manager will keep portfolio return within a band of 75 basis points (bps) of the benchmark's return, on a quarterly basis.[7] To quantify this expectation further, you will be satisfied if portfolio return is within the 75 bps band 90 percent of the time. The manager meets the objective in six out of eight quarters. Of course, six out of eight quarters is a 75 percent success rate. But how does the manager's record precisely relate to your expectation of a 90 percent success rate and the sample size, 8 observations? To answer this question, you must find the probability that, given an assumed true or underlying success rate of 90 percent, performance could be as bad as or worse than that delivered. Calculate the probability (by hand or with a spreadsheet).

Specifically, you want to find the probability that portfolio return is within the 75 bps band in six or fewer quarters out of the eight in the sample. With $n = 8$ and $p = 0.90$, this probability is $F(6) = p(6) + p(5) + p(4) + p(3) + p(2) + p(1) + p(0)$. Start with

$$p(6) = (8!/6!2!)(0.90^6)(0.10^2) = 28(0.005314) = 0.148803$$

and work through the other probabilities:

$$p(5) = (8!/5!3!)(0.90^5)(0.10^3) = 56(0.00059) = 0.033067$$
$$p(4) = (8!/4!4!)(0.90^4)(0.10^4) = 70(0.000066) = 0.004593$$
$$p(3) = (8!/3!5!)(0.90^3)(0.10^5) = 56(0.000007) = 0.000408$$
$$p(2) = (8!/2!6!)(0.90^2)(0.10^6) = 28(0.000001) = 0.000023$$
$$p(1) = (8!/1!7!)(0.90^1)(0.10^7) = 8(0.00000009) = 0.00000072$$
$$p(0) = (8!/0!8!)(0.90^0)(0.10^8) = 1(0.00000001) = 0.00000001$$

Summing all these probabilities, you conclude that $F(6) = 0.148803 + 0.033067 + 0.004593 + 0.000408 + 0.000023 + 0.00000072 + 0.00000001 = 0.186895$ or 18.7 percent. There is a moderate 18.7 percent probability that the manager would show the record he did (or a worse record) if he had the skill to meet your expectations 90 percent of the time.

You can use other evaluation concepts such as tracking error or tracking risk, defined as the standard deviation of return differences between a portfolio and its benchmark, to assess the manager's performance. The calculation above would be only one input into any conclusions that you reach concerning the manager's performance. But to answer problems involving success rates, you need to be skilled in using the binomial distribution.

---

7 A basis point is one-hundredth of 1 percent (0.01 percent).

Two descriptors of a distribution that are often used in investments are the mean and the variance (or the standard deviation, the positive square root of variance).[8] Table 4 gives the expressions for the mean and variance of binomial random variables.

Table 4 Mean and Variance of Binomial Random Variables

	Mean	Variance
Bernoulli, $B(1, p)$	p	$p(1 - p)$
Binomial, $B(n, p)$	np	$np(1 - p)$

Because a single Bernoulli random variable, $Y \sim B(1, p)$, takes on the value 1 with probability p and the value 0 with probability $1 - p$, its mean or weighted-average outcome is p. Its variance is $p(1 - p)$.[9] A general binomial random variable, $B(n, p)$, is the sum of n Bernoulli random variables, and so the mean of a $B(n, p)$ random variable is np. Given that a $B(1, p)$ variable has variance $p(1 - p)$, the variance of a $B(n, p)$ random variable is n times that value, or $np(1 - p)$, assuming that all the trials (Bernoulli random variables) are independent. We can illustrate the calculation for two binomial random variables with differing probabilities as follows:

Random Variable	Mean	Variance
$B(n = 5, p = 0.50)$	$2.50 = 5(0.50)$	$1.25 = 5(0.50)(0.50)$
$B(n = 5, p = 0.10)$	$0.50 = 5(0.10)$	$0.45 = 5(0.10)(0.90)$

For a $B(n = 5, p = 0.50)$ random variable, the expected number of successes is 2.5 with a standard deviation of $1.118 = (1.25)^{1/2}$; for a $B(n = 5, p = 0.10)$ random variable, the expected number of successes is 0.50 with a standard deviation of $0.67 = (0.45)^{1/2}$.

EXAMPLE 6

The Expected Number of Defaults in a Bond Portfolio

Suppose as a bond analyst you are asked to estimate the number of bond issues expected to default over the next year in an unmanaged high-yield bond portfolio with 25 US issues from distinct issuers. The credit ratings of the bonds in the portfolio are tightly clustered around Moody's B2/Standard & Poor's B, meaning that the bonds are speculative with respect to the capacity to pay interest and repay principal. The estimated annual default rate for B2/B rated bonds is 10.7 percent.

1. Over the next year, what is the expected number of defaults in the portfolio, assuming a **binomial model** for defaults?
2. Estimate the standard deviation of the number of defaults over the coming year.
3. Critique the use of the binomial probability model in this context.

8 The mean (or arithmetic mean) is the sum of all values in a distribution or dataset, divided by the number of values summed. The variance is a measure of dispersion about the mean. See the reading on statistical concepts and market returns for further details on these concepts.

9 We can show that $p(1 - p)$ is the variance of a Bernoulli random variable as follows, noting that a Bernoulli random variable can take on only one of two values, 1 or 0: $\sigma^2(Y) = E[(Y - EY)^2] = E[(Y - p)^2] = (1 - p)^2 p + (0 - p)^2(1 - p) = (1 - p)[(1 - p)p + p^2] = p(1 - p)$.

Solution to 1:

For each bond, we can define a Bernoulli random variable equal to 1 if the bond defaults during the year and zero otherwise. With 25 bonds, the expected number of defaults over the year is np = 25(0.107) = 2.675 or approximately 3.

Solution to 2:

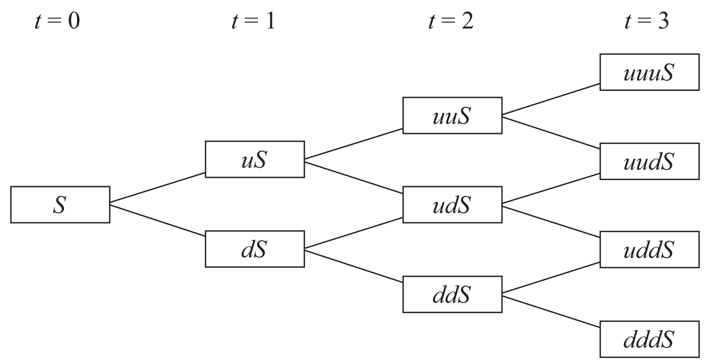
The variance is $np(1 - p)$ = 25(0.107)(0.893) = 2.388775. The standard deviation is $(2.388775)^{1/2}$ = 1.55. Thus, a two standard deviation confidence interval about the expected number of defaults would run from approximately 0 to approximately 6, for example.

Solution to 3:

An assumption of the binomial model is that the trials are independent. In this context, a trial relates to whether an individual bond issue will default over the next year. Because the issuing companies probably share exposure to common economic factors, the trials may not be independent. Nevertheless, for a quick estimate of the expected number of defaults, the binomial model may be adequate.

Earlier, we looked at a simple one-period model for stock price movement. Now we extend the model to describe stock price movement on three consecutive days. Each day is an independent trial. The stock moves up with constant probability p (the **up transition probability**); if it moves up, u is 1 plus the rate of return for an up move. The stock moves down with constant probability $1 - p$ (the **down transition probability**); if it moves down, d is 1 plus the rate of return for a down move. We graph stock price movement in Figure 2, where we now associate each of the n = 3 stock price moves with time indexed by t. The shape of the graph suggests why it is a called a **binomial tree**. Each boxed value from which successive moves or outcomes branch in the tree is called a **node**; in this example, a node is potential value for the stock price at a specified time.

Figure 2 A Binomial Model of Stock Price Movement



We see from the tree that the stock price at t = 3 has four possible values: $uuuS$, $uudS$, $uddS$, and $dddS$. The probability that the stock price equals *any* one of these four values is given by the binomial distribution. For example, three sequences of moves result in a final stock price of $uudS$: These are uud, udu, and duu. These sequences have two up moves out of three moves in total; the combination formula confirms that the number of ways to get two up moves (successes) in three periods (trials) is $3!/(3 - 2)!2! = 3$. Next note that each of these sequences, *uud*, *udu*, and *duu*, has probability $p^2(1 - p)$. So $P(S_3 = uudS) = 3p^2(1 - p)$, where S_3 indicates the stock's price after three moves.

The binomial random variable in this application is the number of up moves. Final stock price distribution is a function of the initial stock price, the *number* of up moves, and the *size* of the up moves and down moves. We cannot say that stock price itself is a binomial random variable; rather, it is a function of a binomial random variable, as well as of u and d, and initial price. This richness is actually one key to why this way of modeling stock price is useful: It allows us to choose values of these parameters to approximate various distributions for stock price (using a large number of time periods).[10] One distribution that can be approximated is the lognormal, an important continuous distribution model for stock price that we will discuss later. The flexibility extends further. In the tree shown above, the transition probabilities are the same at each node: p for an up move and $1 - p$ for a down move. That standard formula describes a process in which stock return volatility is constant through time. Option experts, however, sometimes model changing volatility through time using a binomial tree in which the probabilities for up and down moves differ at different nodes.

The binomial tree also supplies the possibility of testing a condition or contingency at any node. This flexibility is useful in investment applications such as option pricing. Consider an American call option on a dividend-paying stock. (Recall that an American option can be exercised at any time before expiration, at any node on the tree.) Just before an ex-dividend date, it may be optimal to exercise an American call option on stock to buy the stock and receive the dividend.[11] If we model stock price with a binomial tree, we can test, at each node, whether exercising the option is optimal. Also, if we know the value of the call at the four terminal nodes at $t = 3$ and we have a model for discounting values by one period, we can step backward one period to $t = 2$ to find the call's value at the three nodes there. Continuing back recursively, we can find the call's value today. This type of recursive operation is easily programmed on a computer. As a result, binomial trees can value options even more complex than American calls on stock.[12]

3 CONTINUOUS RANDOM VARIABLES

In the previous section, we considered discrete random variables (i.e., random variables whose set of possible outcomes is countable). In contrast, the possible outcomes of continuous random variables are never countable. If 1.250 is one possible value of a continuous random variable, for example, we cannot name the next higher or lower possible value. Technically, the range of possible outcomes of a continuous random variable is the real line (all real numbers between $-\infty$ and $+\infty$) or some subset of the real line.

In this section, we focus on the two most important continuous distributions in investment work, the normal and lognormal. As we did with discrete distributions, we introduce the topic through the uniform distribution.

---

10 For example, we can split 20 days into 100 subperiods, taking care to use compatible values for u and d.

11 Cash dividends represent a reduction of a company's assets. Early exercise may be optimal because the exercise price of options is typically not reduced by the amount of cash dividends, so cash dividends negatively affect the position of an American call option holder.

12 See Chance and Brooks (2016) for more information on option pricing models.

3.1 Continuous Uniform Distribution

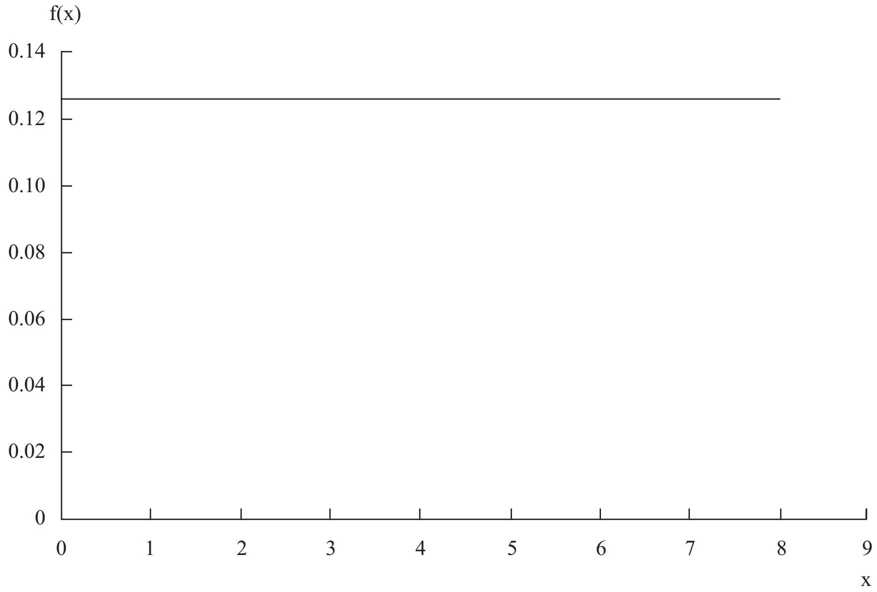
The continuous uniform distribution is the simplest continuous probability distribution. The uniform distribution has two main uses. As the basis of techniques for generating random numbers, the uniform distribution plays a role in Monte Carlo simulation. As the probability distribution that describes equally likely outcomes, the uniform distribution is an appropriate probability model to represent a particular kind of uncertainty in beliefs in which all outcomes appear equally likely.

The pdf for a uniform random variable is

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{for } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

For example, with $a = 0$ and $b = 8$, $f(x) = 1/8$ or 0.125. We graph this density in Figure 3.

Figure 3 Continuous Uniform Distribution



The graph of the density function plots as a horizontal line with a value of 0.125.

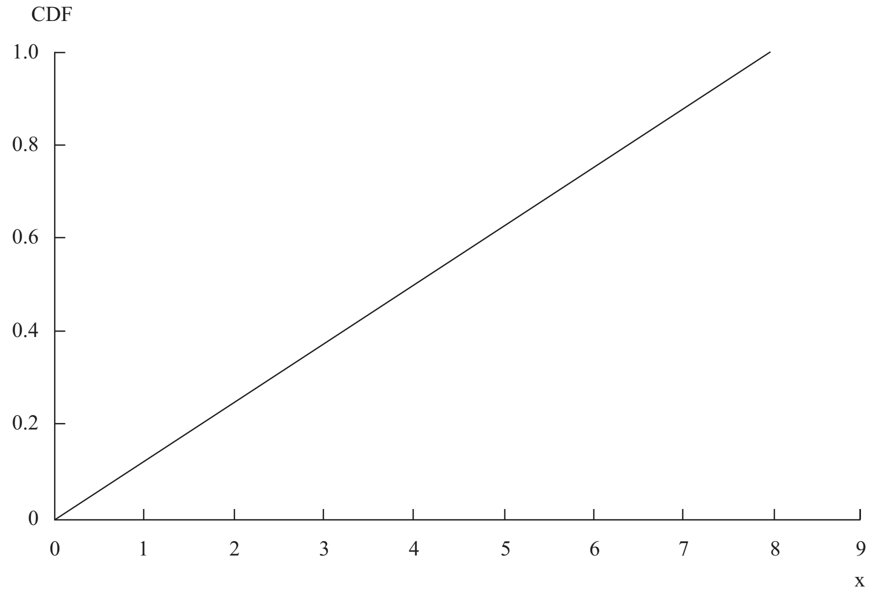
What is the probability that a uniform random variable with limits $a = 0$ and $b = 8$ is less than or equal to 3, or $F(3) = P(X \leq 3)$? When we were working with the discrete uniform random variable with possible outcomes 1, 2, …, 8, we summed individual probabilities: $p(1) + p(2) + p(3) = 0.375$. In contrast, the probability that a continuous uniform random variable, or any continuous random variable, assumes any given fixed value is 0. To illustrate this point, consider the narrow interval 2.510 to 2.511. Because that interval holds an infinity of possible values, the sum of the probabilities of values in that interval alone would be infinite if each individual value in it had a positive probability. To find the probability $F(3)$, we find the area under the curve graphing the pdf, between 0 to 3 on the x axis. In calculus, this operation is called integrating the probability function $f(x)$ from 0 to 3. This area under the curve is a rectangle with base $3 - 0 = 3$ and height 1/8. The area of this rectangle equals base times height: $3(1/8) = 3/8$ or 0.375. So $F(3) = 3/8$ or 0.375.

The interval from 0 to 3 is three-eighths of the total length between the limits of 0 and 8, and $F(3)$ is three-eighths of the total probability of 1. The middle line of the expression for the cdf captures this relationship.

$$F(x) = \begin{cases} 0 \text{ for } x \le a \\ \dfrac{x-a}{b-a} \text{ for } a < x < b \\ 1 \text{ for } x \ge b \end{cases}$$

For our problem, $F(x) = 0$ for $x \le 0$, $F(x) = x/8$ for $0 < x < 8$, and $F(x) = 1$ for $x \ge 8$. We graph this cdf in Figure 4.

Figure 4 Continuous Uniform Cumulative Distribution



The mathematical operation that corresponds to finding the area under the curve of a pdf $f(x)$ from a to b is the integral of $f(x)$ from a to b:

$$P(a \le X \le b) = \int_a^b f(x)dx \qquad \textbf{(2)}$$

where $\int dx$ is the symbol for summing $\int$ over small changes dx, and the limits of integration (a and b) can be any real numbers or $-\infty$ and $+\infty$. All probabilities of continuous random variables can be computed using Equation 2. For the uniform distribution example considered above, $F(7)$ is Equation 2 with lower limit $a = 0$ and upper limit $b = 7$. The integral corresponding to the cdf of a uniform distribution reduces to the three-line expression given previously. To evaluate Equation 2 for nearly all other continuous distributions, including the normal and lognormal, we rely on spreadsheet functions, computer programs, or tables of values to calculate probabilities. Those tools use various numerical methods to evaluate the integral in Equation 2.

Recall that the probability of a continuous random variable equaling any fixed point is 0. This fact has an important consequence for working with the cumulative distribution function of a continuous random variable: For any continuous random variable X, $P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$, because the probabilities at the endpoints a and b are 0. For discrete random variables, these relations of equality are not true, because probability accumulates at points.

EXAMPLE 7

Probability That a Lending Facility Covenant Is Breached

You are evaluating the bonds of a below-investment-grade borrower at a low point in its business cycle. You have many factors to consider, including the terms of the company's bank lending facilities. The contract creating a bank lending facility such as an unsecured line of credit typically has clauses known as covenants. These covenants place restrictions on what the borrower can do. The company will be in breach of a covenant in the lending facility if the interest coverage ratio, EBITDA/interest, calculated on EBITDA over the four trailing quarters, falls below 2.0. EBITDA is earnings before interest, taxes, depreciation, and amortization.[13] Compliance with the covenants will be checked at the end of the current quarter. If the covenant is breached, the bank can demand immediate repayment of all borrowings on the facility. That action would probably trigger a liquidity crisis for the company. With a high degree of confidence, you forecast interest charges of \$25 million. Your estimate of EBITDA runs from \$40 million on the low end to \$60 million on the high end.

Address two questions (treating projected interest charges as a constant):

1 If the outcomes for EBITDA are equally likely, what is the probability that EBITDA/interest will fall below 2.0, breaching the covenant?

2 Estimate the mean and standard deviation of EBITDA/interest. For a continuous uniform random variable, the mean is given by $\mu = (a + b)/2$ and the variance is given by $\sigma^2 = (b - a)^2/12$.

Solution to 1:

EBITDA/interest is a continuous uniform random variable because all outcomes are equally likely. The ratio can take on values between 1.6 = (\$40 million)/(\$25 million) on the low end and 2.4 = (\$60 million/\$25 million) on the high end. The range of possible values is 2.4 – 1.6 = 0.8. What fraction of the possible values falls below 2.0, the level that triggers default? The distance between 2.0 and 1.6 is 0.40; the value 0.40 is one-half the total length of 0.8, or 0.4/0.8 = 0.50. So, the probability that the covenant will be breached is 50 percent.

Solution to 2:

In Solution 1, we found that the lower limit of EBITDA/interest is 1.6. This lower limit is a. We found that the upper limit is 2.4. This upper limit is b. Using the formula given above,

$$\mu = (a + b)/2 = (1.6 + 2.4)/2 = 2.0$$

The variance of the interest coverage ratio is

$$\sigma^2 = (b - a)^2/12 = (2.4 - 1.6)^2/12 = 0.053333$$

The standard deviation is the positive square root of the variance, $0.230940 = (0.053333)^{1/2}$. The standard deviation is not particularly useful as a risk measure for a uniform distribution, however. The probability that lies within various standard deviation bands around the mean is sensitive to different specifications of the upper and lower limits (although Chebyshev's inequality is always satisfied).[14] Here, a one standard deviation interval around the mean of 2.0 runs from 1.769

13 For a detailed discussion on EBITDA, see the Level I CFA Program curriculum reading "Financial Reporting Quality."

14 Chebyshev's inequality is discussed in the reading on statistical concepts and market returns.

to 2.231 and captures 0.462/0.80 = 0.5775 or 57.8 percent of the probability. A two standard deviation interval runs from 1.538 to 2.462, which extends past both the lower and upper limits of the random variable.

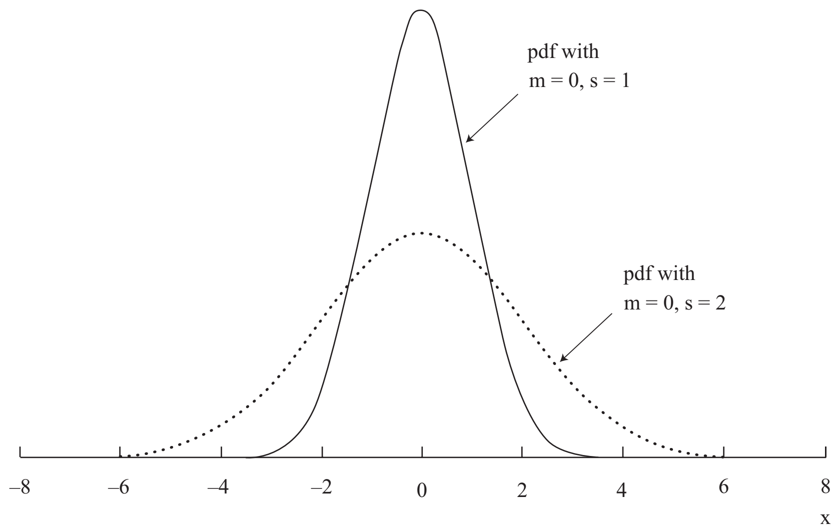
3.2 The Normal Distribution

The normal distribution may be the most extensively used probability distribution in quantitative work. It plays key roles in modern portfolio theory and in several risk management technologies. Because it has so many uses, the normal distribution must be thoroughly understood by investment professionals.

The role of the normal distribution in statistical inference and regression analysis is vastly extended by a crucial result known as the central limit theorem. The central limit theorem states that the sum (and mean) of a large number of independent random variables is approximately normally distributed.[15]

The French mathematician Abraham de Moivre (1667–1754) introduced the normal distribution in 1733 in developing a version of the central limit theorem. As Figure 5 shows, the normal distribution is symmetrical and bell-shaped. The range of possible outcomes of the normal distribution is the entire real line: all real numbers lying between $-\infty$ and $+\infty$. The tails of the bell curve extend without limit to the left and to the right.

Figure 5 Two Normal Distributions



The defining characteristics of a normal distribution are as follows:

- The normal distribution is completely described by two parameters—its mean, μ, and variance, σ^2. We indicate this as $X \sim N(\mu, \sigma^2)$ (read "X follows a normal distribution with mean μ and variance σ^2"). We can also define a normal distribution in terms of the mean and the standard deviation, σ (this is often convenient because σ is measured in the same units as X and μ). As a consequence, we can answer any probability question about a normal random variable if we know its mean and variance (or standard deviation).

15 The central limit theorem is discussed further in the reading on sampling.

- The normal distribution has a skewness of 0 (it is symmetric). The normal distribution has a kurtosis of 3; its excess kurtosis (kurtosis − 3.0) equals 0.[16] As a consequence of symmetry, the mean, median, and the mode are all equal for a normal random variable.
- A linear combination of two or more normal random variables is also normally distributed.

These bullet points concern a single variable or univariate normal distribution: the distribution of one normal random variable. A **univariate distribution** describes a single random variable. A **multivariate distribution** specifies the probabilities for a group of related random variables. You will encounter the **multivariate normal distribution** in investment work and reading and should know the following about it.

When we have a group of assets, we can model the distribution of returns on each asset individually, or the distribution of returns on the assets as a group. "As a group" means that we take account of all the statistical interrelationships among the return series. One model that has often been used for security returns is the multivariate normal distribution. A multivariate normal distribution for the returns on n stocks is completely defined by three lists of parameters:

- the list of the mean returns on the individual securities (n means in total);
- the list of the securities' variances of return (n variances in total); and
- the list of all the distinct pairwise return correlations: $n(n - 1)/2$ distinct correlations in total.[17]

The need to specify correlations is a distinguishing feature of the multivariate normal distribution in contrast to the univariate normal distribution.

The statement "assume returns are normally distributed" is sometimes used to mean a joint normal distribution. For a portfolio of 30 securities, for example, portfolio return is a weighted average of the returns on the 30 securities. A weighted average is a linear combination. Thus, portfolio return is normally distributed if the individual security returns are (joint) normally distributed. To review, in order to specify the normal distribution for portfolio return, we need the means, variances, and the distinct pairwise correlations of the component securities.

With these concepts in mind, we can return to the normal distribution for one random variable. The curves graphed in Figure 5 are the normal density function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \text{ for } -\infty < x < +\infty \qquad (3)$$

The two densities graphed in Figure 5 correspond to a mean of $\mu = 0$ and standard deviations of $\sigma = 1$ and $\sigma = 2$. The normal density with $\mu = 0$ and $\sigma = 1$ is called the **standard normal distribution** (or **unit normal distribution**). Plotting two normal distributions with the same mean and different standard deviations helps us appreciate why standard deviation is a good measure of dispersion for the normal distribution: Observations are much more concentrated around the mean for the normal distribution with $\sigma = 1$ than for the normal distribution with $\sigma = 2$.

16 If we have a sample of size n from a normal distribution, we may want to know the possible variation in sample skewness and kurtosis. For a normal random variable, the standard deviation of sample skewness is $6/n$ and the standard deviation of sample kurtosis is $24/n$.

17 For example, a distribution with two stocks (a bivariate normal distribution) has two means, two variances, and one correlation: 2(2 − 1)/2. A distribution with 30 stocks has 30 means, 30 variances, and 435 distinct correlations: 30(30 − 1)/2. The return correlation of Dow Chemical with American Express stock is the same as the correlation of American Express with Dow Chemical stock, so these are counted as one distinct correlation.

Although not literally accurate, the normal distribution can be considered an approximate model for returns. Nearly all the probability of a normal random variable is contained within three standard deviations of the mean. For realistic values of mean return and return standard deviation for many assets, the normal probability of outcomes below −100 percent is very small. Whether the approximation is useful in a given application is an empirical question. For example, the normal distribution is a closer fit for quarterly and yearly holding period returns on a diversified equity portfolio than it is for daily or weekly returns.[18] A persistent departure from normality in most equity return series is kurtosis greater than 3, the fat-tails problem. So when we approximate equity return distributions with the normal distribution, we should be aware that the normal distribution tends to underestimate the probability of extreme returns.[19] Option returns are skewed. Because the normal is a symmetrical distribution, we should be cautious in using the normal distribution to model the returns on portfolios containing significant positions in options.

The normal distribution, however, is less suitable as a model for asset prices than as a model for returns. A normal random variable has no lower limit. This characteristic has several implications for investment applications. An asset price can drop only to 0, at which point the asset becomes worthless. As a result, practitioners generally do not use the normal distribution to model the distribution of asset prices. Also note that moving from any level of asset price to 0 translates into a return of −100 percent. Because the normal distribution extends below 0 without limit, it cannot be literally accurate as a model for asset returns.

Having established that the normal distribution is the appropriate model for a variable of interest, we can use it to make the following probability statements:

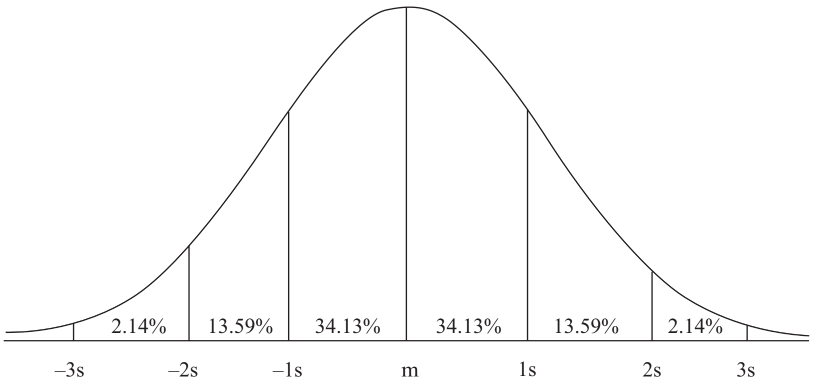
- Approximately 50 percent of all observations fall in the interval $\mu \pm (2/3)\sigma$.
- Approximately 68 percent of all observations fall in the interval $\mu \pm \sigma$.
- Approximately 95 percent of all observations fall in the interval $\mu \pm 2\sigma$.
- Approximately 99 percent of all observations fall in the interval $\mu \pm 3\sigma$.

One, two, and three standard deviation intervals are illustrated in Figure 6. The intervals indicated are easy to remember but are only approximate for the stated probabilities. More-precise intervals are $\mu \pm 1.96\sigma$ for 95 percent of the observations and $\mu \pm 2.58\sigma$ for 99 percent of the observations.

18 See Fama (1976) and Campbell, Lo, and MacKinlay (1997).

19 Fat tails can be modeled by a mixture of normal random variables or by a Student's *t*-distribution with a relatively small number of degrees of freedom. See Kon (1984) and Campbell, Lo, and MacKinlay (1997). We discuss the Student's *t*-distribution in the reading on sampling and estimation.

Figure 6 Units of Standard Deviation



In general, we do not observe the population mean or the population standard deviation of a distribution, so we need to estimate them.[20] We estimate the population mean, μ, using the sample mean, $\overline{X}$ (sometimes denoted as $\hat{\mu}$) and estimate the population standard deviation, σ, using the sample standard deviation, s (sometimes denoted as $\hat{\sigma}$).

There are as many different normal distributions as there are choices for mean (μ) and variance (σ^2). We can answer all of the above questions in terms of any normal distribution. Spreadsheets, for example, have functions for the normal cdf for any specification of mean and variance. For the sake of efficiency, however, we would like to refer all probability statements to a single normal distribution. The standard normal distribution (the normal distribution with μ = 0 and σ = 1) fills that role.

There are two steps in **standardizing** a random variable X: Subtract the mean of X from X, then divide that result by the standard deviation of X. If we have a list of observations on a normal random variable, X, we subtract the mean from each observation to get a list of deviations from the mean, then divide each deviation by the standard deviation. The result is the standard normal random variable, Z. (Z is the conventional symbol for a standard normal random variable.) If we have $X \sim N(\mu,\sigma^2)$ (read "X follows the normal distribution with parameters μ and σ^2"), we standardize it using the formula

$$Z = (X - \mu)/\sigma \tag{4}$$

Suppose we have a normal random variable, X, with μ = 5 and σ = 1.5. We standardize X with $Z = (X - 5)/1.5$. For example, a value $X = 9.5$ corresponds to a standardized value of 3, calculated as $Z = (9.5 - 5)/1.5 = 3$. The probability that we will observe a value as small as or smaller than 9.5 for $X \sim N(5,1.5)$ is exactly the same as the probability that we will observe a value as small as or smaller than 3 for $Z \sim N(0,1)$. We can answer all probability questions about X using standardized values and probability tables for Z. We generally do not know the population mean and standard deviation, so we often use the sample mean $\overline{X}$ for μ and the sample standard deviation s for σ.

20 A population is all members of a specified group, and the population mean is the arithmetic mean computed for the population. A sample is a subset of a population, and the sample mean is the arithmetic mean computed for the sample. For more information on these concepts, see the reading on statistical concepts and market returns.

Standard normal probabilities can also be computed with spreadsheets, statistical and econometric software, and programming languages. Tables of the cumulative distribution function for the standard normal random variable are in the back of this book. Table 5 shows an excerpt from those tables. $N(x)$ is a conventional notation for the cdf of a standard normal variable.[21]

Table 5 $P(Z \le x) = N(x)$ for $x \ge 0$ or $P(Z \le z) = N(z)$ for $z \ge 0$

x or *z*	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.10	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.20	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.30	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.40	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.50	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224

To find the probability that a standard normal variable is less than or equal to 0.24, for example, locate the row that contains 0.20, look at the 0.04 column, and find the entry 0.5948. Thus, $P(Z \le 0.24) = 0.5948$ or 59.48 percent.

The following are some of the most frequently referenced values in the standard normal table:

- The 90th percentile point is 1.282: $P(Z \le 1.282) = N(1.282) = 0.90$ or 90 percent, and 10 percent of values remain in the right tail.
- The 95th percentile point is 1.65: $P(Z \le 1.65) = N(1.65) = 0.95$ or 95 percent, and 5 percent of values remain in the right tail. Note the difference between the use of a percentile point when dealing with one tail rather than two tails. Earlier, we used 1.65 standard deviations for the 90 percent confidence interval, where 5 percent of values lie outside that interval on each of the two sides. Here we use 1.65 because we are concerned with the 5 percent of values that lie only on one side, the right tail.
- The 99th percentile point is 2.327: $P(Z \le 2.327) = N(2.327) = 0.99$ or 99 percent, and 1 percent of values remain in the right tail.

The tables that we give for the normal cdf include probabilities for $x \le 0$. Many sources, however, give tables only for $x \ge 0$. How would one use such tables to find a normal probability? Because of the symmetry of the normal distribution, we can find all probabilities using tables of the cdf of the standard normal random variable, $P(Z \le x) = N(x)$, for $x \ge 0$. The relations below are helpful for using tables for $x \ge 0$, as well as in other uses:

- For a non-negative number x, use $N(x)$ from the table. Note that for the probability to the right of x, we have $P(Z \ge x) = 1.0 - N(x)$.
- For a negative number $-x$, $N(-x) = 1.0 - N(x)$: Find $N(x)$ and subtract it from 1. All the area under the normal curve to the left of x is $N(x)$. The balance, $1.0 - N(x)$, is the area and probability to the right of x. By the symmetry of the normal distribution around its mean, the area and the probability to the right of x are equal to the area and the probability to the left of $-x$, $N(-x)$.
- For the probability to the right of $-x$, $P(Z \ge -x) = N(x)$.

21 Another often-seen notation for the cdf of a standard normal variable is $\Phi(x)$.

EXAMPLE 8

Probabilities for a Common Stock Portfolio

Assume the portfolio mean return is 12 percent and the standard deviation of return estimate is 22 percent per year.

You want to calculate the following probabilities, assuming that a normal distribution describes returns. (You can use the excerpt from the table of normal probabilities to answer these questions.)

1. What is the probability that portfolio return will exceed 20 percent?
2. What is the probability that portfolio return will be between 12 percent and 20 percent? In other words, what is P(12% ≤ Portfolio return ≤ 20%)?
3. You can buy a one-year T-bill that yields 5.5 percent. This yield is effectively a one-year risk-free interest rate. What is the probability that your portfolio's return will be equal to or less than the risk-free rate?

If X is portfolio return, standardized portfolio return is $Z = (X - \overline{X})/s = (X - 12\%)/22\%$. We use this expression throughout the solutions.

Solution to 1:

For X = 20%, Z = (20% − 12%)/22% = 0.363636. You want to find $P(Z > 0.363636)$. First note that $P(Z > x) = P(Z \geq x)$ because the normal is a continuous distribution. Recall that $P(Z \geq x) = 1.0 - P(Z \leq x)$ or $1 - N(x)$. Rounding 0.363636 to 0.36, according to the table, $N(0.36) = 0.6406$. Thus, 1 − 0.6406 = 0.3594. The probability that portfolio return will exceed 20 percent is about 36 percent if your normality assumption is accurate.

Solution to 2:

P(12% ≤ Portfolio return ≤ 20%) = N(Z corresponding to 20%) − N(Z corresponding to 12%). For the first term, Z = (20% − 12%)/22% = 0.36 approximately, and $N(0.36) = 0.6406$ (as in Solution 1). To get the second term immediately, note that 12 percent is the mean, and for the normal distribution 50 percent of the probability lies on either side of the mean. Therefore, N(Z corresponding to 12%) must equal 50 percent. So P(12% ≤ Portfolio return ≤ 20%) = 0.6406 − 0.50 = 0.1406 or approximately 14 percent.

Solution to 3:

If X is portfolio return, then we want to find P(Portfolio return ≤ 5.5%). This question is more challenging than Parts 1 or 2, but when you have studied the solution below you will have a useful pattern for calculating other shortfall probabilities.

There are three steps, which involve standardizing the portfolio return: First, subtract the portfolio mean return from each side of the inequality: P(Portfolio return − 12% ≤ 5.5% − 12%). Second, divide each side of the inequality by the standard deviation of portfolio return: P[(Portfolio return − 12%)/22% ≤ (5.5% − 12%)/22%] = $P(Z \leq -0.295455) = N(-0.295455)$. Third, recognize that on the left-hand side we have a standard normal variable, denoted by Z. As we pointed out above, $N(-x) = 1 - N(x)$. Rounding −0.29545 to −0.30 for use with the excerpted table, we have $N(-0.30) = 1 - N(0.30) = 1 - 0.6179 = 0.3821$, roughly 38 percent. The probability that your portfolio will underperform the one-year risk-free rate is about 38 percent.

We can get the answer above quickly by subtracting the mean portfolio return from 5.5 percent, dividing by the standard deviation of portfolio return, and evaluating the result (−0.295455) with the standard normal cdf.

3.3 Applications of the Normal Distribution

Modern portfolio theory (MPT) makes wide use of the idea that the value of investment opportunities can be meaningfully measured in terms of mean return and variance of return. In economic theory, **mean–variance analysis** holds exactly when investors are risk averse; when they choose investments so as to maximize expected utility, or satisfaction; and when either 1) returns are normally distributed, or 2) investors have quadratic utility functions.[22] Mean–variance analysis can still be useful, however—that is, it can hold approximately—when either assumption 1 or 2 is violated. Because practitioners prefer to work with observables such as returns, the proposition that returns are at least approximately normally distributed has played a key role in much of MPT.

Mean–variance analysis generally considers risk symmetrically in the sense that standard deviation captures variability both above and below the mean.[23] An alternative approach evaluates only downside risk. We discuss one such approach, safety-first rules, as it provides an excellent illustration of the application of normal distribution theory to practical investment problems. **Safety-first rules** focus on **shortfall risk**, the risk that portfolio value will fall below some minimum acceptable level over some time horizon. The risk that the assets in a defined benefit plan will fall below plan liabilities is an example of a shortfall risk.

Suppose an investor views any return below a level of R_L as unacceptable. Roy's safety-first criterion states that the optimal portfolio minimizes the probability that portfolio return, R_P, falls below the threshold level, R_L.[24] In symbols, the investor's objective is to choose a portfolio that minimizes $P(R_P < R_L)$. When portfolio returns are normally distributed, we can calculate $P(R_P < R_L)$ using the number of standard deviations that R_L lies below the expected portfolio return, $E(R_P)$. The portfolio for which $E(R_P) - R_L$ is largest relative to standard deviation minimizes $P(R_P < R_L)$. Therefore, if returns are normally distributed, the safety-first optimal portfolio *maximizes* the safety-first ratio (SFRatio):

$$\text{SFRatio} = [E(R_P) - R_L]/\sigma_P$$

The quantity $E(R_P) - R_L$ is the distance from the mean return to the shortfall level. Dividing this distance by σ_P gives the distance in units of standard deviation. There are two steps in choosing among portfolios using Roy's criterion (assuming normality):[25]

1. Calculate each portfolio's SFRatio.
2. Choose the portfolio with the highest SFRatio.

For a portfolio with a given safety-first ratio, the probability that its return will be less than R_L is $N(-\text{SFRatio})$, and the safety-first optimal portfolio has the lowest such probability. For example, suppose an investor's threshold return, R_L, is 2 percent. He is presented with two portfolios. Portfolio 1 has an expected return of 12 percent with a standard deviation of 15 percent. Portfolio 2 has an expected return of 14 percent with a standard deviation of 16 percent. The SFRatios are 0.667 = (12 − 2)/15 and 0.75 = (14 − 2)/16 for Portfolios 1 and 2, respectively. For the superior Portfolio 2, the probability that portfolio return will be less than 2 percent is $N(-0.75) = 1 - N(0.75) = 1 - 0.7734 = 0.227$ or about 23 percent, assuming that portfolio returns are normally distributed.

22 Utility functions are mathematical representations of attitudes toward risk and return.

23 We shall discuss mean–variance analysis in detail in the readings on portfolio concepts.

24 A.D. Roy (1952) introduced this criterion.

25 If there is an asset offering a risk-free return over the time horizon being considered, and if R_L is less than or equal to that risk-free rate, then it is optimal to be fully invested in the risk-free asset. Holding the risk-free asset in this case eliminates the chance that the threshold return is not met.

You may have noticed the similarity of SFRatio to the Sharpe ratio. If we substitute the risk-free rate, R_F, for the critical level R_L, the SFRatio becomes the Sharpe ratio. The safety-first approach provides a new perspective on the Sharpe ratio: When we evaluate portfolios using the Sharpe ratio, the portfolio with the highest Sharpe ratio is the one that minimizes the probability that portfolio return will be less than the risk-free rate (given a normality assumption).

EXAMPLE 9

The Safety-First Optimal Portfolio for a Client

You are researching asset allocations for a client in Canada with a C$800,000 portfolio. Although her investment objective is long-term growth, at the end of a year she may want to liquidate C$30,000 of the portfolio to fund educational expenses. If that need arises, she would like to be able to take out the C$30,000 without invading the initial capital of C$800,000. Table 6 shows three alternative allocations.

Table 6 Mean and Standard Deviation for Three Allocations (in Percent)

	A	B	C
Expected annual return	25	11	14
Standard deviation of return	27	8	20

Address these questions (assume normality for Parts 2 and 3):

1 Given the client's desire not to invade the C$800,000 principal, what is the shortfall level, R_L? Use this shortfall level to answer Part 2.

2 According to the safety-first criterion, which of the three allocations is the best?

3 What is the probability that the return on the safety-first optimal portfolio will be less than the shortfall level?

Solution to 1:

Because C$30,000/C$800,000 is 3.75 percent, for any return less than 3.75 percent the client will need to invade principal if she takes out C$30,000. So R_L = 3.75 percent.

Solution to 2:

To decide which of the three allocations is safety-first optimal, select the alternative with the highest ratio $[E(R_P) - R_L]/\sigma_P$:

Allocation A: 0.787037 = (25 – 3.75)/27

Allocation B: 0.90625 = (11 – 3.75)/8

Allocation C: 0.5125 = (14 – 3.75)/20

Allocation B, with the largest ratio (0.90625), is the best alternative according to the safety-first criterion.

Solution to 3:

To answer this question, note that $P(R_B < 3.75) = N(-0.90625)$. We can round 0.90625 to 0.91 for use with tables of the standard normal cdf. First, we calculate $N(-0.91) = 1 - N(0.91) = 1 - 0.8186 = 0.1814$ or about 18.1 percent. Using a spreadsheet function for the standard normal cdf on −0.90625 without rounding, we get 18.24 percent or about 18.2 percent. The safety-first optimal portfolio has a roughly 18 percent chance of not meeting a 3.75 percent return threshold.

Several points are worth noting. First, if the inputs were even slightly different, we could get a different ranking. For example, if the mean return on B were 10 rather than 11 percent, A would be superior to B. Second, if meeting the 3.75 percent return threshold were a necessity rather than a wish, C$830,000 in one year could be modeled as a liability. Fixed income strategies such as cash flow matching could be used to offset or immunize the C$830,000 quasi-liability.

Roy's safety-first rule was the earliest approach to addressing shortfall risk. The standard mean-variance portfolio selection process can also accommodate a shortfall risk constraint.[26]

In many investment contexts besides Roy's safety-first criterion, we use the normal distribution to estimate a probability. For example, Kolb, Gay, and Hunter (1985) developed an expression based on the standard normal distribution for the probability that a futures trader will exhaust his liquidity because of losses in a futures contract. Another arena in which the normal distribution plays an important role is financial risk management. Financial institutions such as investment banks, security dealers, and commercial banks have formal systems to measure and control financial risk at various levels, from trading positions to the overall risk for the firm.[27] Two mainstays in managing financial risk are Value at Risk (VaR) and stress testing/scenario analysis. **Stress testing/scenario analysis**, a complement to VaR, refers to a set of techniques for estimating losses in extremely unfavorable combinations of events or scenarios. **Value at Risk** (VaR) is a money measure of the minimum value of losses expected over a specified time period (for example, a day, a quarter, or a year) at a given level of probability (often 0.05 or 0.01). Suppose we specify a one-day time horizon and a level of probability of 0.05, which would be called a 95 percent one-day VaR.[28] If this VaR equaled €5 million for a portfolio, there would be a 0.05 probability that the portfolio would lose €5 million or more in a single day (assuming our assumptions were correct). One of the basic approaches to estimating VaR, the variance-covariance or analytical method, assumes that returns follow a normal distribution. For more information on VaR, see Chance and Brooks (2016).

3.4 The Lognormal Distribution

Closely related to the normal distribution, the lognormal distribution is widely used for modeling the probability distribution of share and other asset prices. For example, the lognormal appears in the Black–Scholes–Merton option pricing model. The Black–Scholes–Merton model assumes that the price of the asset underlying the option is lognormally distributed.

26 See Leibowitz and Henriksson (1989), for example.

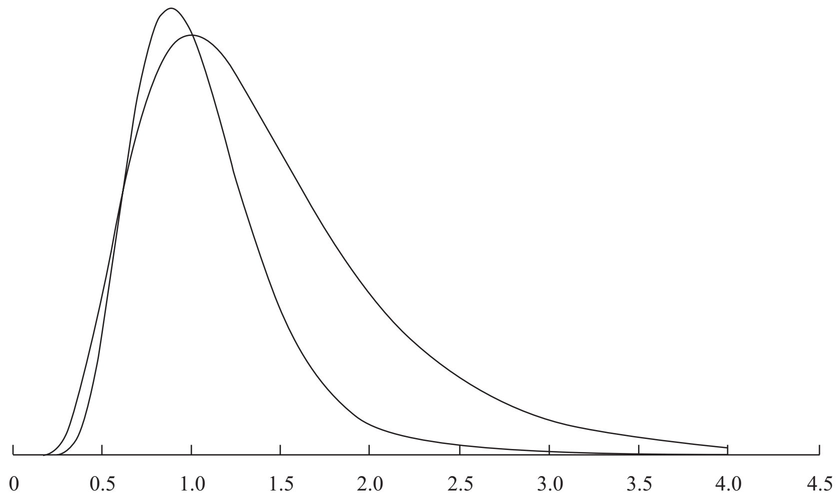
27 **Financial risk** is risk relating to asset prices and other financial variables. The contrast is to other, nonfinancial risks (for example, relating to operations and technology), which require different tools to manage.

28 In 95 percent one-day VaR, the 95 percent refers to the confidence in the value of VaR and is equal to 1 − 0.05; this is a traditional way to state VaR.

A random variable Y follows a lognormal distribution if its natural logarithm, ln Y, is normally distributed. The reverse is also true: If the natural logarithm of random variable Y, ln Y, is normally distributed, then Y follows a lognormal distribution. If you think of the term lognormal as "the log is normal," you will have no trouble remembering this relationship.

The two most noteworthy observations about the lognormal distribution are that it is bounded below by 0 and it is skewed to the right (it has a long right tail). Note these two properties in the graphs of the pdfs of two lognormal distributions in Figure 7. Asset prices are bounded from below by 0. In practice, the lognormal distribution has been found to be a usefully accurate description of the distribution of prices for many financial assets. On the other hand, the normal distribution is often a good approximation for returns. For this reason, both distributions are very important for finance professionals.

Figure 7 Two Lognormal Distributions



Like the normal distribution, the lognormal distribution is completely described by two parameters. Unlike the other distributions we have considered, a lognormal distribution is defined in terms of the parameters of a *different* distribution. The two parameters of a lognormal distribution are the mean and standard deviation (or variance) of its associated normal distribution: the mean and variance of ln Y, given that Y is lognormal. Remember, we must keep track of two sets of means and standard deviations (or variances): the mean and standard deviation (or variance) of the associated normal distribution (these are the parameters), and the mean and standard deviation (or variance) of the lognormal variable itself.

The expressions for the mean and variance of the lognormal variable itself are challenging. Suppose a normal random variable X has expected value μ and variance σ^2. Define $Y = \exp(X)$. Remember that the operation indicated by $\exp(X)$ or e^X is the opposite operation from taking logs.[29] Because $\ln Y = \ln [\exp(X)] = X$ is normal (we assume X is normal), Y is lognormal. What is the expected value of $Y = \exp(X)$? A guess might be that the expected value of Y is $\exp(\mu)$. The expected value is actually $\exp(\mu + 0.50\sigma^2)$, which is larger than $\exp(\mu)$ by a factor of $\exp(0.50\sigma^2) > 1$.[30] To get some

29 The quantity $e \approx 2.7182818$.

30 Note that $\exp(0.50\sigma^2) > 1$ because $\sigma^2 > 0$.

insight into this concept, think of what happens if we increase σ^2. The distribution spreads out; it can spread upward, but it cannot spread downward past 0. As a result, the center of its distribution is pushed to the right—the distribution's mean increases.[31]

The expressions for the mean and variance of a lognormal variable are summarized below, where μ and σ^2 are the mean and variance of the associated normal distribution (refer to these expressions as needed, rather than memorizing them):

- Mean (μ_L) of a lognormal random variable = $\exp(\mu + 0.50\sigma^2)$
- Variance ($\sigma_L{}^2$) of a lognormal random variable = $\exp(2\mu + \sigma^2) \times [\exp(\sigma^2) - 1]$

We now explore the relationship between the distribution of stock return and stock price. In the following we show that if a stock's continuously compounded return is normally distributed, then future stock price is necessarily lognormally distributed.[32] Furthermore, we show that stock price may be well described by the lognormal distribution even when continuously compounded returns do not follow a normal distribution. These results provide the theoretical foundation for using the lognormal distribution to model prices.

To outline the presentation that follows, we first show that the stock price at some future time T, S_T, equals the current stock price, S_0, multiplied by e raised to power $r_{0,T}$, the continuously compounded return from 0 to T; this relationship is expressed as $S_T = S_0\exp(r_{0,T})$. We then show that we can write $r_{0,T}$ as the sum of shorter-term continuously compounded returns and that if these shorter-period returns are normally distributed, then $r_{0,T}$ is normally distributed (given certain assumptions) or approximately normally distributed (not making those assumptions). As S_T is proportional to the log of a normal random variable, S_T is lognormal.

To supply a framework for our discussion, suppose we have a series of equally spaced observations on stock price: $S_0, S_1, S_2, \ldots, S_T$. Current stock price, S_0, is a known quantity and so is nonrandom. The future prices (such as S_1), however, are random variables. The **price relative**, S_1/S_0, is an ending price, S_1, over a beginning price, S_0; it is equal to 1 plus the holding period return on the stock from $t = 0$ to $t = 1$:

$$S_1/S_0 = 1 + R_{0,1}$$

For example, if S_0 = \$30 and S_1 = \$34.50, then S_1/S_0 = \$34.50/\$30 = 1.15. Therefore, $R_{0,1}$ = 0.15 or 15 percent. In general, price relatives have the form

$$S_{t+1}/S_t = 1 + R_{t,t+1}$$

where $R_{t,t+1}$ is the rate of return from t to $t + 1$.

An important concept is the continuously compounded return associated with a holding period return such as $R_{0,1}$. The **continuously compounded return** associated with a holding period is the natural logarithm of 1 plus that holding period return, or equivalently, the natural logarithm of the ending price over the beginning price (the price relative).[33] For example, if we observe a one-week holding period return of 0.04, the equivalent continuously compounded return, called the one-week continuously compounded return, is ln(1.04) = 0.039221; €1.00 invested for one week at 0.039221 continuously compounded gives €1.04, equivalent to a 4 percent one-week holding period return. The continuously compounded return from t to $t + 1$ is

$$r_{t,t+1} = \ln(S_{t+1}/S_t) = \ln(1 + R_{t,t+1}) \tag{5}$$

31 Luenberger (1998) is the source of this explanation.

32 Continuous compounding treats time as essentially continuous or unbroken, in contrast to discrete compounding, which treats time as advancing in discrete finite intervals. Continuously compounded returns are the model for returns in so-called **continuous time** finance models such as the Black–Scholes–Merton option pricing model. See the reading on the time value of money for more information on compounding.

33 In this reading we use lowercase r to refer specifically to continuously compounded returns.

For our example, $r_{0,1} = \ln(S_1/S_0) = \ln(1 + R_{0,1}) = \ln(\$34.50/\$30) = \ln(1.15) = 0.139762$. Thus, 13.98 percent is the continuously compounded return from $t = 0$ to $t = 1$. The continuously compounded return is smaller than the associated holding period return. If our investment horizon extends from $t = 0$ to $t = T$, then the continuously compounded return to T is

$$r_{0,T} = \ln(S_T/S_0)$$

Applying the function exp to both sides of the equation, we have $\exp(r_{0,T}) = \exp[\ln(S_T/S_0)] = S_T/S_0$, so

$$S_T = S_0\exp(r_{0,T})$$

We can also express S_T/S_0 as the product of price relatives:

$$S_T/S_0 = (S_T/S_{T-1})(S_{T-1}/S_{T-2})\ldots(S_1/S_0)$$

Taking logs of both sides of this equation, we find that continuously compounded return to time T is the sum of the one-period continuously compounded returns:

$$r_{0,T} = r_{T-1,T} + r_{T-2,T-1} + \ldots r_{0,1} \quad (6)$$

Using holding period returns to find the ending value of a \$1 investment involves the multiplication of quantities (1 + holding period return). Using continuously compounded returns involves addition.

A key assumption in many investment applications is that returns are **independently and identically distributed (IID)**. Independence captures the proposition that investors cannot predict future returns using past returns (i.e., weak-form market efficiency). Identical distribution captures the assumption of stationarity.[34]

Assume that the one-period continuously compounded returns (such as $r_{0,1}$) are IID random variables with mean μ and variance σ^2 (but making no normality or other distributional assumption). Then

$$E(r_{0,T}) = E(r_{T-1,T}) + E(r_{T-2,T-1}) + \ldots + E(r_{0,1}) = \mu T \quad (7)$$

(we add up μ for a total of T times) and

$$\sigma^2(r_{0,T}) = \sigma^2 T \quad (8)$$

(as a consequence of the independence assumption). The variance of the T holding period continuously compounded return is T multiplied by the variance of the one-period continuously compounded return; also, $\sigma(r_{0,T}) = \sigma\sqrt{T}$. If the one-period continuously compounded returns on the right-hand side of Equation 6 are normally distributed, then the T holding period continuously compounded return, $r_{0,T}$, is also normally distributed with mean μT and variance $\sigma^2 T$. This relationship is so because a linear combination of normal random variables is also normal. But even if the one-period continuously compounded returns are not normal, their sum, $r_{0,T}$, is approximately normal according to a result in statistics known as the central limit theorem.[35] Now compare $S_T = S_0\exp(r_{0,T})$ to $Y = \exp(X)$, where X is normal and Y is lognormal (as we discussed above). Clearly, we can model future stock price S_T as a lognormal random variable because $r_{0,T}$ should be at least approximately normal. This assumption of normally distributed returns is the basis in theory for the lognormal distribution as a model for the distribution of prices of shares and other assets.

---

34 Stationarity implies that the mean and variance of return do not change from period to period.

35 We mentioned the central limit theorem earlier in our discussion of the normal distribution. To give a somewhat fuller statement of it, according to the central limit theorem the sum (as well as the mean) of a set of independent, identically distributed random variables with finite variances is normally distributed, whatever distribution the random variables follow. We discuss the central limit theorem in the reading on sampling.

Continuously compounded returns play a role in many option pricing models, as mentioned earlier. An estimate of volatility is crucial for using option pricing models such as the Black–Scholes–Merton model. **Volatility** measures the standard deviation of the continuously compounded returns on the underlying asset.[36] In practice, we very often estimate volatility using a historical series of continuously compounded daily returns. We gather a set of daily holding period returns and then use Equation 5 to convert them into continuously compounded daily returns. We then compute the standard deviation of the continuously compounded daily returns and annualize that number using Equation 8.[37] (By convention, volatility is stated as an annualized measure.)[38] Example 10 illustrates the estimation of volatility for the shares of Astra International.

EXAMPLE 10

Volatility as Used in Option Pricing Models

Suppose you are researching Astra International (Indonesia Stock Exchange: ASII) and are interested in Astra's price action in a week in which international economic news had significantly affected the Indonesian stock market. You decide to use volatility as a measure of the variability of Astra shares during that week. Table 7 shows closing prices during that week.

Table 7 Astra International Daily Closing Prices

Day	Closing Price (IDR)
Monday	6,950
Tuesday	7,000
Wednesday	6,850
Thursday	6,600
Friday	6,350

Use the data in Table 7 to do the following:

1. Estimate the volatility of Astra shares. (Annualize volatility based on 250 days in a year.)
2. Identify the probability distribution for Astra share prices if continuously compounded daily returns follow the normal distribution.

36 Volatility is also called the instantaneous standard deviation, and as such is denoted σ. The underlying asset, or simply the underlying, is the asset underlying the option. For more information on these concepts, see Chance and Brooks (2016).

37 To compute the standard deviation of a set or sample of n returns, we sum the squared deviation of each return from the mean return and then divide that sum by $n - 1$. The result is the sample variance. Taking the square root of the sample variance gives the sample standard deviation. To review the calculation of standard deviation, see the reading on statistical concepts and market returns.

38 Annualizing is often done on the basis of 250 days in a year, the approximate number of days markets are open for trading. The 250-day number may lead to a better estimate of volatility than the 365-day number. Thus if daily volatility were 0.01, we would state volatility (on an annual basis) as $0.01\sqrt{250} = 0.1581$.

Solution to 1:

First, use Equation 5 to calculate the continuously compounded daily returns; then find their standard deviation in the usual way. (In the calculation of sample variance to get sample standard deviation, use a divisor of 1 less than the sample size.)

ln(7,000/6,950) = 0.007168

ln(6,850/7,000) = –0.021661

ln(6,600/6,850) = –0.037179

ln(6,350/6,600) = –0.038615

Sum = –0.090287

Mean = –0.022572

Variance = 0.000452

Standard Deviation = 0.021261

The standard deviation of continuously compounded daily returns is 0.021261. Equation 8 states that $\hat{\sigma}\left(r_{0,T}\right) = \hat{\sigma}\sqrt{T}$. In this example, $\hat{\sigma}$ is the sample standard deviation of one-period continuously compounded returns. Thus, $\hat{\sigma}$ refers to 0.021261. We want to annualize, so the horizon T corresponds to one year. As $\hat{\sigma}$ is in days, we set T equal to the number of trading days in a year (250).

We find that annualized volatility for Astra stock that week was 33.6 percent, calculated as $0.021261\sqrt{250} = 0.336165$.

Note that the sample mean, –0.022572, is a possible estimate of the mean, μ, of the continuously compounded one-period or daily returns. The sample mean can be translated into an estimate of the expected continuously compounded annual return using Equation 7: $\hat{\mu}T = -0.022572\left(250\right)$ (using 250 to be consistent with the calculation of volatility). But four observations are far too few to estimate expected returns. The variability in the daily returns overwhelms any information about expected return in a series this short.

Solution to 2:

Astra share prices should follow the lognormal distribution if the continuously compounded daily returns on Astra shares follow the normal distribution.

We have shown that the distribution of stock price is lognormal, given certain assumptions. What are the mean and variance of S_T if S_T follows the lognormal distribution? Earlier in this section, we gave bullet-point expressions for the mean and variance of a lognormal random variable. In the bullet-point expressions, the $\hat{\mu}$ and $\hat{\sigma}^2$ would refer, in the context of this discussion, to the mean and variance of the T horizon (not the one-period) continuously compounded returns (assumed to follow a normal distribution), compatible with the horizon of S_T.[39] Related to the use of mean and variance (or standard deviation), earlier in this reading we used those quantities to construct intervals in which we expect to find a certain percentage of the observations of a normally distributed random variable. Those intervals were symmetric about the mean. Can we state similar, symmetric intervals for a lognormal random variable?

39 The expression for the mean is $E(S_T) = S_0 \exp[E(r_{0,T}) + 0.5\sigma^2(r_{0,T})]$, for example.

Unfortunately, we cannot. Because the lognormal distribution is not symmetric, such intervals are more complicated than for the normal distribution, and we will not discuss this specialist topic here.[40]

Finally, we have presented the relation between the mean and variance of continuously compounded returns associated with different time horizons (see Equations 7 and 8), but how are the means and variances of holding period returns and continuously compounded returns related? As analysts, we typically think in terms of holding period returns rather than continuously compounded returns, and we may desire to convert means and standard deviations of holding period returns to means and standard deviations of continuously compounded returns for an option application, for example. To effect such conversions (and those in the other direction, from a continuous compounding to a holding period basis), we can use the expressions in Ferguson (1993).

MONTE CARLO SIMULATION

4

With an understanding of probability distributions, we are now prepared to learn about a computer-based technique in which probability distributions play an integral role. The technique is called Monte Carlo simulation. **Monte Carlo simulation** in finance involves the use of a computer to represent the operation of a complex financial system. A characteristic feature of Monte Carlo simulation is the generation of a large number of random samples from a specified probability distribution or distributions to represent the role of risk in the system.

Monte Carlo simulation has several quite distinct uses. One use is in planning. Stanford University researcher Sam Savage provided the following neat picture of that role: "What is the last thing you do before you climb on a ladder? You shake it, and that is Monte Carlo simulation."[41] Just as shaking a ladder helps us assess the risks in climbing it, Monte Carlo simulation allows us to experiment with a proposed policy before actually implementing it. For example, investment performance can be evaluated with reference to a benchmark or a liability. Defined benefit pension plans often invest assets with reference to plan liabilities. Pension liabilities are a complex random process. In a Monte Carlo asset-liability financial planning study, the functioning of pension assets and liabilities is simulated over time, given assumptions about how assets are invested, the work force, and other variables. A key specification in this and all Monte Carlo simulations is the probability distributions of the various sources of risk (including interest rates and security market returns, in this case). The implications of different investment policy decisions on the plan's funded status can be assessed through simulated time. The experiment can be repeated for another set of assumptions. We can view Example 11 below as coming under this heading. In that example, market return series are not long enough to address researchers' questions on stock market timing, so the researchers simulate market returns to find answers to their questions.

Monte Carlo simulation is also widely used to develop estimates of VaR. In this application, we simulate the portfolio's profit and loss performance for a specified time horizon. Repeated trials within the simulation (each trial involving a draw of random observations from a probability distribution) produce a frequency distribution for changes in portfolio value. The point that defines the cutoff for the least favorable 5 percent of simulated changes is an estimate of 95 percent VaR, for example.

40 See Hull (2017) for a discussion of lognormal confidence intervals.
41 *Business Week*, 22 January 2001.

In an extremely important use, Monte Carlo simulation is a tool for valuing complex securities, particularly some European-style options for which no analytic pricing formula is available.[42] For other securities, such as mortgage-backed securities with complex embedded options, Monte Carlo simulation is also an important modeling resource.

Researchers use Monte Carlo simulation to test their models and tools. How critical is a particular assumption to the performance of a model? Because we control the assumptions when we do a simulation, we can run the model through a Monte Carlo simulation to examine a model's sensitivity to a change in our assumptions.

To understand the technique of Monte Carlo simulation, let us present the process as a series of steps.[43] To illustrate the steps, we take the case of using Monte Carlo simulation to value a type of option for which no analytic pricing formula is available, an Asian call option on a stock. An **Asian call option** is a European-style option with a value at maturity equal to the difference between the stock price at maturity and the average stock price during the life of the option, or \$0, whichever is greater. For instance, if the final stock price is \$34 with an average value of \$31 over the life of the option, the value of the option at maturity is \$3 (the greater of \$34 – \$31 = \$3 and \$0). Steps 1 through 3 of the process describe specifying the simulation; Steps 4 through 7 describe running the simulation.

1 Specify the quantities of interest (option value, for example, or the funded status of a pension plan) in terms of underlying variables. The underlying variable or variables could be stock price for an equity option, the market value of pension assets, or other variables relating to the pension benefit obligation for a pension plan. Specify the starting values of the underlying variables.

   To illustrate the steps, we are using the case of valuing an Asian call option on stock. We use C_{iT} to represent the value of the option at maturity T. The subscript i in C_{iT} indicates that C_{iT} is a value resulting from the ith **simulation trial**, each simulation trial involving a drawing of random values (an iteration of Step 4).

2 Specify a time grid. Take the horizon in terms of calendar time and split it into a number of subperiods, say K in total. Calendar time divided by the number of subperiods, K, is the time increment, Δt.

3 Specify distributional assumptions for the risk factors that drive the underlying variables. For example, stock price is the underlying variable for the Asian call, so we need a model for stock price movement. Say we choose the following model for changes in stock price, where Z_k stands for the standard normal random variable:

$$\Delta(\text{Stock price}) = (\mu \times \text{Prior stock price} \times \Delta t) + (\sigma \times \text{Prior stock price} \times Z_k)$$

   In the way that we are using the term, Z_k is a risk factor in the simulation. Through our choice of μ and σ, we control the distribution of stock price. Although this example has one risk factor, a given simulation may have multiple risk factors.

4 Using a computer program or spreadsheet function, draw K random values of each risk factor. In our example, the spreadsheet function would produce a draw of K values of the standard normal variable Z_k: $Z_1, Z_2, Z_3, \ldots, Z_K$.

42 A **European-style** option or **European option** is an option exercisable only at maturity.

43 The steps should be viewed as providing an overview of Monte Carlo simulation rather than as a detailed recipe for implementing a Monte Carlo simulation in its many varied applications.

5 Calculate the underlying variables using the random observations generated in Step 4. Using the above model of stock price dynamics, the result is K observations on changes in stock price. An additional calculation is needed to convert those changes into K stock prices (using initial stock price, which is given). Another calculation produces the average stock price during the life of the option (the sum of K stock prices divided by K).

6 Compute the quantities of interest. In our example, the first calculation is the value of an Asian call at maturity, C_{iT}. A second calculation discounts this terminal value back to the present to get the call value as of today, C_{i0}. We have completed one simulation trial. (The subscript i in C_{i0} stands for the ith simulation trial, as it does in C_{iT}.) In a Monte Carlo simulation, a running tabulation is kept of statistics relating to the distribution of the quantities of interest, including their mean value and standard deviation, over the simulation trials to that point.

7 Iteratively go back to Step 4 until a specified number of trials, I, is completed. Finally, produce statistics for the simulation. The key value for our example is the mean value of C_{i0} for the total number of simulation trials. This mean value is the Monte Carlo estimate of the value of the Asian call.

How many simulation trials should be specified? In general, we need to increase the number of trials by a factor of 100 to get each extra digit of accuracy. Depending on the problem, tens of thousands of trials may be needed to obtain accuracy to two decimal places (as required for option value, for example). Conducting a large number of trials is not necessarily a problem, given today's computing power. The number of trials needed can be reduced using variance reduction procedures, a topic outside the scope of this reading.[44]

In Step 4 of our example, a computer function produced a set of random observations on a standard normal random variable. Recall that for a uniform distribution, all possible numbers are equally likely. The term **random number generator** refers to an algorithm that produces uniformly distributed random numbers between 0 and 1. In the context of computer simulations, the term **random number** refers to an observation drawn from a uniform distribution.[45] For other distributions, the term "random observation" is used in this context.

It is a remarkable fact that random observations from any distribution can be produced using the uniform random variable with endpoints 0 and 1. To see why this is so, consider the inverse transformation method of producing random observations. Suppose we are interested in obtaining random observations for a random variable, X, with cumulative distribution function $F(x)$. Recall that $F(x)$ evaluated at x is a number between 0 and 1. Suppose a random outcome of this random variable is 3.21 and that $F(3.21) = 0.25$ or 25 percent. Define an inverse of F, call it F^{-1}, that can do the following: Substitute the probability 0.25 into F^{-1} and it returns the random outcome 3.21. In other words, $F^{-1}(0.25) = 3.21$. To generate random observations on X, the steps are 1) generate a uniform random number, r, between 0 and 1 using the random number generator and 2) evaluate $F^{-1}(r)$ to obtain a random observation on X. Random observation generation is a field of study in itself, and we have briefly discussed the inverse transformation method here just to illustrate a point. As a generalist

44 For details on this and other technical aspects of Monte Carlo simulation, see Hillier (2014).

45 The numbers that random number generators produce depend on a seed or initial value. If the same seed is fed to the same generator, it will produce the same sequence. All sequences eventually repeat. Because of this predictability, the technically correct name for the numbers produced by random number generators is **pseudo-random numbers**. Pseudo-random numbers have sufficient qualities of randomness for most practical purposes.

you do not need to address the technical details of converting random numbers into random observations, but you do need to know that random observations from any distribution can be generated using a uniform random variable.

In Examples 11 and 12, we give an application of Monte Carlo simulation to a question of great interest to investment practice: the potential gains from market timing.

EXAMPLE 11

Potential Gains from Market Timing: A Monte Carlo Simulation (1)

All active investors want to achieve superior performance. One possible source of superior performance is market timing ability. How accurate does an investor need to be as a bull- and bear-market forecaster for market timing to be profitable? What size gains compared with a buy-and-hold strategy accrue to a given level of accuracy? Because of the variability in asset returns, a huge amount of return data is needed to find statistically reliable answers to these questions. Chua, Woodward, and To (1987) thus selected Monte Carlo simulation to address the potential gains from market timing. They were interested in the perspective of a Canadian investor.

To understand their study, suppose that at the beginning of a year, an investor predicts that the next year will see either a bull market or bear market. If the prediction is *bull market,* the investor puts all her money in stocks and earns the market return for that year. On the other hand, if the prediction is *bear market,* the investor holds T-bills and earns the T-bill return. After the fact, a market is categorized as *bull market* if the stock market return, R_{Mt}, minus T-bill return, R_{Ft}, is positive for the year; otherwise, the market is classed as *bear market.* The investment results of a market timer can be compared with those of a buy-and-hold investor. A buy-and-hold investor earns the market return every year. For Chua et al., one quantity of interest was the gain from market timing. They defined this quantity as the market timer's average return minus the average return to a buy-and-hold investor.

To simulate market returns, Chua et al. generated 10,000 random standard normal observations, Z_t. At the time of the study, Canadian stocks had a historical mean annual return of 12.95 percent with a standard deviation of 18.30 percent. To reflect these parameters, the simulated market returns are $R_{Mt} = 0.1830Z_t + 0.1295$, $t = 1, 2, \ldots, 10{,}000$. Using a second set of 10,000 random standard normal observations, historical return parameters for Canadian T-bills, as well as the historical correlation of T-bill and stock returns, the authors generated 10,000 T-bill returns.

An investor can have different skills in forecasting bull and bear markets. Chua et al. characterized market timers by accuracy in forecasting bull markets and accuracy in forecasting bear markets. For example, bull market forecasting accuracy of 50 percent means that when the timer forecasts *bull market* for the next year, she is right just half the time, indicating no skill. Suppose an investor has 60 percent accuracy in forecasting *bull market* and 80 percent accuracy in forecasting *bear market* (a 60–80 timer). We can simulate how an investor would fare. After generating the first observation on $R_{Mt} - R_{Ft}$, we know whether that observation is a bull or bear market. If the observation is *bull market,* then 0.60 (forecast accuracy for bull markets) is compared with a random number (between 0 and 1). If the random number is less than 0.60, which occurs with a 60 percent probability, then the market timer is assumed to have correctly predicted *bull market* and her return for that first observation is the market return. If the random number is greater than 0.60, then the market timer is assumed to have made an error and predicted *bear market;* her return for that observation is the risk-free rate. In a similar fashion, if that first observation is *bear market,* the timer has an 80 percent chance of being right in forecasting *bear market* based on a random number draw. In either case, her return is compared with the market return to record her gain versus a buy-and-hold strategy. That process is one simulation trial. The simulated mean return earned by the timer is the average return earned by the timer over all trials in the simulation.

To increase our understanding of the process, consider a hypothetical Monte Carlo simulation with four trials for the 60–80 timer (who, to reiterate, has 60 percent accuracy in forecasting bull markets and 80 percent accuracy in forecasting bear markets). Table 8 gives data for the simulation. Let us look at Trials 1 and 2. In Trial 1, the first random number drawn leads to a market return of 0.121. Because the market return, 0.121, exceeded the T-bill return, 0.050, we have a bull market. We generate a random number, 0.531, which we then compare with the timer's bull market accuracy, 0.60. Because 0.531 is less than 0.60, the timer is assumed to have made a correct bull market forecast and thus to have invested in stocks. Thus the timer earns the stock market return, 0.121, for that trial. In the second trial we observe another bull market, but because the random number 0.725 is greater than 0.60, the timer is assumed to have made an error and predicted a bear market; therefore, the timer earned the T-bill return, 0.081, rather than higher stock market return.

Table 8 Hypothetical Simulation for a 60–80 Market Timer

	After Draws for Z_t and for the T-bill Return			Simulation Results		
Trial	R_{Mt}	R_{Ft}	**Bull or Bear Market?**	**Value of *X***	**Timer's Prediction Correct?**	**Return Earned by Timer**
1	0.121	0.050	Bull	0.531	Yes	0.121
2	0.092	0.081	Bull	0.725	No	0.081
3	–0.020	0.034	Bear	0.786	Yes	0.034
4	0.052	0.055	***A***	0.901	***B***	***C***
						$\bar{R}$ = ***D***

Note: $\bar{R}$ is the mean return earned by the timer over the four simulation trials.

Using the data in Table 8, determine the values of *A*, *B*, *C*, and *D*.

Solution:

The value of *A* is *Bear* because the stock market return was less than the T-bill return in Trial 4. The value of *B* is *No.* Because we observe a bear market, we compare the random number 0.901 with 0.80, the timer's bear-market forecasting accuracy. Because 0.901 is greater than 0.8, the timer is assumed to have made an error. The value of *C* is 0.052, the return on the stock market, because the timer made an error and invested in the stock market and earned 0.052 rather than the higher T-bill return of 0.055. The value of *D* is $\bar{R}$ = (0.121 + 0.081 + 0.034 + 0.052) = 0.288/4 = 0.072. Note that we could calculate other statistics besides the mean, such as the standard deviation of the returns earned by the timer over the four trials in the simulation.

EXAMPLE 12

Potential Gains from Market Timing: A Monte Carlo Simulation (2)

Having discussed the plan of the Chua et al. study and illustrated the method for a hypothetical Monte Carlo simulation with four trials, we conclude our presentation of the study.

The hypothetical simulation in Example 11 had four trials, far too few to reach statistically precise conclusions. The simulation of Chua et al. incorporated 10,000 trials. Chua et al. specified bull- and bear-market prediction skill levels of 50, 60, 70, 80, 90, and 100 percent. Table 9 presents a very small excerpt from their simulation results for the no transaction costs case (transaction costs were also examined). Reading across the row, the timer with 60 percent bull market and 80 percent bear market forecasting accuracy had a mean annual gain from market timing of −1.12 percent per year. On average, the buy-and-hold investor out-earned this skillful timer by 1.12 percentage points. There was substantial variability in gains across the simulation trials, however: The standard deviation of the gain was 14.77 percent, so in many trials (but not on average) the gain was positive. Row 3 (win/loss) is the ratio of profitable switches between stocks and T-bills to unprofitable switches. This ratio was a favorable 1.2070 for the 60–80 timer. (When transaction costs were considered, however, fewer switches are profitable: The win–loss ratio was 0.5832 for the 60–80 timer.)

Table 9 Gains from Stock Market Timing (No Transaction Costs)

Bull Market Accuracy (%)		Bear Market Accuracy (%) 50	60	70	80	90	100
60	Mean (%)	−2.50	−1.99	−1.57	−1.12	−0.68	−0.22
	S.D. (%)	13.65	14.11	14.45	14.77	15.08	15.42
	Win/Loss	0.7418	0.9062	1.0503	1.2070	1.3496	1.4986

Source: Chua, Woodward, and To (1987), Table II (excerpt).

The authors concluded that the cost of not being invested in the market during bull market years is high. Because a buy-and-hold investor never misses a bull market year, she has 100 percent forecast accuracy for bull markets (at the cost of 0 percent accuracy for bear markets). Given their definitions and assumptions, the authors also concluded that successful market timing requires a minimum accuracy of 80 percent in forecasting both bull and bear markets. Market timing is a continuing area of interest and study, and other perspectives exist. However, this example illustrates how Monte Carlo simulation is used to address important investment issues.

The analyst chooses the probability distributions in Monte Carlo simulation. By contrast, **historical simulation** samples from a historical record of returns (or other underlying variables) to simulate a process. The concept underlying historical simulation (also called **back simulation**) is that the historical record provides the most direct evidence on distributions (and that the past applies to the future). For example, refer to Step 2 in the outline of Monte Carlo simulation above and suppose the time increment is one day. Further, suppose we base the simulation on the record of daily stock returns over the last five years. In one type of historical simulation, we randomly draw K returns from that record to generate one simulation trial. We put back the observations into the sample, and in the next trial we again randomly sample with replacement. The simulation results directly reflect frequencies in the data. A drawback of this approach is that any risk not represented in the time period selected (for example, a stock market crash) will not be reflected in the simulation. Compared with Monte Carlo simulation, historical simulation does not lend itself to "what if" analyses. Nevertheless, historic simulation is an established alternative simulation methodology.

Monte Carlo simulation is a complement to analytical methods. It provides only statistical estimates, not exact results. Analytical methods, where available, provide more insight into cause-and-effect relationships. For example, the Black–Scholes–Merton option pricing model for the value of a European call option is an analytical method, expressed as a formula. It is a much more efficient method for valuing such a call than is Monte Carlo simulation. As an analytical expression, the Black–Scholes–Merton model permits the analyst to quickly gauge the sensitivity of call value to changes in current stock price and the other variables that determine call value. In contrast, Monte Carlo simulations do not directly provide such precise insights. However, only some types of options can be priced with analytical expressions. As financial product innovations proceed, the field of applications for Monte Carlo simulation continues to grow.

SUMMARY

In this reading, we have presented the most frequently used probability distributions in investment analysis and the Monte Carlo simulation.

- A probability distribution specifies the probabilities of the possible outcomes of a random variable.
- The two basic types of random variables are discrete random variables and continuous random variables. Discrete random variables take on at most a countable number of possible outcomes that we can list as x_1, x_2, … In contrast, we cannot describe the possible outcomes of a continuous random variable Z with a list z_1, z_2, … because the outcome $(z_1 + z_2)/2$, not in the list, would always be possible.
- The probability function specifies the probability that the random variable will take on a specific value. The probability function is denoted $p(x)$ for a discrete random variable and $f(x)$ for a continuous random variable. For any probability function $p(x)$, $0 \leq p(x) \leq 1$, and the sum of $p(x)$ over all values of X equals 1.
- The cumulative distribution function, denoted $F(x)$ for both continuous and discrete random variables, gives the probability that the random variable is less than or equal to x.
- The discrete uniform and the continuous uniform distributions are the distributions of equally likely outcomes.
- The binomial random variable is defined as the number of successes in n Bernoulli trials, where the probability of success, p, is constant for all trials and the trials are independent. A Bernoulli trial is an experiment with two outcomes, which can represent success or failure, an up move or a down move, or another binary (two-fold) outcome.

- A binomial random variable has an expected value or mean equal to np and variance equal to $np(1 - p)$.
- A binomial tree is the graphical representation of a model of asset price dynamics in which, at each period, the asset moves up with probability p or down with probability $(1 - p)$. The binomial tree is a flexible method for modeling asset price movement and is widely used in pricing options.
- The normal distribution is a continuous symmetric probability distribution that is completely described by two parameters: its mean, μ, and its variance, σ^2.
- A univariate distribution specifies the probabilities for a single random variable. A multivariate distribution specifies the probabilities for a group of related random variables.
- To specify the normal distribution for a portfolio when its component securities are normally distributed, we need the means, standard deviations, and all the distinct pairwise correlations of the securities. When we have those statistics, we have also specified a multivariate normal distribution for the securities.
- For a normal random variable, approximately 68 percent of all possible outcomes are within a one standard deviation interval about the mean, approximately 95 percent are within a two standard deviation interval about the mean, and approximately 99 percent are within a three standard deviation interval about the mean.
- A normal random variable, X, is standardized using the expression $Z = (X - \mu)/\sigma$, where μ and σ are the mean and standard deviation of X. Generally, we use the sample mean $\bar{X}$ as an estimate of μ and the sample standard deviation s as an estimate of σ in this expression.
- The standard normal random variable, denoted Z, has a mean equal to 0 and variance equal to 1. All questions about any normal random variable can be answered by referring to the cumulative distribution function of a standard normal random variable, denoted $N(x)$ or $N(z)$.
- Shortfall risk is the risk that portfolio value will fall below some minimum acceptable level over some time horizon.
- Roy's safety-first criterion, addressing shortfall risk, asserts that the optimal portfolio is the one that minimizes the probability that portfolio return falls below a threshold level. According to Roy's safety-first criterion, if returns are normally distributed, the safety-first optimal portfolio P is the one that maximizes the quantity $[E(R_P) - R_L]/\sigma_P$, where R_L is the minimum acceptable level of return.
- A random variable follows a lognormal distribution if the natural logarithm of the random variable is normally distributed. The lognormal distribution is defined in terms of the mean and variance of its associated normal distribution. The lognormal distribution is bounded below by 0 and skewed to the right (it has a long right tail).
- The lognormal distribution is frequently used to model the probability distribution of asset prices because it is bounded below by zero.
- Continuous compounding views time as essentially continuous or unbroken; discrete compounding views time as advancing in discrete finite intervals.
- The continuously compounded return associated with a holding period is the natural log of 1 plus the holding period return, or equivalently, the natural log of ending price over beginning price.

- If continuously compounded returns are normally distributed, asset prices are lognormally distributed. This relationship is used to move back and forth between the distributions for return and price. Because of the central limit theorem, continuously compounded returns need not be normally distributed for asset prices to be reasonably well described by a lognormal distribution.
- Monte Carlo simulation involves the use of a computer to represent the operation of a complex financial system. A characteristic feature of Monte Carlo simulation is the generation of a large number of random samples from specified probability distribution(s) to represent the operation of risk in the system. Monte Carlo simulation is used in planning, in financial risk management, and in valuing complex securities. Monte Carlo simulation is a complement to analytical methods but provides only statistical estimates, not exact results.
- Historical simulation is an established alternative to Monte Carlo simulation that in one implementation involves repeated sampling from a historical data series. Historical simulation is grounded in actual data but can reflect only risks represented in the sample historical data. Compared with Monte Carlo simulation, historical simulation does not lend itself to "what if" analyses.

REFERENCES

Campbell, John, Andrew Lo, and A. Craig MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

Chance, Don M., and Robert Brooks. 2016. *An Introduction to Derivatives and Risk Management*, 10th ed. Mason, OH: South-Western.

Chua, Jess H., Richard S. Woodward, and Eric C. To. 1987. "Potential Gains from Stock Market Timing in Canada." *Financial Analysts Journal*, vol. 43, no. 5:50–56.

Cox, Jonathan, Stephen Ross, and Mark Rubinstein. 1979. "Options Pricing: A Simplified Approach." *Journal of Financial Economics*, vol. 7:229–263.

Fama, Eugene. 1976. *Foundations of Finance*. New York: Basic Books.

Ferguson, Robert. 1993. "Some Formulas for Evaluating Two Popular Option Strategies." *Financial Analysts Journal*, vol. 49, no. 5:71–76.

Hillier, Frederick S. 2014. *Introduction to Operations Research*, 10th edition. New York: McGraw-Hill.

Hull, John. 2017. *Options, Futures, and Other Derivatives*, 10th edition. Upper Saddle River, NJ: Pearson.

Kolb, Robert W., Gerald D. Gay, and William C. Hunter. 1985. "Liquidity Requirements for Financial Futures Investments." *Financial Analysts Journal*, vol. 41, no. 3:60–68.

Kon, Stanley J. 1984. "Models of Stock Returns—A Comparison." *Journal of Finance*, vol. 39:147–165.

Leibowitz, Martin, and Roy Henriksson. 1989. "Portfolio Optimization with Shortfall Constraints: A Confidence-Limit Approach to Managing Downside Risk." *Financial Analysts Journal*, vol. 45, no. 2:34–41.

Luenberger, David G. 1998. *Investment Science*. New York: Oxford University Press.

Roy, A.D. 1952. "Safety-First and the Holding of Assets." *Econometrica*, vol. 20:431–439.

PRACTICE PROBLEMS

1 A European put option on stock conveys the right to sell the stock at a prespecified price, called the exercise price, at the maturity date of the option. The value of this put at maturity is (exercise price – stock price) or \$0, whichever is greater. Suppose the exercise price is \$100 and the underlying stock trades in ticks of \$0.01. At any time before maturity, the terminal value of the put is a random variable.

  A Describe the distinct possible outcomes for terminal put value. (Think of the put's maximum and minimum values and its minimum price increments.)

  B Is terminal put value, at a time before maturity, a discrete or continuous random variable?

  C Letting Y stand for terminal put value, express in standard notation the probability that terminal put value is less than or equal to \$24. No calculations or formulas are necessary.

2 Define the term "binomial random variable." Describe the types of problems for which the binomial distribution is used.

3 The value of the cumulative distribution function $F(x)$, where x is a particular outcome, for a discrete uniform distribution:

  A sums to 1.

  B lies between 0 and 1.

  C decreases as x increases.

4 For a binomial random variable with five trials, and a probability of success on each trial of 0.50, the distribution will be:

  A skewed.

  B uniform.

  C symmetric.

5 In a discrete uniform distribution with 20 potential outcomes of integers 1 to 20, the probability that X is greater than or equal to 3 but less than 6, $P(3 \le X < 6)$, is:

  A 0.10.

  B 0.15.

  C 0.20.

6 Over the last 10 years, a company's annual earnings increased year over year seven times and decreased year over year three times. You decide to model the number of earnings increases for the next decade as a binomial random variable.

  A What is your estimate of the probability of success, defined as an increase in annual earnings?

  For Parts B, C, and D of this problem, assume the estimated probability is the actual probability for the next decade.

  B What is the probability that earnings will increase in exactly 5 of the next 10 years?

  C Calculate the expected number of yearly earnings increases during the next 10 years.

D Calculate the variance and standard deviation of the number of yearly earnings increases during the next 10 years.

E The expression for the probability function of a binomial random variable depends on two major assumptions. In the context of this problem, what must you assume about annual earnings increases to apply the binomial distribution in Part B? What reservations might you have about the validity of these assumptions?

7 A portfolio manager annually outperforms her benchmark 60% of the time. Assuming independent annual trials, what is the probability that she will outperform her benchmark four or more times over the next five years?

A 0.26

B 0.34

C 0.48

8 You are examining the record of an investment newsletter writer who claims a 70 percent success rate in making investment recommendations that are profitable over a one-year time horizon. You have the one-year record of the newsletter's seven most recent recommendations. Four of those recommendations were profitable. If all the recommendations are independent and the newsletter writer's skill is as claimed, what is the probability of observing four or fewer profitable recommendations out of seven in total?

9 You are forecasting sales for a company in the fourth quarter of its fiscal year. Your low-end estimate of sales is €14 million, and your high-end estimate is €15 million. You decide to treat all outcomes for sales between these two values as equally likely, using a continuous uniform distribution.

A What is the expected value of sales for the fourth quarter?

B What is the probability that fourth-quarter sales will be less than or equal to €14,125,000?

10 State the approximate probability that a normal random variable will fall within the following intervals:

A Mean plus or minus one standard deviation.

B Mean plus or minus two standard deviations.

C Mean plus or minus three standard deviations.

11 Find the area under the normal curve up to $z = 0.36$; that is, find $P(Z \leq 0.36)$. Interpret this value.

12 If the probability that a portfolio outperforms its benchmark in any quarter is 0.75, the probability that the portfolio outperforms its benchmark in three or fewer quarters over the course of a year is *closest* to:

A 0.26

B 0.42

C 0.68

13 In futures markets, profits or losses on contracts are settled at the end of each trading day. This procedure is called marking to market or daily resettlement. By preventing a trader's losses from accumulating over many days, marking to market reduces the risk that traders will default on their obligations. A futures markets trader needs a liquidity pool to meet the daily mark to market. If liquidity is exhausted, the trader may be forced to unwind his position at an unfavorable time.

Suppose you are using financial futures contracts to hedge a risk in your portfolio. You have a liquidity pool (cash and cash equivalents) of λ dollars per contract and a time horizon of T trading days. For a given size liquidity pool, λ, Kolb, Gay, and Hunter (1985) developed an expression for the probability stating that you will exhaust your liquidity pool within a T-day horizon as a result of the daily mark to market. Kolb et al. assumed that the expected change in futures price is 0 and that futures price changes are normally distributed. With σ representing the standard deviation of daily futures price changes, the standard deviation of price changes over a time horizon to day T is $\sigma\sqrt{T}$, given continuous compounding. With that background, the Kolb et al. expression is

Probability of exhausting liquidity pool = $2[1 - N(x)]$

where $x = \lambda \big/ \left(\sigma\sqrt{T}\right)$. Here x is a standardized value of λ. $N(x)$ is the standard normal cumulative distribution function. For some intuition about $1 - N(x)$ in the expression, note that the liquidity pool is exhausted if losses exceed the size of the liquidity pool at any time up to and including T; the probability of that event happening can be shown to be proportional to an area in the right tail of a standard normal distribution, $1 - N(x)$.

Using the Kolb et al. expression, answer the following questions:

A Your hedging horizon is five days, and your liquidity pool is $2,000 per contract. You estimate that the standard deviation of daily price changes for the contract is $450. What is the probability that you will exhaust your liquidity pool in the five-day period?

B Suppose your hedging horizon is 20 days, but all the other facts given in Part A remain the same. What is the probability that you will exhaust your liquidity pool in the 20-day period?

14 Which of the following is characteristic of the normal distribution?

A Asymmetry

B Kurtosis of 3

C Definitive limits or boundaries

15 Which of the following assets *most likely* requires the use of a multivariate distribution for modeling returns?

A A call option on a bond

B A portfolio of technology stocks

C A stock in a market index

16 The total number of parameters that fully characterizes a multivariate normal distribution for the returns on two stocks is:

A 3.

B 4.

C 5.

17 A client has a portfolio of common stocks and fixed-income instruments with a current value of £1,350,000. She intends to liquidate £50,000 from the portfolio at the end of the year to purchase a partnership share in a business. Furthermore, the client would like to be able to withdraw the £50,000 without reducing the initial capital of £1,350,000. The following table shows four alternative asset allocations.

Mean and Standard Deviation for Four Allocations (in Percent)

	A	B	C	D
Expected annual return	16	12	10	9
Standard deviation of return	24	17	12	11

Address the following questions (assume normality for Parts B and C):

A Given the client's desire not to invade the £1,350,000 principal, what is the shortfall level, R_L? Use this shortfall level to answer Part B.

B According to the safety-first criterion, which of the allocations is the best?

C What is the probability that the return on the safety-first optimal portfolio will be less than the shortfall level, R_L?

Please refer to Exhibit 1 for Questions 18 and 19

Exhibit 1 ***Z*-Table Values, $P(Z \le z) = N(z)$ for $z \ge 0$**

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224

18 A portfolio has an expected mean return of 8 percent and standard deviation of 14 percent. The probability that its return falls between 8 and 11 percent is *closest* to:

A 8.3%

B 14.8%.

C 58.3%.

19 A portfolio has an expected return of 7% with a standard deviation of 13%. For an investor with a minimum annual return target of 4%, the probability that the portfolio return will fail to meet the target is *closest* to:

A 33%.

B 41%.

C 59%.

20 **A** Define Monte Carlo simulation and explain its use in finance.

B Compared with analytical methods, what are the strengths and weaknesses of Monte Carlo simulation for use in valuing securities?

21 A standard lookback call option on stock has a value at maturity equal to (Value of the stock at maturity – Minimum value of stock during the life of the option prior to maturity) or \$0, whichever is greater. If the minimum value reached prior to maturity was \$20.11 and the value of the stock at maturity is \$23, for example, the call is worth \$23 – \$20.11 = \$2.89. Briefly discuss how you might use Monte Carlo simulation in valuing a lookback call option.

22 Which of the following is a continuous random variable?

A The value of a futures contract quoted in increments of \$0.05

B The total number of heads recorded in 1 million tosses of a coin

C The rate of return on a diversified portfolio of stocks over a three-month period

23 X is a discrete random variable with possible outcomes $X = \{1,2,3,4\}$. Three functions $f(x)$, $g(x)$, and $h(x)$ are proposed to describe the probabilities of the outcomes in X.

	Probability Function		
$X = x$	**$f(x) = P(X = x)$**	**$g(x) = P(X = x)$**	**$h(x) = P(X = x)$**
1	–0.25	0.20	0.20
2	0.25	0.25	0.25
3	0.50	0.50	0.30
4	0.25	0.05	0.35

The conditions for a probability function are satisfied by:

A $f(x)$.

B $g(x)$.

C $h(x)$.

24 The cumulative distribution function for a discrete random variable is shown in the following table.

$X = x$	**Cumulative Distribution Function $F(x) = P(X \leq x)$**
1	0.15
2	0.25
3	0.50
4	0.60
5	0.95
6	1.00

The probability that X will take on a value of either 2 or 4 is *closest* to:

A 0.20.

B 0.35.

C 0.85.

25 Which of the following events can be represented as a Bernoulli trial?

A The flip of a coin

B The closing price of a stock

C The picking of a random integer between 1 and 10

26 The weekly closing prices of Mordice Corporation shares are as follows:

Date	Closing Price (€)
1 August	112
8 August	160
15 August	120

The continuously compounded return of Mordice Corporation shares for the period August 1 to August 15 is *closest to*:

A 6.90%

B 7.14%

C 8.95%

27 A stock is priced at $100.00 and follows a one-period binomial process with an up move that equals 1.05 and a down move that equals 0.97. If 1 million Bernoulli trials are conducted, and the average terminal stock price is $102.00, the probability of an up move (p) is *closest* to:

A 0.375.

B 0.500.

C 0.625.

28 A call option on a stock index is valued using a three-step binomial tree with an up move that equals 1.05 and a down move that equals 0.95. The current level of the index is $190, and the option exercise price is $200. If the option value is positive when the stock price exceeds the exercise price at expiration and $0 otherwise, the number of terminal nodes with a positive payoff is:

A one.

B two.

C three.

29 A random number between zero and one is generated according to a continuous uniform distribution. What is the probability that the first number generated will have a value of exactly 0.30?

A 0%

B 30%

C 70%

30 A Monte Carlo simulation can be used to:

A directly provide precise valuations of call options.

B simulate a process from historical records of returns.

C test the sensitivity of a model to changes in assumptions.

31 A limitation of Monte Carlo simulation is:

A its failure to do "what if" analysis.

B that it requires historical records of returns

C its inability to independently specify cause-and-effect relationships.

32 Which parameter equals zero in a normal distribution?

A Kurtosis

B Skewness

C Standard deviation

33 An analyst develops the following capital market projections.

	Stocks	Bonds
Mean Return	10%	2%
Standard Deviation	15%	5%

Assuming the returns of the asset classes are described by normal distributions, which of the following statements is correct?

A Bonds have a higher probability of a negative return than stocks.

B On average, 99% of stock returns will fall within two standard deviations of the mean.

C The probability of a bond return less than or equal to 3% is determined using a Z-score of 0.25.

34 A client holding a £2,000,000 portfolio wants to withdraw £90,000 in one year without invading the principal. According to Roy's safety-first criterion, which of the following portfolio allocations is optimal?

	Allocation A	Allocation B	Allocation C
Expected annual return	6.5%	7.5%	8.5%
Standard deviation of returns	8.35%	10.21%	14.34%

A Allocation A

B Allocation B

C Allocation C

35 In contrast to normal distributions, lognormal distributions:

A are skewed to the left.

B have outcomes that cannot be negative.

C are more suitable for describing asset returns than asset prices.

36 The lognormal distribution is a more accurate model for the distribution of stock prices than the normal distribution because stock prices are:

A symmetrical.

B unbounded.

C non-negative.

37 The price of a stock at t = 0 is \$208.25 and at t = 1 is \$186.75. The continuously compounded rate of return for the stock from t = 0 to t = 1 is *closest* to:

A –10.90%.

B –10.32%.

C 11.51%.

SOLUTIONS

1 **A** The put's minimum value is \$0. The put's value is \$0 when the stock price is at or above \$100 at the maturity date of the option. The put's maximum value is \$100 = \$100 (the exercise price) – \$0 (the lowest possible stock price). The put's value is \$100 when the stock is worthless at the option's maturity date. The put's minimum price increments are \$0.01. The possible outcomes of terminal put value are thus \$0.00, \$0.01, \$0.02, …, \$100.

B The price of the underlying has minimum price fluctuations of \$0.01: These are the minimum price fluctuations for terminal put value. For example, if the stock finishes at \$98.20, the payoff on the put is \$100 – \$98.20 = \$1.80. We can specify that the nearest values to \$1.80 are \$1.79 and \$1.81. With a continuous random variable, we cannot specify the nearest values. So, we must characterize terminal put value as a discrete random variable.

C The probability that terminal put value is less than or equal to \$24 is $P(Y \leq 24)$ or $F(24)$, in standard notation, where F is the cumulative distribution function for terminal put value.

2 A binomial random variable is defined as the number of successes in n Bernoulli trials (a trial that produces one of two outcomes). The binomial distribution is used to make probability statements about a record of successes and failures or about anything with binary (twofold) outcomes.

3 B is correct. The value of the cumulative distribution function lies between 0 and 1 for any x: $0 \leq F(x) \leq 1$.

4 C is correct. The binomial distribution is symmetric when the probability of success on a trial is 0.50, but it is asymmetric or skewed otherwise. Here it is given that $p = 0.50$

5 B is correct. The probability of any outcome is 0.05, $P(1) = 1/20 = 0.05$. The probability that X is greater than or equal to 3 but less than 6, which is expressed as $P(3 \leq X < 6) = P(3) + P(4) + P(5) = 0.05 + 0.05 + 0.05 = 0.15$

6 **A** The probability of an earnings increase (success) in a year is estimated as 7/10 = 0.70 or 70 percent, based on the record of the past 10 years.

B The probability that earnings will increase in 5 out of the next 10 years is about 10.3 percent. Define a binomial random variable X, counting the number of earnings increases over the next 10 years. From Part A, the probability of an earnings increase in a given year is $p = 0.70$ and the number of trials (years) is $n = 10$. Equation 1 gives the probability that a binomial random variable has x successes in n trials, with the probability of success on a trial equal to p.

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x}$$

For this example,

$$\binom{10}{5} 0.7^5 0.3^{10-5} = \frac{10!}{(10-5)!5!} 0.7^5 0.3^{10-5}$$
$$= 252 \times 0.16807 \times 0.00243 = 0.102919$$

We conclude that the probability that earnings will increase in exactly 5 of the next 10 years is 0.1029, or approximately 10.3 percent.

C The expected number of yearly increases is $E(X) = np = 10 \times 0.70 = 7$.

D The variance of the number of yearly increases over the next 10 years is $\sigma^2 = np(1 - p) = 10 \times 0.70 \times 0.30 = 2.1$. The standard deviation is 1.449 (the positive square root of 2.1).

E You must assume that 1) the probability of an earnings increase (success) is constant from year to year and 2) earnings increases are independent trials. If current and past earnings help forecast next year's earnings, Assumption 2 is violated. If the company's business is subject to economic or industry cycles, neither assumption is likely to hold.

7 B is correct. To calculate the probability of 4 years of outperformance, use the formula:

$$p(x) = P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} = \frac{n!}{(n - x)!x!} p^x (1 - p)^{n-x}$$

Using this formula to calculate the probability in 4 of 5 years, $n = 5$, $x = 4$ and $p = 0.60$.

Therefore,

$$p(4) = \frac{5!}{(5 - 4)!4!} 0.6^4 (1 - 0.6)^{5-4} = [120/24](0.1296)(0.40) = 0.2592$$

$$p(5) = \frac{5!}{(5 - 5)!5!} 0.6^5 (1 - 0.6)^{5-5} = [120/120](0.0778)(1) = 0.0778$$

The probability of outperforming 4 or more times is $p(4) + p(5) = 0.2592 + 0.0778 = 0.3370$

8 The observed success rate is 4/7 = 0.571, or 57.1 percent. The probability of four or fewer successes is $F(4) = p(4) + p(3) + p(2) + p(1) + p(0)$, where $p(4)$, $p(3)$, $p(2)$, $p(1)$, and $p(0)$ are respectively the probabilities of 4, 3, 2, 1, and 0 successes, according to the binomial distribution with $n = 7$ and $p = 0.70$. We have

$p(4) = (7!/4!3!)(0.70^4)(0.30^3) = 35(0.006483) = 0.226895$

$p(3) = (7!/3!4!)(0.70^3)(0.30^4) = 35(0.002778) = 0.097241$

$p(2) = (7!/2!5!)(0.70^2)(0.30^5) = 21(0.001191) = 0.025005$

$p(1) = (7!/1!6!)(0.70^1)(0.30^6) = 7(0.000510) = 0.003572$

$p(0) = (7!/0!7!)(0.70^0)(0.30^7) = 1(0.000219) = 0.000219$

Summing all these probabilities, you conclude that $F(4) = 0.226895 + 0.097241 + 0.025005 + 0.003572 + 0.000219 = 0.352931$, or 35.3 percent.

9 A The expected value of fourth-quarter sales is €14,500,000, calculated as (€14,000,000 + €15,000,000)/2. With a continuous uniform random variable, the mean or expected value is the midpoint between the smallest and largest values. (See Example 7.)

B The probability that fourth-quarter sales will be less than €14,125,000 is 0.125 or 12.5 percent, calculated as (€14,125,000 – €14,000,000)/(€15,000,000 – €14,000,000).

10 A Approximately 68 percent of all outcomes of a normal random variable fall within plus or minus one standard deviation of the mean.

B Approximately 95 percent of all outcomes of a normal random variable fall within plus or minus two standard deviations of the mean.

C Approximately 99 percent of all outcomes of a normal random variable fall within plus or minus three standard deviations of the mean.

11 The area under the normal curve for $z = 0.36$ is 0.6406 or 64.06 percent. The following table presents an excerpt from the tables of the standard normal cumulative distribution function in the back of this volume. To locate $z = 0.36$, find 0.30 in the fourth row of numbers, then look at the column for 0.06 (the second decimal place of 0.36). The entry is 0.6406.

$P(Z \le x) = N(x)$ for $x \ge 0$ or $P(Z \le z) = N(z)$ for $z \ge 0$

x or z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.10	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.20	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.30	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	**0.6406**	0.6443	0.6480	0.6517
0.40	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.50	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224

The interpretation of 64.06 percent for $z = 0.36$ is that 64.06 percent of observations on a standard normal random variable are smaller than or equal to the value 0.36. (So 100% − 64.06% = 35.94% of the values are greater than 0.36.)

12 C is correct. The probability that the performance is at or below the expectation is calculated by finding $F(3) = p(3) + p(2) + p(1)$ using the formula:

$$p(x) = P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} = \frac{n!}{(n - x)!x!} p^x (1 - p)^{n-x}$$

Using this formula,

$$p(3) = \frac{4!}{(4 - 3)!3!} 0.75^3 (1 - 0.75)^{4-3} = [24/6](0.42)(0.25) = 0.42$$

$$p(2) = \frac{4!}{(4 - 2)!2!} 0.75^2 (1 - 0.75)^{4-2} = [24/4](0.56)(0.06) = 0.20$$

$$p(1) = \frac{4!}{(4 - 1)!1!} 0.75^1 (1 - 0.75)^{4-1} = [24/6](0.75)(0.02) = 0.06$$

$$p(0) = \frac{4!}{(4 - 0)!0!} 0.75^0 (1 - 0.75)^{4-0} = [24/24](1)(0.004) = 0.004$$

Therefore,

$F(3) = p(3) + p(2) + p(1) + p(0) = 0.42 + 0.20 + 0.06 + 0.004 = 0.684$ or approximately 68 percent

13 **A** The probability of exhausting the liquidity pool is 4.7 percent. First calculate $x = \lambda / \left(\sigma\sqrt{T}\right) = \$2{,}000 / \left(\$450\sqrt{5}\right) = 1.987616$. We can round this value to 1.99 to use the standard normal tables in the back of this book. Using those tables, we find that $N(1.99) = 0.9767$. Thus, the probability of exhausting the liquidity pool is $2[1 - N(1.99)] = 2(1 - 0.9767) = 0.0466$ or about 4.7 percent.

B The probability of exhausting the liquidity pool is now 32.2 percent. The calculation follows the same steps as those in Part A. We calculate $x = \lambda/\left(\sigma\sqrt{T}\right) = \$2{,}000/\left(\$450\sqrt{20}\right) = 0.993808$. We can round this value to 0.99 to use the standard normal tables in the back of this book. Using those tables, we find that $N(0.99) = 0.8389$. Thus, the probability of exhausting the liquidity pool is $2[1 - N(0.99)] = 2(1 - 0.8389) = 0.3222$ or about 32.2 percent. This is a substantial probability that you will run out of funds to meet mark to market.

In their paper, Kolb et al. call the probability of exhausting the liquidity pool the probability of ruin, a traditional name for this type of calculation.

14 B is correct. The normal distribution has a skewness of 0, a kurtosis of 3, and a mean, median and mode that are all equal.

15 B is correct. Multivariate distributions specify the probabilities for a group of related random variables. A portfolio of technology stocks represents a group of related assets. Accordingly, statistical interrelationships must be considered, resulting in the need to use a multivariate normal distribution.

16 C is correct. A bivariate normal distribution (two stocks) will have two means, two variances and one correlation. A multivariate normal distribution for the returns on n stocks will have n means, n variances and $n(n - 1)/2$ distinct correlations.

17 **A** Because £50,000/£1,350,000 is 3.7 percent, for any return less than 3.7 percent the client will need to invade principal if she takes out £50,000. So R_L = 3.7 percent.

B To decide which of the allocations is safety-first optimal, select the alternative with the highest ratio $[E(R_P) - R_L]/\sigma_P$:

Allocation A: 0.5125 = (16 – 3.7)/24

Allocation B: 0.488235 = (12 – 3.7)/17

Allocation C: 0.525 = (10 – 3.7)/12

Allocation D: 0.481818 = (9 – 3.7)/11

Allocation C, with the largest ratio (0.525), is the best alternative according to the safety-first criterion.

C To answer this question, note that $P(R_C < 3.7) = N(-0.525)$. We can round 0.525 to 0.53 for use with tables of the standard normal cdf. First, we calculate $N(-0.53) = 1 - N(0.53) = 1 - 0.7019 = 0.2981$ or about 30 percent. The safety-first optimal portfolio has a roughly 30 percent chance of not meeting a 3.7 percent return threshold.

18 A is correct. P(8% ≤ Portfolio return ≤ 11%) = N(Z corresponding to 11%) – N(Z corresponding to 8%). For the first term, Z = (11% – 8%)/14% = 0.21 approximately, and using the table of cumulative normal distribution given in the problem, $N(0.21) = 0.5832$. To get the second term immediately, note that 8 percent is the mean, and for the normal distribution 50 percent of the probability lies on either side of the mean. Therefore, N(Z corresponding to 8%) must equal 50 percent. So P(8% ≤ Portfolio return ≤ 11%) = 0.5832 – 0.50 = 0.0832 or approximately 8.3 percent.

19 B is correct. There are three steps, which involve standardizing the portfolio return: First, subtract the portfolio mean return from each side of the inequality: P(Portfolio return – 7%) ≤ 4% – 7%). Second, divide each side of the inequality by the standard deviation of portfolio return: P[(Portfolio return

$-$ 7%)/13% ≤ (4% $-$ 7%)/13%] = $P(Z \leq -0.2308) = N(-0.2308)$. Third, recognize that on the left-hand side we have a standard normal variable, denoted by Z and $N(-x) = 1 - N(x)$. Rounding -0.2308 to -0.23 for use with the cumulative distribution function (cdf) table, we have $N(-0.23) = 1 - N(0.23) = 1 - 0.5910 = 0.409$, approximately 41 percent. The probability that the portfolio will underperform the target is about 41 percent.

20 **A** Elements that should appear in a definition of Monte Carlo simulation are that it makes use of a computer; that it is used to represent the operation of a complex system, or in some applications, to find an approximate solution to a problem; and that it involves the generation of a large number of random samples from a specified probability distribution. The exact wording can vary, but one definition follows:

Monte Carlo simulation in finance involves the use of a computer to represent the operation of a complex financial system. In some important applications, Monte Carlo simulation is used to find an approximate solution to a complex financial problem. An integral part of Monte Carlo simulation is the generation of a large number of random samples from a probability distribution.

B *Strengths.* Monte Carlo simulation can be used to price complex securities for which no analytic expression is available, particularly European-style options.

Weaknesses. Monte Carlo simulation provides only statistical estimates, not exact results. Analytic methods, when available, provide more insight into cause-and-effect relationships than does Monte Carlo simulation.

21 In the text, we described how we could use Monte Carlo simulation to value an Asian option, a complex European-style option. Just as we can calculate the average value of the stock over a simulation trial to value an Asian option, we can also calculate the minimum value of the stock over a simulation trial. Then, for a given simulation trial, we can calculate the terminal value of the call, given the minimum value of the stock for the simulation trial. We can then discount back this terminal value to the present to get the value of the call today ($t = 0$). The average of these $t = 0$ values over all simulation trials is the Monte Carlo simulated value of the lookback call option.

22 C is correct. The rate of return is a random variable because the future outcomes are uncertain, and it is continuous because it can take on an unlimited number of outcomes.

23 B is correct. The function $g(x)$ satisfies the conditions of a probability function. All of the values of $g(x)$ are between 0 and 1, and the values of $g(x)$ all sum to 1.

24 A is correct. The probability that X will take on a value of 4 or less is: $F(4) = P(X \leq 4) = p(1) + p(2) + p(3) + p(4) = 0.60$. The probability that X will take on a value of 3 or less is: $F(3) = P(X \leq 3) = p(1) + p(2) + \mathrm{p}(3) = 0.50$. So, the probability that X will take on a value of 4 is: $F(4) - F(3) = p(4) = 0.10$. The probability of $X = 2$ can be found using the same logic: $F(2) - F(1) = p(2) = 0.25 - 0.15 = 0.10$. The probability of X taking on a value of 2 or 4 is: $p(2) + p(4) = 0.10 + 0.10 = 0.20$.

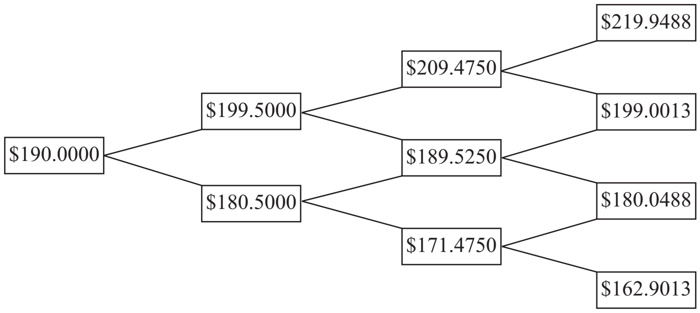
25 A is correct. A trial, such as a coin flip, will produce one of two outcomes. Such a trial is a Bernoulli trial.

26 A is correct. The continuously compounded return of an asset over a period is equal to the natural log of period's change. In this case:

$$\ln(120/112) = 6.90\%$$

27 C is correct. The probability of an up move (p) can be found by solving the equation: $(p)uS + (1 - p)dS = (p)105 + (1 - p)97 = 102$. Solving for p gives $8p = 5$, so that $p = 0.625$.

28 A is correct. Only the top node value of \$219.9488 exceeds \$200.



29 A is correct. The probability of generating a random number equal to any fixed point under a continuous uniform distribution is zero.

30 C is correct. A characteristic feature of Monte Carlo simulation is the generation of a large number of random samples from a specified probability distribution or distributions to represent the role of risk in the system.

31 C is correct. Monte Carlo simulation is a complement to analytical methods. Monte Carlo simulation provides statistical estimates and not exact results. Analytical methods, when available, provide more insight into cause-and-effect relationships

32 B is correct. A normal distribution has a skewness of zero (it is symmetrical around the mean). A non-zero skewness implies asymmetry in a distribution.

33 A is correct. The chance of a negative return falls in the area to the left of 0% under a standard normal curve. By standardizing the returns and standard deviations of the two assets, the likelihood of either asset experiencing a negative return may be determined: Z-score (standardized value) = $(X - \mu)/\sigma$

Z-score for a bond return of 0% = $(0 - 2)/5 = -0.40$.

Z-score for a stock return of 0% = $(0 - 10)/15 = -0.67$.

For bonds, a 0% return falls 0.40 standard deviations below the mean return of 2%. In contrast, for stocks, a 0% return falls 0.67 standard deviations below the mean return of 10%. A standard deviation of 0.40 is less than a standard deviation of 0.67. Negative returns thus occupy more of the left tail of the bond distribution than the stock distribution. Thus, bonds are more likely than stocks to experience a negative return.

34 B is correct. Allocation B has the highest safety-first ratio. The threshold return level R_L for the portfolio is £90,000/£2,000,000 = 4.5%, thus any return less than R_L = 4.5% will invade the portfolio principal. To compute the allocation that is safety-first optimal, select the alternative with the highest ratio:

$$\frac{\left[E\left(R_P - R_L\right)\right]}{\sigma_P}$$

$$\text{Allocation A} = \frac{6.5 - 4.5}{8.35} = 0.240$$

$$\text{Allocation B} = \frac{7.5 - 4.5}{10.21} = 0.294$$

$$\text{Allocation C} = \frac{8.5 - 4.5}{14.34} = 0.279$$

35 B is correct. By definition, lognormal random variables cannot have negative values.

36 C is correct. A lognormal distributed variable has a lower bound of zero. The lognormal distribution is also right skewed, which is a useful property in describing asset prices.

37 A is correct. The continuously compounded return from $t = 0$ to $t = 1$ is $r_{0,1} = \ln(S_1/S_0) = \ln(186.75/208.25) = -0.10897 = -10.90\%$.