CHAPTER 12

Simulation

Chapter 1 discussed how the calculations in a spreadsheet can be viewed as a mathematical model that defines a functional relationship between various input variables (or independent variables) and one or more bottom-line performance measures (or dependent variables). The following equation expresses this relationship:

$$Y = f(X_1, X_2, ..., X_k)$$

In many spreadsheets, the values of various input cells are determined by the person using the spreadsheet. These input cells correspond to the independent variables $X_1, X_2, ..., X_k$ in the above equation. Various formulas (represented by $f(\cdot)$ above) are entered in other cells of the spreadsheet to transform the values of the input cells into some bottom-line output (denoted by Y above). Simulation is a technique that is helpful in analyzing models where the value to be assumed by one or more independent variables is uncertain.

12.1 RANDOM VARIABLES AND RISK

In order to compute a value for the bottom-line performance measure of a spreadsheet model, each input cell must be assigned a specific value so that all the related calculations can be performed. However, some uncertainty often exists regarding the value that should be assumed by one or more independent variables (or input cells) in the spreadsheet. This is particularly true in spreadsheet models that represent future conditions. A **random variable** is any variable whose value cannot be predicted or set with certainty. Thus, many input variables in a spreadsheet model represent random variables whose actual values cannot be predicted with certainty.

For example, projections of the cost of raw materials, future interest rates, future numbers of employees, and expected product demand are random variables because their true values are unknown and will be determined in the future. If we cannot say with certainty what value one or more input variables in a model will assume, we also cannot say with certainty what value the dependent variable will assume. This uncertainty associated with the value of the dependent variable introduces an element of risk to the decision-making problem. Specifically, if the dependent variable represents some bottom-line performance measure that managers use to make decisions, and its value is uncertain, any decisions made on the basis of this value are based on uncertain (or incomplete) information. When such a decision is made, some chance exists that the decision will not produce the intended results. This chance, or uncertainty, represents an element of **risk** in the decision-making problem.

The term "risk" also implies the *potential* for loss. The fact that a decision's outcome is uncertain does not mean that the decision is particularly risky. For example, whenever we put money into a soft drink machine, there is a chance that the machine will take our money and not deliver the product. However, most of us would not consider this risk to be particularly great. From past experience, we know that the chance of not receiving the product is small. But even if the machine takes our money and does not deliver the product, most of us would not consider this to be a tremendous loss. Thus, the amount of risk involved in a given decision-making situation is a function of the uncertainty in the outcome of the decision and the magnitude of the potential loss. A proper assessment of the risk present in a decision-making situation should address both of these issues, as the examples in this chapter will demonstrate.

12.2 WHY ANALYZE RISK?

Many spreadsheets built by business people contain *estimated* values for the uncertain input variables in their models. If a manager cannot say with certainty what value a particular cell in a spreadsheet will assume, this cell most likely represents a random variable. Ordinarily, the manager will attempt to make an informed guess about the values such cells will assume. The manager hopes that inserting the expected, or most likely, values for all the uncertain cells in a spreadsheet will provide the most likely value for the cell containing the bottom-line performance measure (Y). The problem with this type of analysis is that it tells the decision maker nothing about the *variability* of the performance measure.

For example, in analyzing a particular investment opportunity, we might determine that the expected return on a $1,000 investment is $10,000 within two years. But how much variability exists in the possible outcomes? If all the potential outcomes are scattered closely around $10,000 (say from $9,000 to $11,000), then the investment opportunity might still be attractive. If, on the other hand, the potential outcomes are scattered widely around $10,000 (say from –$30,000 to +$50,000), then the investment opportunity might be unattractive. Although these two scenarios might have the same expected or average value, the risks involved are quite different. Thus, even if we can determine the expected outcome of a decision using a spreadsheet, it is just as important, if not more so, to consider the risk involved in the decision.

12.3 METHODS OF RISK ANALYSIS

Several techniques are available to help managers analyze risk. Three of the most common are best-case/worst-case analysis, what-if analysis, and simulation. Of these methods, simulation is the most powerful and, therefore, is the technique we will focus on in this chapter. Although the other techniques might not be effective in risk analysis, they are probably used more often than simulation by most managers in business today. This is largely due to the fact that most managers are unaware of the spreadsheet's ability to perform simulation and of the benefits provided by this technique. Before discussing simulation, let's first briefly look at the other methods of risk analysis to understand their strengths and weaknesses.

12.3.1 Best-Case/Worst-Case Analysis

If we don't know what value a particular cell in a spreadsheet will assume, we could enter a number that we think is the most likely value for the uncertain cell. If we enter such numbers for all the uncertain cells in the spreadsheet, we can easily calculate the most likely value of the bottom-line performance measure. (This is also called the **base-case** scenario.) However, this scenario gives us no information about how far away the actual outcome might be from this expected or most likely value.

One simple solution to this problem is to calculate the value of the bottom-line performance measure using the **best-case,** or most optimistic, and **worst-case,** or most pessimistic, values for the uncertain input cells. These additional scenarios show the range of possible values that might be assumed by the bottom-line performance measure. As indicated in the earlier example about the $1,000 investment, knowing the range of possible outcomes is very helpful in assessing the risk involved in different alternatives. However, simply knowing the best-case and worst-case outcomes tells us nothing about the distribution of possible values within this range, nor does it tell us the probability of either scenario occurring.

Figure 12.1 displays several probability distributions that might be associated with the value of a bottom-line performance measure within a given range. Each of these distributions describe variables that have identical ranges and similar average values. But each distribution is very different in terms of the risk it represents to the decision maker. The appeal of best-case/worst-case analysis is that it is easy to do. Its weakness is that it tells us nothing about the shape of the distribution associated with the bottom-line performance measure. As we'll see later, knowing the shape of the distribution of the bottom-line performance measure can be critically important in helping us answer a number of managerial questions.

12.3.2 What-If Analysis

Prior to the introduction of electronic spreadsheets in the early 1980s, the use of best-case/worst-case analysis was often the only feasible way for a manager to analyze the risk associated with a decision. This process was extremely time consuming, error prone, and tedious, using only a piece of paper, pencil, and calculator to recalculate the performance measure of a model using different values for the uncertain inputs. The arrival of personal computers and electronic spreadsheets made it much easier for a manager to play out a large number of scenarios in addition to the best and worst cases—which is the essence of what-if analysis.

In **what-if analysis**, a manager changes the values of the uncertain input variables to see what happens to the bottom-line performance measure. By making a series of such changes, a manager can gain some insight into how sensitive the performance measure is to changes to the input variables. Although many managers perform this type of manual what-if analysis, it has three major flaws.

First, if the values selected for the independent variables are based on only the manager's judgment, the resulting sample values of the performance measure are

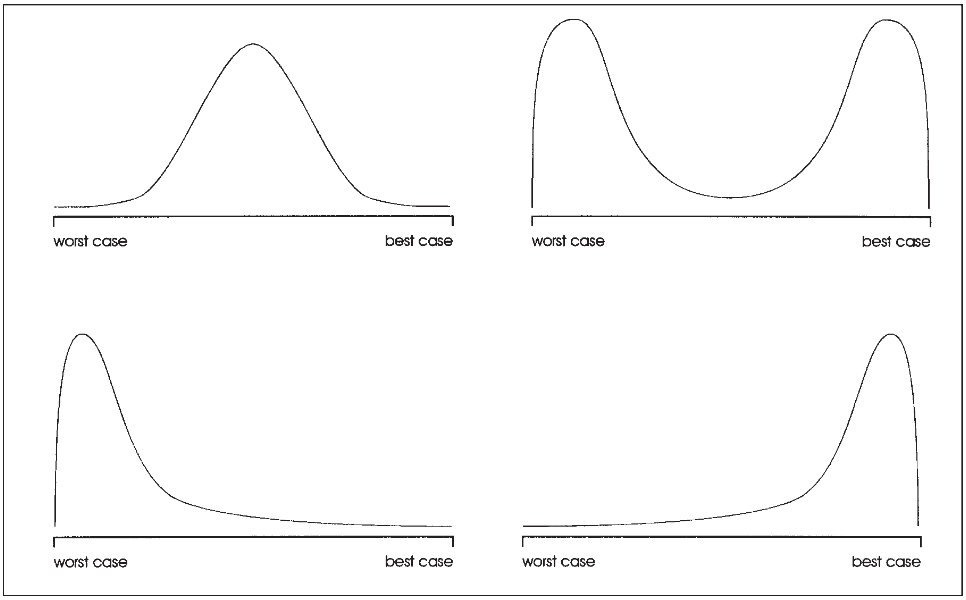


Figure 12.1 Possible distributions of performance measure values within a given range.

likely to be biased. That is, if several uncertain variables can each assume some range of values, it would be difficult to ensure that the manager tests a fair, or representative, sample of all possible combinations of these values. To select values for the uncertain variables that correctly reflect their random variations, the values must be randomly selected from a distribution, or pool, of values that reflects the appropriate range of possible values as well as the appropriate relative frequencies of these variables.

Second, hundreds or thousands of what-if scenarios might be required to create a valid representation of the underlying variability in the bottom-line performance measure. No one would want to perform these scenarios manually nor would anyone be able to make sense of the resulting stream of numbers that would flash by on the screen.

The third problem with what-if analysis is that the insight the manager might gain from playing out various scenarios is of little value when recommending a decision to top management. What-if analysis simply does not supply the manager with the tangible evidence (facts and figures) needed to justify why a given decision was made or recommended. Additionally, what-if analysis does not address the problem identified in our earlier discussion of best-case/worst-case analysis—it does not allow us to estimate the distribution of the performance measure in a formal enough manner. Thus, what-if analysis is a step in the right direction, but it's not quite a large enough step to allow managers to analyze risk effectively in the decisions they face.

12.3.3 Simulation

Simulation is a technique that measures and describes various characteristics of the bottom-line performance measure of a model when one or more values for the independent variables are uncertain. If any independent variables in a model are random variables, the dependent variable (Y) also represents a random variable. The objective in simulation is to *describe* the distribution and characteristics of the possible values of the bottom-line performance measure Y, given the possible values and behavior of the independent variables $X_1, X_2, ..., X_k$.

The idea behind simulation is similar to the notion of playing out many what-if scenarios. The difference is that the process of assigning values to the cells in the spreadsheet that represent random variables is automated so that: (1) the values are assigned in a nonbiased way, and (2) the spreadsheet user is relieved of the burden of determining these values. With simulation, we repeatedly and randomly generate sample values for each uncertain input variable ($X_1, X_2, ..., X_k$) in our model and then compute the resulting value of our bottom-line performance measure (Y). We can then use the sample values of Y to estimate the true distribution and other characteristics of the performance measure Y. For example, we can use the sample observations to construct a frequency distribution of the performance measure, to estimate the range of values over which the performance measure might vary, to estimate the performance measure mean and variance, and to estimate the probability that the actual value of the performance measure will be greater than (or less than) a particular value. All these measures provide greater insight into the risk associated with a given decision than a single value calculated based on the expected values for the uncertain independent variables.

12.4 A CORPORATE HEALTH INSURANCE EXAMPLE

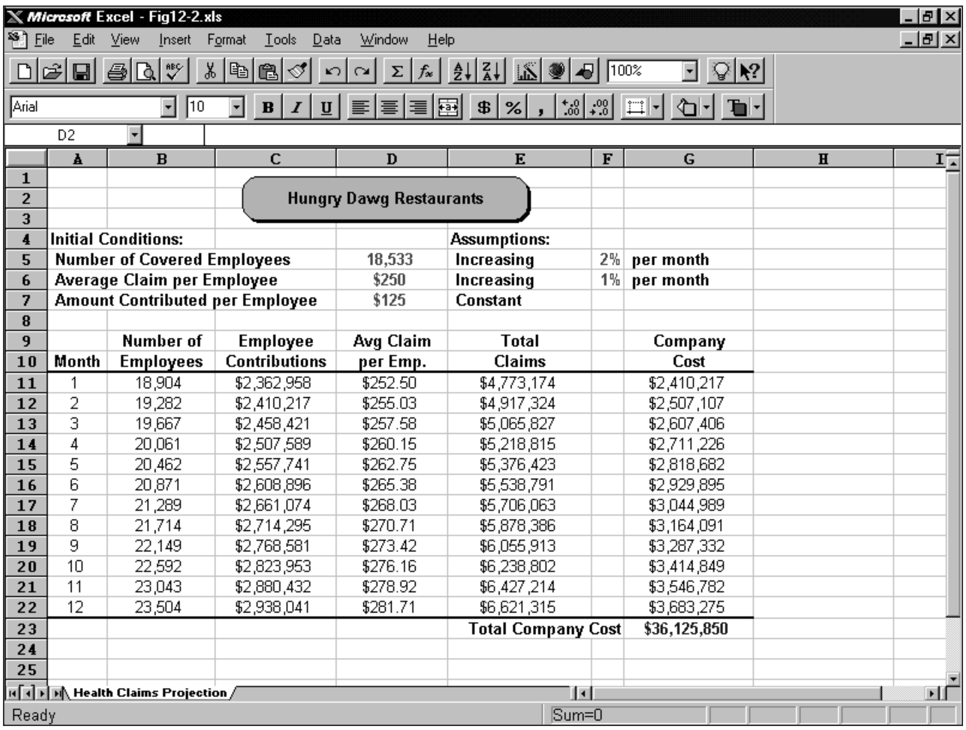
The following example demonstrates the mechanics of preparing a spreadsheet model for risk analysis using simulation. The example presents a fairly simple model to illustrate the process and provide a sense of the amount of effort involved. However, the process for performing simulation is basically the same regardless of the size of the model.

Lisa Pon has just been hired as an analyst in the corporate planning department of Hungry Dawg Restaurants. Her first assignment is to determine how much money the company needs to accrue in the coming year to pay for its employees' health insurance claims. Hungry Dawg is a large, growing chain of restaurants that specializes in traditional southern foods. The company has become large enough that it no longer buys insurance from a private insurance company. The company is now self-insured, meaning that it pays health insurance claims with its own money (although it contracts with an outside company to handle the administrative details of processing claims and writing checks).

The money the company uses to pay claims comes from two sources: employee contributions (or premiums deducted from employees' paychecks) and company funds (the company must pay whatever costs are not covered by employee contributions). Each employee covered by the health plan contributes $125 per month. However, the number of employees covered by the plan changes from month to month as employees are hired and fired, quit, or simply add or drop health insurance coverage. A total of 18,533 employees were covered by the plan last month. The average monthly health claim per covered employee was $250 last month.

An example of how most analysts would model this problem is shown in Figure 12.2 (and in the file FIG12-2.xls on your data disk). The spreadsheet begins with a listing of the initial conditions and assumptions for the problem. For example, cell D5 indicates that 18,533 employees are covered currently by the health plan, and cell D6 indicates that the average monthly claim per covered employee is $250. The average monthly contribution per employee is $125, as shown in cell D7. The values in cells D5 and D6 are unlikely to stay the same for the entire year. Thus, we need to make some assumptions about the rate at which these values are likely to increase during the year. For example, we might assume that the number of covered employees will increase by about 2% per month, and that the average claim per employee will increase at a rate of 1% per month. These assumptions are reflected in cells F5 and F6. The average contribution per employee is assumed to be constant over the coming year.

Using the assumed rate of increase in the number of covered employees (cell F5), we can create formulas for cells B11 through B22 that cause the number of covered employees to increase by the assumed amount each month. (The details of these formulas are covered later.) The expected monthly employee contributions shown in column C are calculated as $125 times the number of employees in each month. We can use the assumed rate of increase in average monthly claims (cell F6) to create formulas for cells D11 through D22 that cause the average claim per employee to increase at the assumed rate. The total claims for each month (shown in column E) are calculated as the average claim figures in column D times the number of employees for each month in column B. Because the company must pay for any claims that are not covered by the employee contributions, the company cost figures in column G are calculated as the total claims minus the employee contributions (column E minus column C). Finally, cell G23 sums the company cost figures listed in column G,



	A	B	C	D	E	F	G
2			Hungry Dawg Restaurants				
4	Initial Conditions:				Assumptions:		
5	Number of Covered Employees			18,533	Increasing	2%	per month
6	Average Claim per Employee			$250	Increasing	1%	per month
7	Amount Contributed per Employee			$125	Constant		
9		Number of	Employee	Avg Claim	Total		Company
10	Month	Employees	Contributions	per Emp.	Claims		Cost
11	1	18,904	$2,362,958	$252.50	$4,773,174		$2,410,217
12	2	19,282	$2,410,217	$255.03	$4,917,324		$2,507,107
13	3	19,667	$2,458,421	$257.58	$5,065,827		$2,607,406
14	4	20,061	$2,507,589	$260.15	$5,218,815		$2,711,226
15	5	20,462	$2,557,741	$262.75	$5,376,423		$2,818,682
16	6	20,871	$2,608,896	$265.38	$5,538,791		$2,929,895
17	7	21,289	$2,661,074	$268.03	$5,706,063		$3,044,989
18	8	21,714	$2,714,295	$270.71	$5,878,386		$3,164,091
19	9	22,149	$2,768,581	$273.42	$6,055,913		$3,287,332
20	10	22,592	$2,823,953	$276.16	$6,238,802		$3,414,849
21	11	23,043	$2,880,432	$278.92	$6,427,214		$3,546,782
22	12	23,504	$2,938,041	$281.71	$6,621,315		$3,683,275
23					Total Company Cost		$36,125,850

Figure 12.2 Original corporate health insurance model with expected values for uncertain variables.

and shows that the company can expect to contribute $36,125,850 of its revenues toward paying the health insurance claims of its employees in the coming year.

12.4.1 A Critique of the Base-Case Model

Now let's consider the model we just described. The example model assumes that the number of covered employees will increase by *exactly* 2% each month and that the average claim per covered employee will increase by *exactly* 1% each month. Although these values might be reasonable approximations of what might happen, they are unlikely to reflect exactly what will happen. In fact, the number of employees covered by the health plan each month is likely to vary randomly around the average increase per month—that is, the number might decrease in some months and increase by more than 2% in others. Similarly, the average claim per covered employee might be lower than expected in certain months and higher than expected in others.

Both of these figures are likely to exhibit some uncertainty or random behavior, even if they do move in the general upward direction assumed throughout the year. So we cannot say with certainty that the total cost figure of $36,125,850 is exactly what the company will have to contribute toward health claims in the coming year. It is simply a prediction of what might happen. The actual outcome could be smaller or larger than this estimate. Using the original model, we have no idea how much larger or smaller the actual result could be—nor do we have any idea how the actual values are distributed around this estimate. We do not know if there is a 10%, 50%, or 90% chance of the actual total costs exceeding this estimate. To determine the variability or risk inherent in the bottom-line performance measure of total company costs, we'll apply the technique of simulation to our model.

12.5 RANDOM NUMBER GENERATORS

To perform simulation in an electronic spreadsheet, we must first place a **random number generator** (RNG) formula in each cell that represents a random, or uncertain, independent variable. Each RNG provides a sample observation from an appropriate distribution that represents the range and frequency of possible values for the variable. Once the RNGs are in place, new sample values are provided automatically by the RNGs each time the spreadsheet is recalculated. We can recalculate the spreadsheet n times, where n is the desired number of replications or scenarios, and the value of the bottom-line performance measure will be stored after each replication. We can analyze these stored observations to gain insight into the behavior and characteristics of the performance measure.

The process of simulation involves a lot of work, but, fortunately, the spreadsheet can do most of the work for us fairly easily. As mentioned earlier, the first step is to place an RNG formula in each cell that contains an uncertain value. Each formula will generate (or return) a number that represents a randomly selected value from a distribution, or pool, of values. The distributions from which these samples are taken should be representative of the underlying pool of values expected to occur in each uncertain cell.

12.5.1 The RAND( ) Function

Excel, like most other spreadsheet packages, includes a built-in function called RAND( ) that provides the foundation for creating various RNGs. The **RAND( )**

function returns a uniformly distributed random number between 0.0 and 1.0 whenever the spreadsheet is recalculated (technically, $0 \leq \text{RAND( )} < 0.99\overline{9}$). If you enter the RAND( ) formula in some cell in a spreadsheet, and press the recalculate key (function key [F9]) repeatedly, a series of random numbers between 0.0 and 1.0 appears in the cell.

The RAND( ) function enables us to do some interesting modeling. As a simple example, suppose that we want to simulate the toss of a fair coin in a spreadsheet. When tossing a fair coin, there are two possible outcomes: heads or tails. If the coin is fair (that is, not weighted or biased toward one outcome over the other), we expect that each outcome has an equal probability of occurring each time we toss the coin. That is, the probability of heads is 0.5 and the probability of tails is 0.5 on any toss. However, before any given toss we cannot say with certainty whether the observed outcome will be heads or tails.

Using the RAND( ) function, we can create a formula that simulates the process of tossing the coin. Suppose that the value 1 represents the occurrence of heads and the value 0 represents the occurrence of tails. Now consider the following formula:

IF(RAND( )<0.5,1,0)

Whenever the spreadsheet is recalculated, the RAND( ) function will return a random value between 0 and 1. If the value returned by RAND( ) is less than 0.5, the previous IF( ) function will return the value 1 (representing heads); otherwise, it will return the value 0 (representing tails). Because the RAND( ) function has a 0.5 probability of returning a value less than 0.5, there is a 50% chance that the IF( ) function will generate the heads value, and a 50% chance that it will generate the tails value each time the spreadsheet is recalculated.

As another example, suppose that we want to simulate rolling a fair, six-sided die using a spreadsheet. In this case, each roll of the die can produce one of six possible outcomes (the value 1, 2, 3, 4, 5, or 6). We need an RNG that randomly generates the integer numbers from 1 to 6 with each value having a 1/6 chance of occurring. Because RAND( ) generates uniformly distributed random numbers between 0.0 and 1.0, 6*RAND( ) would generate uniformly distributed random numbers between 0.0 and 6.0 (technically, $0 \leq 6*\text{RAND( )} \leq 5.99\overline{9}$).

Now suppose that we took this interval from 0.0 to 6.0 and divided it into six equal pieces as follows:

	Lower Limit	*Upper Limit*
1	0.0	$0.99\overline{9}$
2	1.0	$1.99\overline{9}$
3	2.0	$2.99\overline{9}$
4	3.0	$3.99\overline{9}$
5	4.0	$4.99\overline{9}$
6	5.0	$5.99\overline{9}$

Because each of the six intervals is exactly the same size, the value of 6∗RAND( ) is equally likely to fall in each of them. If the value generated by 6∗RAND( ) falls between 0.0 and $0.99\overline{9}$, we could associate this with the outcome of rolling a 1 on the die; if 6∗RAND( ) falls between 1.0 and $1.99\overline{9}$, we could associate this with rolling a 2 on the die, and so forth. This makes sense, but what mathematical function makes this association happen? Consider the following formula:

INT(6*RAND( ))+1

The INT( ) function returns the integer portion of the value inside its parentheses (for example, INT(4.85) = 4, INT(2.13) = 2). The following table summarizes the different outcomes generated using this formula:

*If 6*RAND( ) falls in the interval*	*INT(6*RAND( ))+1 returns the value*
0.0 to $0.99\overline{9}$	1
1.0 to $1.99\overline{9}$	2
2.0 to $2.99\overline{9}$	3
3.0 to $3.99\overline{9}$	4
4.0 to $4.99\overline{9}$	5
5.0 to $5.99\overline{9}$	6

Again, because each interval in the table is exactly the same size, the value of 6∗RAND( ) is equally likely to fall in each interval. Therefore, each value generated by the formula INT(6∗RAND( ))+1 also is equally likely to occur. Thus, the formula INT(6∗RAND( ))+1 accurately simulates the act of rolling a fair, six-sided die.

12.5.2 RNG Functions

The two previous examples represent random variables that follow the **discrete uniform distribution,** which is appropriate for modeling random variables with n distinct possible outcomes, each outcome being equally likely (or having a $1/n$ probability of occurring). The following formula can be used to randomly generate the integers $a, a + 1, a + 2, ..., a + n - 1$ with equal probability of occurrence:

RNG for the discrete uniform distribution: INT(n∗RAND( ))+a **12.2**

This formula is equivalent to the formula used in the die rolling example, where $a = 1$ and $n = 6$. Also, we could have used this same formula with $a = 0$ and $n = 2$ to simulate the toss of a fair coin.

Figure 12.3 (and the file FIG12-3.xls on your data disk) gives an example of the RNG for the discrete uniform distribution and several other formulas that utilize the RAND( ) function to generate random numbers from several other probability distributions. Notice that the numbers shown in column D of this spreadsheet will change each time the spreadsheet is recalculated (by pressing the [F9] function key).

While it is possible to use the formulas shown in Figure 12.3 to generate random numbers in Excel, it is more convenient (and less error prone) to use the Visual Basic for Applications (VBA) macro language to create user defined functions that implement various RNGs. A VBA add-in file named RNG.xla (found on your data disk) was created to simplify the process of using RNGs in Excel. Refer to the box titled "Installing and Using the RNG.xla Add-In" for instructions on installing this add-in on your computer. Figure 12.4 describes the RNG functions implemented in the RNG.xla add-in.

The functions listed in Figure 12.4 allows us to generate a variety of random numbers easily. For example, if we think that the behavior of an uncertain cell could be modeled as a normally distributed random variable with a mean of 125 and standard deviation of 10, then according to Figure 12.4 we could enter the formula =RNGNormal(125,10) in this cell. (The arguments in this function could also be formulas and could refer to

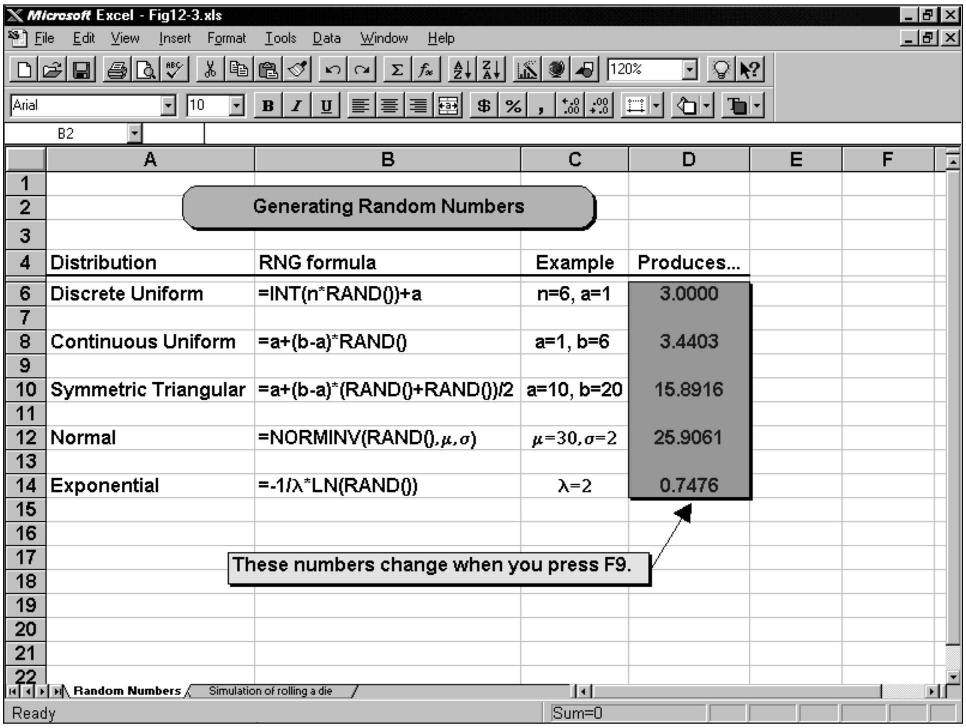


Figure 12.3
Examples of common RNGs constructed with the RAND( ) function.

Installing and Using the RNG.XLA Add-In

In order to access the functions shown in Figure 12.4, you must first install the RNG.xla add-in on your computer. To do this:

1. Copy the RNG.xla file to your hard drive (preferably to the directory C:\MSOffice\Excel\Library).
2. In Excel, select Tools, Add-Ins, click the Browse button, locate the RNG.xla file, and click OK.

This instructs your computer to always open the RNG.xla add-in whenever you start Excel. You can deselect the RNG.xla at any time by using the Tools, Add-Ins command again. Excel will not be able to properly interpret files that make use of the functions in Figure 12.4 unless the RNG.xla add-in is installed on your computer in the manner outlined above.

other cells in the spreadsheet.) Whenever the spreadsheet is recalculated, this function would return a randomly selected value from a normal distribution with a mean of 125 and standard deviation of 10.

Similarly, a cell in our spreadsheet might have a 30% chance of assuming the value 10, a 50% chance of assuming the value 20, and a 20% chance of assuming the value 30. As noted in Figure 12.4, we could use the formula =RNGDiscrete({10,20,30},{0.3,0.5,0.2}) to model this random behavior. (Alternatively, if the values 10, 20, and 30 were entered in cells A1 through A3 and the values 0.3, 0.5, and 0.2 were entered in B1 through B3, we could use the formula =RNGDiscrete(A1:A3,B1:B3).) If we recalculated the

Figure 12.4
Random number functions available in the RNG.xla add-in file.

Distribution	*RNG Function*	*Description*
Binomial	RNGBinomial(n,p)	Returns the number of “successes” in a sample of size n where each trial has a probability p of “success.”
Discrete Uniform	RNGDuniform(min,max)	Returns one of the integers between min and max, inclusive. Each value is equally likely to occur.
General Discrete	RNGDiscrete({x_1,x_2,...x_n}, {p_1,p_2,...p_n})	Returns one of the *n values* represented by the x_i. The value x_i occurs with probability p_i.
Poisson	RNGPoisson(λ)	Returns a random number of events occurring per some unit of measure (for example, arrivals per hour, defects per yard, and so on). The parameter λ represents the average number of events occurring per unit of measure.
Continuous Uniform	RNGUniform(min,max)	Returns a value in the range from a minimum (*min*) to a maximum (*max*). Each value in this range is equally likely to occur.
Chi-square	RNGChisq(λ)	Returns a value from a chi-square distribution with mean λ.
Exponential	RNGExponential(λ)	Returns a value from an exponential distribution with mean λ. Often used to model the time between events or the lifetime of a device with a constant probability of failure.
Normal	RNGNormal(μ,σ)	Returns a value from a normal distribution with mean μ and standard deviation σ.
Truncated Normal	RNGTnormal(μ,σ,min,max)	Same as RNGNormal except the distribution is truncated to the range specified by a minimum (*min*) and a maximum (*max*).
Triangular	RNGTriang(min,most likely,max)	Returns a value from a triangular distribution covering the range specified by a minimum (*min*) and a maximum (*max*). The shape of the distribution is then determined by the size of the most likely value relative to *min* and *max*.

spreadsheet many times, this formula would return the value 10 approximately 30% of the time, the value 20 approximately 50% of the time, and the value 30 approximately 20% of the time.

The arguments required by the RNG functions allow us to generate random numbers from distributions with a wide variety of shapes. Figures 12.5 and 12.6 illustrate some of these distributions.

Troubleshooting

If you install the RNG.xla add-in in the directory C:\MSOffice\Excel\Library, any spreadsheet you create that uses these functions will expect to find the RNG.xla add-in in the same directory on any other computer on which this file is used. If you try to open a file that uses these functions on a computer that does not have RNG.xla installed in the same directory, Excel will display a dialog box saying,

"This document contains links. Reestablish links?"

Answer NO to this question. The file will then be opened, but the cells containing references to RNG functions will all return the "#REF!" error code. To fix this, first make sure the RNG.xla file is installed and loaded (select Tools, Add-Ins and make sure the box labeled RNG is selected.) If the problem still persists, click Edit, Links, Change Source, then locate the RNG.xla file on the computer you are using and click OK. The functions should then work correctly on that computer. Note: If you install the RNG.xla file in the same directory on every computer you use, you should never have this problem.

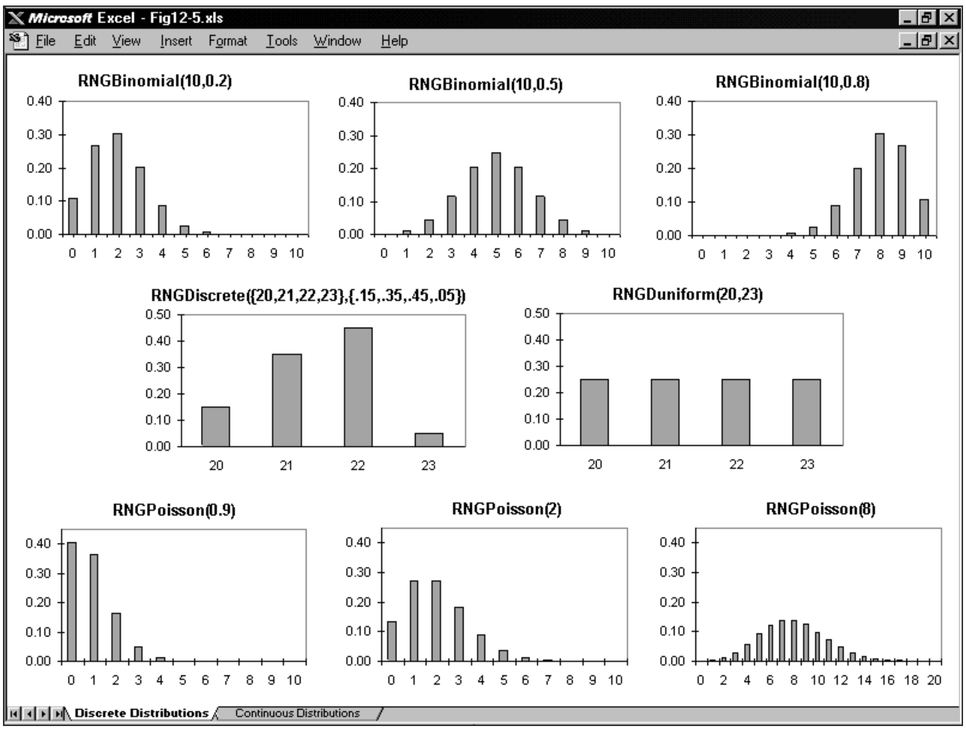


Figure 12.5 Examples of some discrete probability distributions.

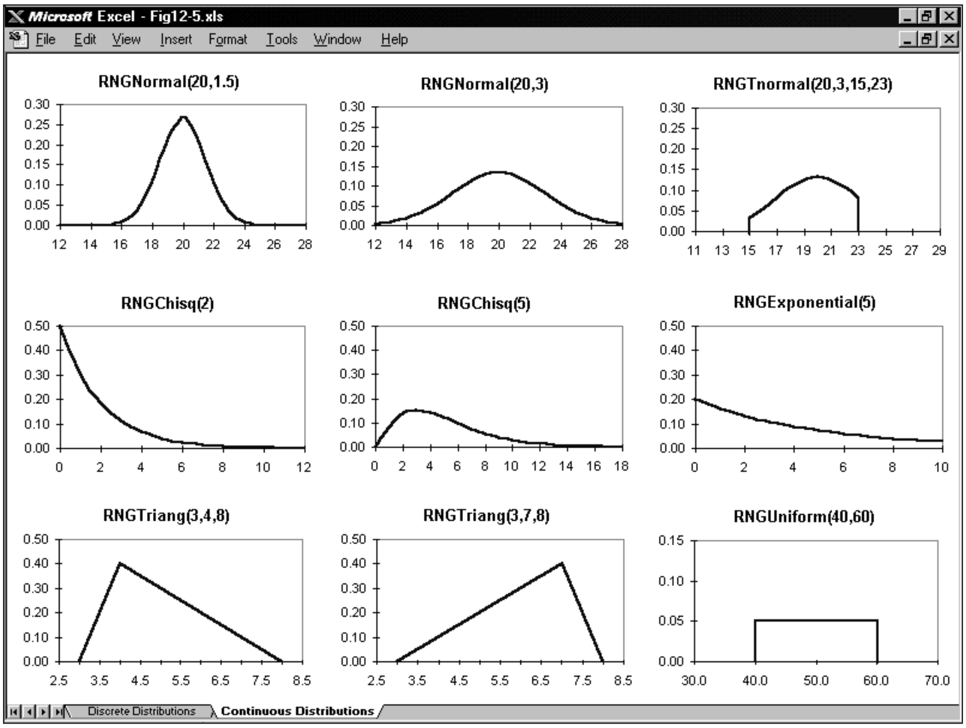


Figure 12.6
Examples of some continuous probability distributions.

12.5.3 *Discrete vs. Continuous Random Variables*

An important distinction exists between the random variables in Figure 12.5 and 12.6. In particular, the RNGs depicted in Figure 12.5 generate *discrete* outcomes, whereas those represented in Figure 12.6 generate *continuous* outcomes. The distinction between discrete and continuous random variables is very important.

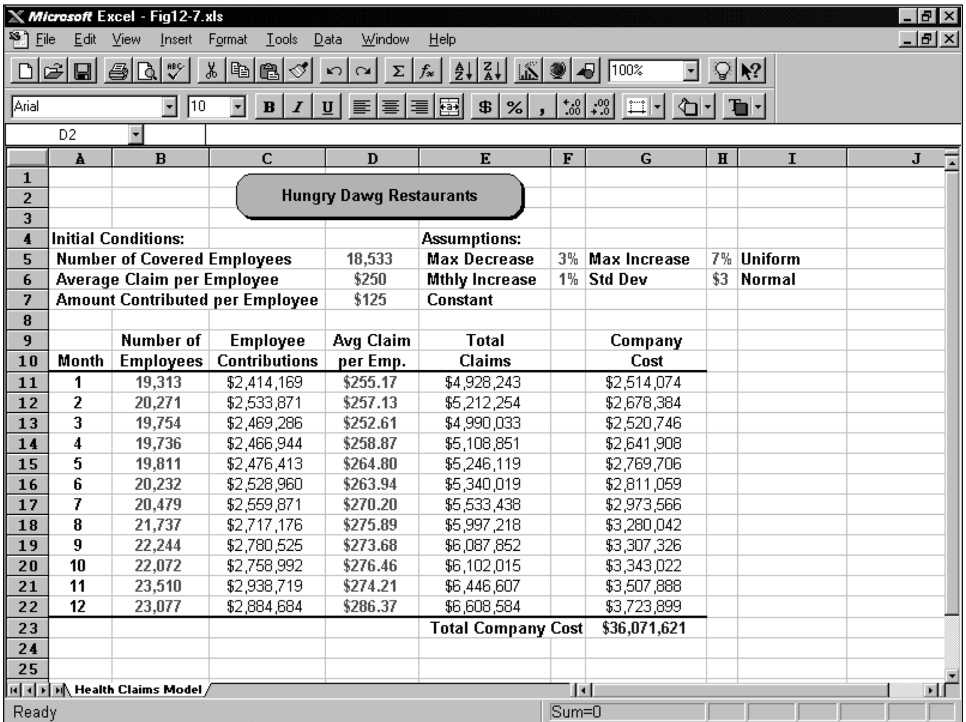
For example, the number of defective tires on a new car is a discrete random variable because it can assume only one of five distinct values: 0, 1, 2, 3, or 4. On the other hand, the amount of fuel in a new car is a continuous random variable because it can assume any value between 0 and the maximum capacity of the fuel tank. Thus, when selecting an RNG for an uncertain variable in a model, it is important to consider whether the variable can assume discrete or continuous values.

12.6 *PREPARING THE MODEL FOR SIMULATION*

To simulate the model for Hungry Dawg Restaurants described earlier, we must first select appropriate RNGs for the uncertain variables in the model. If available, historical data on the uncertain variables could be analyzed to determine appropriate RNGs for these variables. If past data are not available, or if we have reason to expect the future behavior of a variable to be significantly different from the past, then we must use judgment in selecting appropriate RNGs to model the random behavior of the uncertain variables.

For our example problem, let's assume that by analyzing historical data, we determined that the change in the number of covered employees from one month to the next is expected to vary uniformly between a 3% decrease and a 7% increase. (Note that this should cause the *average* change in the number of employees to be a 2% increase, because 0.02 is the midpoint between –0.03 and +0.07.) Further, assume that we can model the average monthly claim per covered employee as a normally distributed random variable with the mean increasing by 1% per month and a standard deviation of approximately $3. (Note that this will cause the average increase in claims per covered employee from one month to the next to be approximately 1%.) These assumptions are reflected in cells F5 through H6 at the top of Figure 12.7 (and in the file FIG12-7.xls on your data disk).

To implement the formula to generate a random number of employees covered by the health plan, we'll use the RNGUniform( ) function shown in Figure 12.4 to sample from a continuous uniform distribution. The RNGUniform( ) function generates random numbers between the minimum and maximum values that we supply.



	A	B	C	D	E	F	G	H	I
2			Hungry Dawg Restaurants						
4	Initial Conditions:				Assumptions:				
5	Number of Covered Employees			18,533	Max Decrease	3%	Max Increase	7%	Uniform
6	Average Claim per Employee			$250	Mthly Increase	1%	Std Dev	$3	Normal
7	Amount Contributed per Employee			$125	Constant				
9		Number of	Employee	Avg Claim	Total		Company		
10	Month	Employees	Contributions	per Emp.	Claims		Cost		
11	1	19,313	$2,414,169	$255.17	$4,928,243		$2,514,074		
12	2	20,271	$2,533,871	$257.13	$5,212,254		$2,678,384		
13	3	19,754	$2,469,286	$252.61	$4,990,033		$2,520,746		
14	4	19,736	$2,466,944	$258.87	$5,108,851		$2,641,908		
15	5	19,811	$2,476,413	$264.80	$5,246,119		$2,769,706		
16	6	20,232	$2,528,960	$263.94	$5,340,019		$2,811,059		
17	7	20,479	$2,559,871	$270.20	$5,533,438		$2,973,566		
18	8	21,737	$2,717,176	$275.89	$5,997,218		$3,280,042		
19	9	22,244	$2,780,525	$273.68	$6,087,852		$3,307,326		
20	10	22,072	$2,758,992	$276.46	$6,102,015		$3,343,022		
21	11	23,510	$2,938,719	$274.21	$6,446,607		$3,507,888		
22	12	23,077	$2,884,684	$286.37	$6,608,584		$3,723,899		
23					Total Company Cost		$36,071,621		

Key Cell Formulas

Cell	*Formula*	*Copied to*
B11	=D5*RNGUniform(1-F5,1+H5)	—
B12	=B11*RNGUniform(1-$F$5,1+$H$5)	B13:B22
C11	=$D$7*B11	C12:C22
D11	=RNGNormal($D$6*(1+$F$6)^A11,$H$6)	D12:D22
E11	=D11*B11	E12:E22
G11	=E11-C11	G12:G22
G23	=SUM(G11:G22)	—

Figure 12.7 Modified corporate health insurance model with RNGs replacing expected values for uncertain variables.

Important Notice

You should install the RNG.xla add-in before opening FIG12-7.xls. Also, once you open the spreadsheet in FIG12-7.xls, the numbers on your screen will not match those shown in Figure 12.7 because these numbers are generated randomly.

In our example problem, the number of employees in any given month should equal the number of employees in the previous month multiplied by a random number between 97% and 107%; that is:

Number of employees in current month = Number of employees in previous month × RNGUniform(0.97,1.07)

Notice that if the RNGUniform( ) function in this equation returns the value 0.97, this formula causes the number of employees in the current month to equal 97% of the number in the previous month (for a 3% decrease). Alternatively, if RNGUniform( ) function returns the value 1.07, this causes the number of employees in the current month to equal 107% of the number in the previous month (for a 7% increase). All the values between these two extremes (between 97% and 107%) are also possible and equally likely to occur. Thus, in Figure 12.7 the following formulas were used to randomly generate the number of employees covered by the health insurance plan each month:

Formula for cell B11: =D5*RNGUniform(1-F5,1+H5)

Formula for cell B12: =B11*RNGUniform(1-$F$5,1+$H$5)
(Copy to B13 through B22.)

To implement the formula to generate the average claims per covered employee in each month, we'll use the RNGNormal( ) function described in Figure 12.4. This formula requires that we supply the value of the mean (μ) and standard deviation (σ) of the distribution from which we want to sample. The assumed $3 standard deviation (σ) for the average monthly claim, shown in cell H6 of Figure 12.7, is constant from month to month. Thus, the only remaining problem is to figure out the proper mean value (μ) for each month.

In this case, the mean for any given month should be 1% larger than the mean in the previous month. For example, the mean for month 1 is:

$$\text{Mean in month 1} = (\text{original mean}) \times 1.01$$

and the mean for month *2* is:

$$\text{Mean in month 2} = (\text{mean in month 1}) \times 1.01$$

If we substitute the previous definition of the mean in month 1 into the above equation, we obtain

$$\text{Mean in month 2} = (\text{original mean}) \times (1.01)^2$$

Similarly, the mean in month 3 is:

$$\text{Mean in month 3} = (\text{mean month 2}) \times 1.01 = (\text{original mean}) \times (1.01)^3$$

So, in general, the mean (μ) for month *n* is:

$$\text{Mean in month } n = (\text{original mean}) \times (1.01)^n$$

Thus, to generate the average claim per covered employee in each month, we'll use the following formula:

Formula for cell D11: =RNGNormal($D$6*(1+$F$6)^A11,$H$6)
(Copy to D12 through D22.)

Note that the term "$D$6*(1+$F$6)^A11" in this formula implements the general definition of the mean (μ) in month *n*.

At this point, the modifications to the model are complete. Each time the recalculate key (the function key [F9]) is pressed, the RNGs will automatically select new values for all the cells in the spreadsheet that represent uncertain (or random) variables. With each recalculation, a new value for the bottom-line performance measure (total company cost) will appear in cell G23. By pressing the recalculate key several times, we can observe representative values of the company's total cost for health claims.

12.7 REPLICATING THE MODEL

The next step in the simulation process involves recalculating the model several hundred times and recording the resulting values generated for the output cell, or bottom-line performance measure. Suppose that we want to perform 300 replications of the model and store the resulting observations of the dependent variable on a new worksheet named Simulation. To create and name this new worksheet:

1. Click the Insert menu.
2. Select Worksheet. Excel inserts a new worksheet in your workbook.
3. Click the Format menu.
4. Click Worksheet.
5. Click Rename.
6. Type Simulation.
7. Click OK.

Because we want to perform 300 replications of our model, we entered the numbers 1, 2, 3, ..., 300 starting in cell A3, as shown in Figure 12.8. This is done as follows:

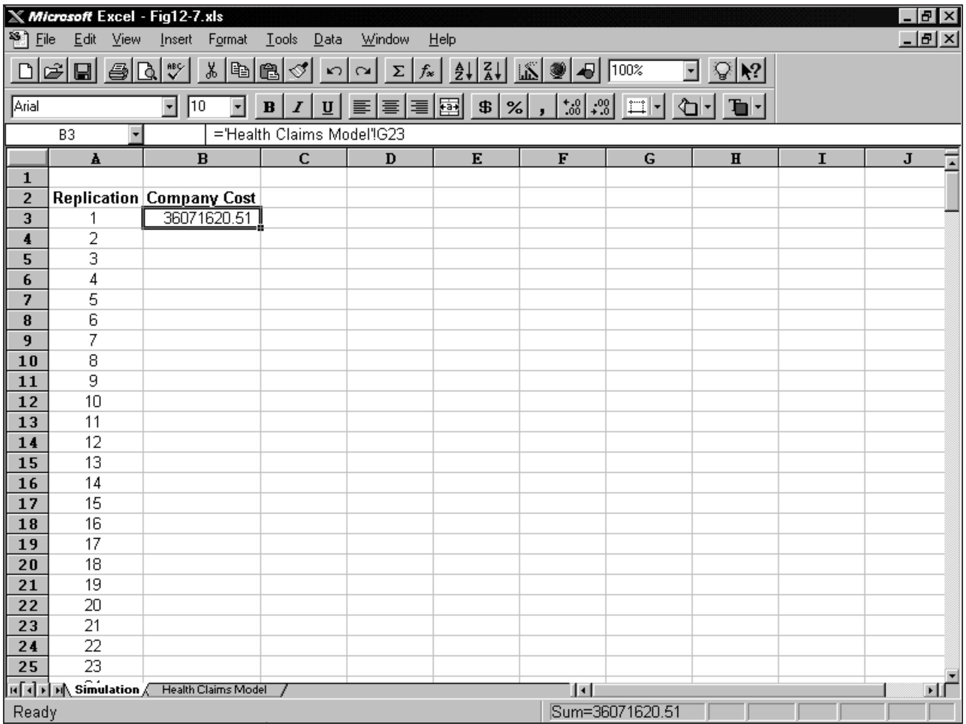
1. Type the starting value (1) in cell A3 and press [Enter].
2. Click cell A3.
3. Click the Edit menu, click Fill, then click Series.
4. Select the Series in Columns option and enter a Stop value of 300.
5. Click OK.

Excel automatically fills the column below the selected cell (A3) with values increasing by 1 (the Step value) until it reaches the Stop value of 300. Because we want to track the company cost value in cell G23 of the Health Claims Model worksheet, we entered the following formula in cell B3:

Formula for cell B3: ='Health Claims Model'!G23

Cell B2 contains the label Company Cost to identify the values that will ultimately appear in column B.

Figure 12.8 Worksheet prepared to simulate the corporate health insurance model.



Key Cell Formulas

Cell	*Formula*	*Copied to*
B3	='Health Claims Model'!G23	—

We can now use the Data Table command to fill in the remainder of column B. Keep in mind that the Data Table command is designed for a purpose other than what we are using it for here. However, we can use this command to "trick" Excel into performing the replications needed in our simulation.

By using the Data Table command, we're instructing Excel to substitute each value in column A into some cell of the spreadsheet, recalculate the spreadsheet, and record in column B the associated value for the output cell (cell B3 in this case). Ordinarily, the values listed in the first column of a data table (column A) represent values that we want Excel to enter into some input cell of the spreadsheet. The resulting data table then shows what happens to the output cell given each of the input cell values. However, for our purposes we'll instruct Excel to enter each value in column A into an input cell that has no impact on the value of the output cell. For example, we might use cell A1 in Figure 12.8 as the input cell. In this way, we can "trick" Excel into recalculating the spreadsheet 300 times while storing the values of the output cell (cell G23 on the Health Claims Model sheet) in column B.

To execute the Data Table command:

1. Select the range A3 through B302. (An easy way to do this is to select cells A3 and B3, then while pressing the shift key, double click on the selection border at the bottom of cell A3.)
2. Click the Data menu.

3. Click Table.
4. In the Table dialog box, enter cell A1 for the Column input cell.
5. Click OK.

Excel substitutes each value in the range A4 through A302 into cell A1, recalculates the workbook, and stores the resulting company cost figures in the adjacent cells in column B. Depending on your computer's speed, this recalculation might take 20 to 30 seconds, or possibly a couple of minutes.

Software Tip

The Data Table command will execute more quickly if you use a separate workbook for each model you are simulating and have only one workbook open while you are doing simulation. So create a separate workbook for each homework problem you do and make sure you close each workbook when you complete each problem. Also, before running a large number of replications, it is a good idea to first verify your model by replicating it a small number of times (say 20 to 50 times). Once you are convinced your model is working correctly, you may increase the size of the data table to the desired sample size.

After running the Data Table command, you should have a list of values in column B representing 300 possible company cost outcomes, similar to those shown in Figure 12.9. The numbers you generate on your computer will *not* match those in Figure 12.9. The procedure demonstrated here generates a *random* sample of 300 observations from an infinite number of possible values. Again, the random sample you generate will be different from the one shown in Figure 12.9, but the characteristics of your sample should be very similar to those of the sample shown in Figure 12.9.

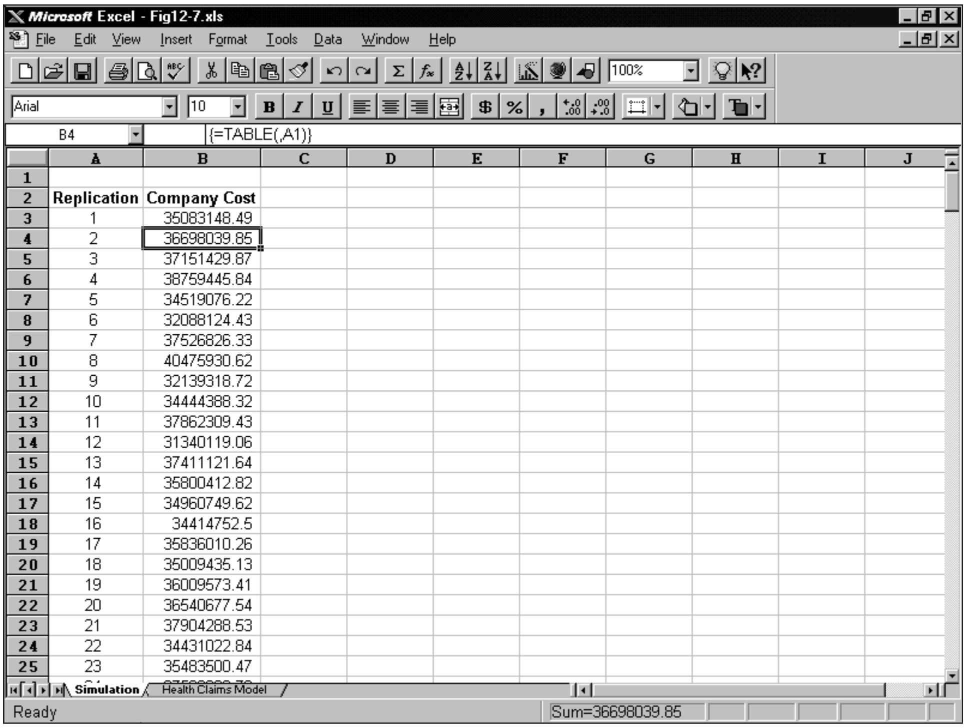
Each new entry Excel created, starting in cell B4, contains the array formula {=TABLE(,A1)}. This is how the Data Table command performs the repeated substitution and recalculation we just described. If we don't change the formulas in column B into values, every time the spreadsheet is recalculated we'll have to wait for this process to reexecute and we'll get a new sample of 300 replications. This wastes time and prevents us from focusing on the results of one batch of 300 observations in order to make decisions. To convert the formulas in column B to values:

1. Select the range B3 through B302.
2. Click the Edit menu.
3. Click Copy.
4. Click the Edit menu.
5. Click Paste Special.
6. Click Values.
7. Click OK.
8. Press [Enter].

The values in column B are now numeric constants that will not change even if the spreadsheet is recalculated.

12.7.1 Determining the Sample Size

You might wonder why we elected to perform 300 replications. Why not 200 or 800? Unfortunately, there is no easy answer to this question. Remember that the goal in



Microsoft Excel - Fig12-7.xls

B4 {=TABLE(,A1)}

	Replication	Company Cost
1		
2	Replication	Company Cost
3	1	35083148.49
4	2	36698039.85
5	3	37151429.87
6	4	38759445.84
7	5	34519076.22
8	6	32088124.43
9	7	37526826.33
10	8	40475930.62
11	9	32139318.72
12	10	34444388.32
13	11	37862309.43
14	12	31340119.06
15	13	37411121.64
16	14	35800412.82
17	15	34960749.62
18	16	34414752.5
19	17	35836010.26
20	18	35009435.13
21	19	36009573.41
22	20	36540677.54
23	21	37904288.53
24	22	34431022.84
25	23	35483500.47

Simulation / Health Claims Model

Ready Sum=36698039.85

Figure 12.9 Results of the Data Table command for the corporate health insurance problem.

Important Software Note

Instead of converting data table to values, another way to prevent the data table from recalculating is to do the following:

1. Click the Tools menu.
2. Click Options.
3. Click the Calculation tab.
4. Click Automatic Except Tables.
5. Click OK.

This tells Excel to recalculate the data tables only when you manually recalculate the spreadsheet by pressing the F9 function key. This can be helpful if you want to run several different simulations under a variety of input conditions.

simulation is to use a sample of observations on a bottom-line performance measure to estimate various characteristics about its behavior. For example, we might want to estimate the mean value of the performance measure and the shape of its probability distribution. However, we saw earlier that different values of the bottom-line performance measure occurred each time we manually recalculated the model in Figure 12.7. Thus, there is an infinite number of possibilities—or an **infinite population**—of total company cost values associated with this model.

We cannot analyze all of these infinite possibilities. But by taking a large enough sample from this infinite population, we can make reasonably accurate estimates about the characteristics of the underlying population of values. The larger the sample we take (that is, the more replications we do), the more accurate our final results will be. But performing many replications takes time and computer resources, so we must make a trade-off in terms of estimation accuracy versus convenience. There is no simple answer to the question of how many replications to perform, but as a minimum, you should always perform at least 100 replications, and more as time and resources permit or accuracy demands.

12.8 DATA ANALYSIS

As mentioned earlier, the objective of performing a simulation is to estimate various characteristics of the performance measure resulting from uncertainty in some or all of the input variables. After performing the replications, we must summarize and analyze the data in order to draw conclusions.

Most spreadsheet packages have built-in functions for performing statistical calculations. Excel also provides a data analysis tool we can use to generate numerous descriptive statistics automatically. To use the data analysis tool:

1. Click the Tools menu.
2. Click Data Analysis.
3. Click Descriptive Statistics.
4. Complete the Descriptive Statistics dialog box, as shown in Figure 12.10.
5. Click OK.

(If the Data Analysis option is not available on your Tools menu, select the Add-Ins option from the Tools menu, then select the Analysis ToolPak option. The Data Analysis option should then appear on your Tools menu. If Analysis ToolPak is not listed among your available add-ins, you must exit Excel, rerun the MSOffice setup program and add the Analysis ToolPak add-in to your installation of Excel.)

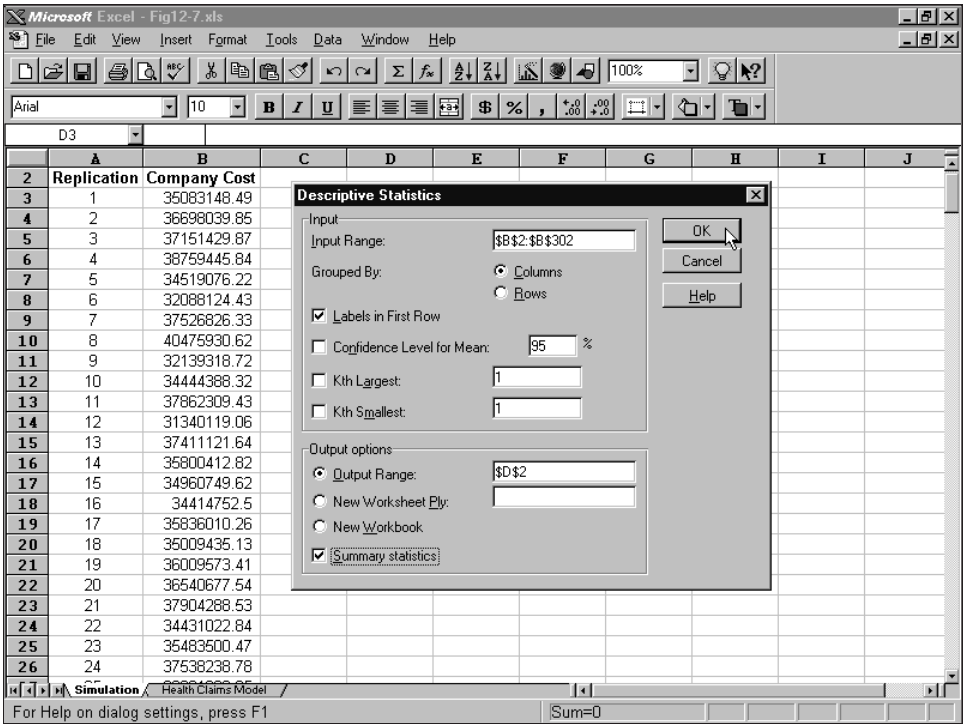


Figure 12.10 Descriptive Statistics dialog box for the corporate health insurance problem.

Figure 12.11 shows the resulting descriptive summary statistics for our sample of company cost data. We could have generated these values using a variety of Excel's built-in statistical functions. For example, we could have calculated the mean value shown in cell E4 using the formula =AVERAGE(B3:B302). However, the Descriptive Statistics command simplifies this process. Note that we can edit the results produced by this command to delete any unnecessary information.

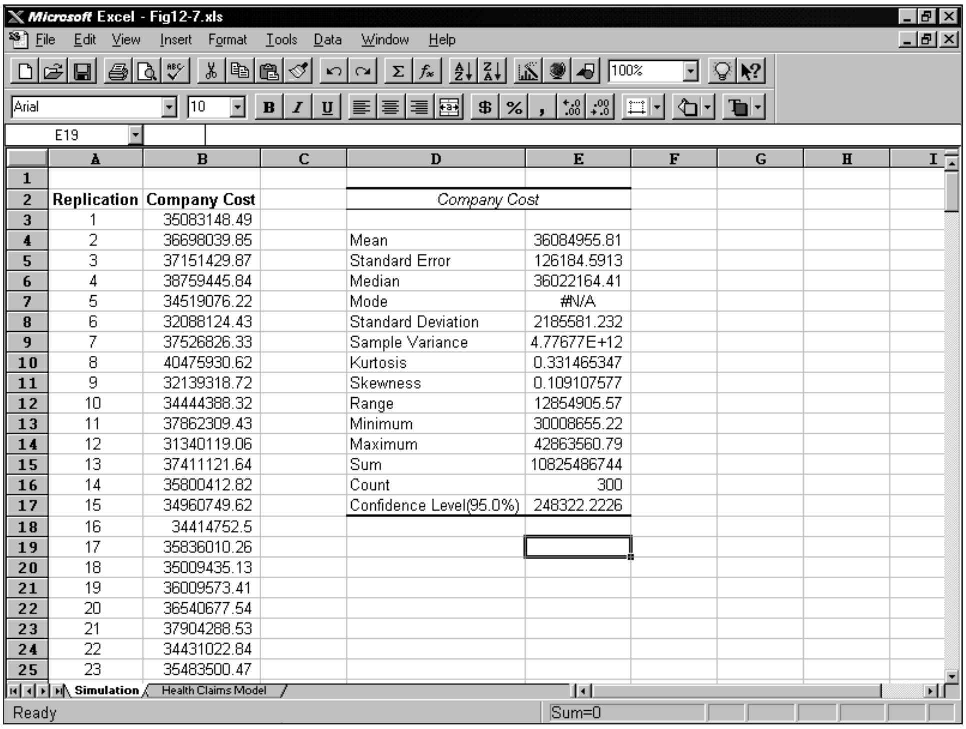
12.8.1 *The Best Case and the Worst Case*

Decision makers usually want to know the best-case and worst-case scenarios to get an idea of the range of possible outcomes they might face. This information is available from the simulation results, as shown by the Minimum and Maximum values listed in Figure 12.11. Although cell E4 indicates that the average value observed in this sample of 300 observations is approximately $36.0 million, cell E13 indicates that the smallest cost observed is about $30.0 million (representing the best case) and cell E14 indicates the largest cost is approximately $42.8 million (representing the worst case). (Note that if you generate your own sample of 300 observations, the statistics you calculate will *not* match those shown in Figure 12.11.) These figures should give the decision maker a good idea about the range of possible cost values that might occur. Note that these values might be difficult to determine manually in a complex model with many uncertain independent variables.

12.8.2 *Determining the Distribution of Outcomes*

Although the data in Figure 12.11 offer some insight into the best and worst possible outcomes of a decision, other factors should also be considered. The best- and worst-case scenarios are the most extreme outcomes, and might not be likely to occur.

Figure 12.11
Statistical results of the simulation data for the corporate health insurance model.



	A	B	C	D	E
1					
2	Replication	Company Cost		*Company Cost*	
3	1	35083148.49			
4	2	36698039.85		Mean	36084955.81
5	3	37151429.87		Standard Error	126184.5913
6	4	38759445.84		Median	36022164.41
7	5	34519076.22		Mode	#N/A
8	6	32088124.43		Standard Deviation	2185581.232
9	7	37526826.33		Sample Variance	4.77677E+12
10	8	40475930.62		Kurtosis	0.331465347
11	9	32139318.72		Skewness	0.109107577
12	10	34444388.32		Range	12854905.57
13	11	37862309.43		Minimum	30008655.22
14	12	31340119.06		Maximum	42863560.79
15	13	37411121.64		Sum	10825486744
16	14	35800412.82		Count	300
17	15	34960749.62		Confidence Level(95.0%)	248322.2226
18	16	34414752.5			
19	17	35836010.26			
20	18	35009435.13			
21	19	36009573.41			
22	20	36540677.54			
23	21	37904288.53			
24	22	34431022.84			
25	23	35483500.47			

Determining the likelihood of these outcomes requires that we know something about the shape of the distribution of our bottom-line performance measure. Thus, we might also want to construct a frequency distribution and histogram for the 300 observations generated for our performance measure. To construct a frequency distribution and histogram:

1. Click the Tools menu.
2. Click Data Analysis.
3. Click Histogram.
4. Complete the Histogram dialog box, as shown in Figure 12.12.
5. Click OK.

The resulting new worksheet, named Histogram, contains a frequency distribution and histogram of our data which, after some simple formatting, appear as shown in Figure 12.13.

The Frequency column in Figure 12.13 indicates the number of observations from our simulation results that fall into the bins defined in column A. For example, the value in cell B2 indicates that one of the 300 replications resulted in a value that is less than or equal to $30,008,655. The value in cell B3 indicates that two of the 300 observations assumed values greater than $30,008,655 and less than or equal to $30,764,826. Similarly, cell B18 indicates that three observations have values between $41,351,219 and $42,107,390, and cell B19 indicates that two observations were greater than $42,107,390. In cell B9, we see that the most frequently occurring values are in the range $34,545,681 to $35,301,852. (Again, your results will *not* match those shown in Figure 12.13.)

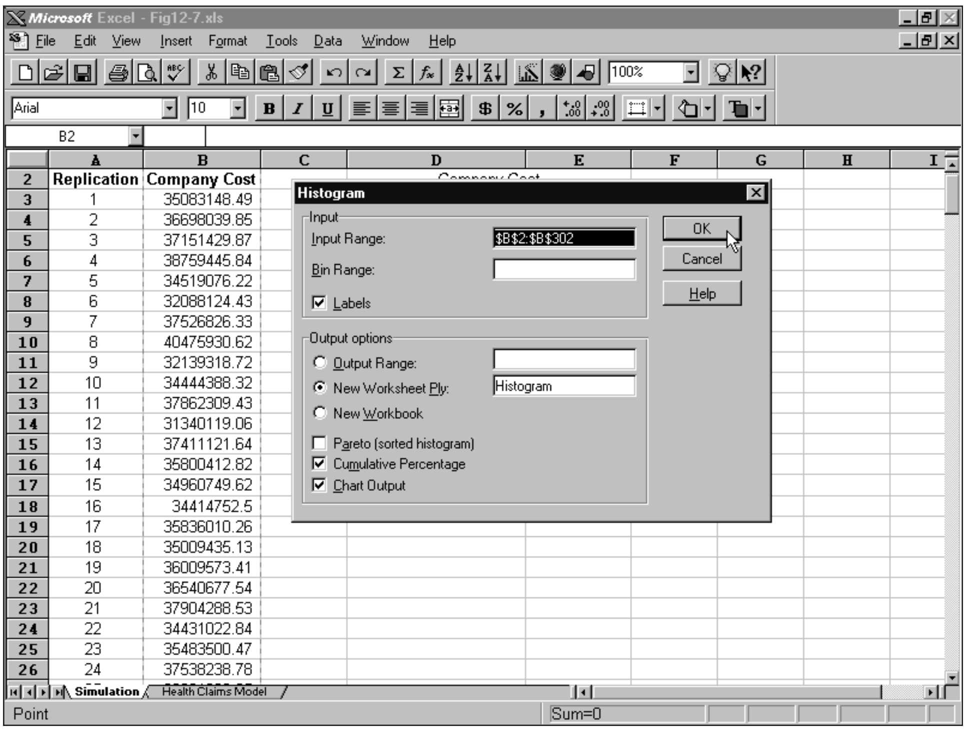
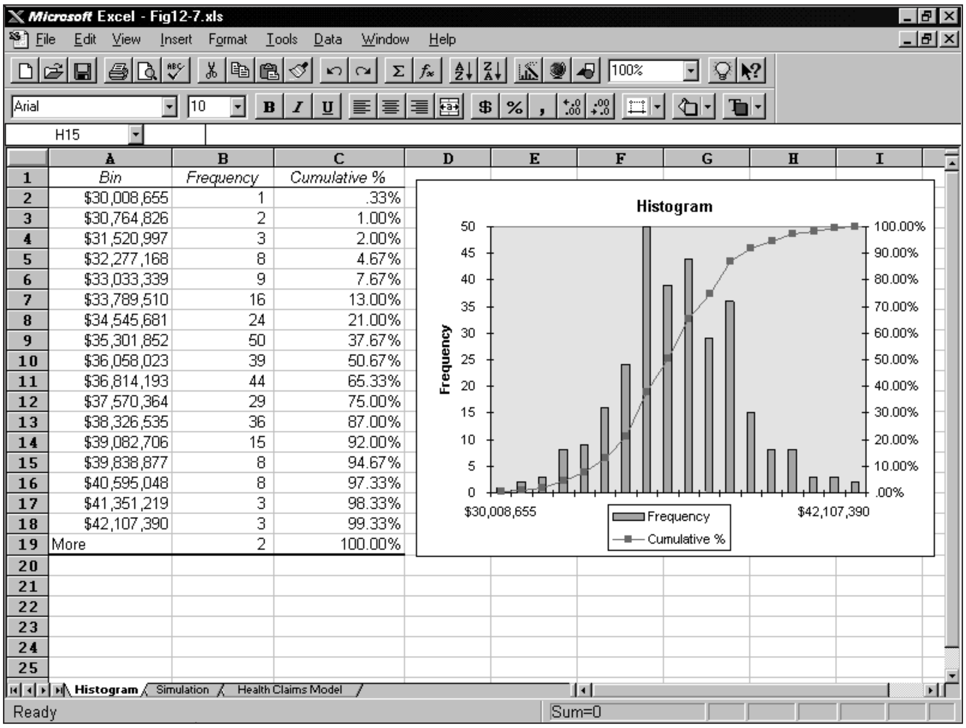


Figure 12.12 Histogram dialog box.

Figure 12.13
Frequency distribution and Histogram for the corporate health insurance model.



Bin	Frequency	Cumulative %
$30,008,655	1	.33%
$30,764,826	2	1.00%
$31,520,997	3	2.00%
$32,277,168	8	4.67%
$33,033,339	9	7.67%
$33,789,510	16	13.00%
$34,545,681	24	21.00%
$35,301,852	50	37.67%
$36,058,023	39	50.67%
$36,814,193	44	65.33%
$37,570,364	29	75.00%
$38,326,535	36	87.00%
$39,082,706	15	92.00%
$39,838,877	8	94.67%
$40,595,048	8	97.33%
$41,351,219	3	98.33%
$42,107,390	3	99.33%
More	2	100.00%

The histogram in Figure 12.13 gives a visual representation of the frequency distribution values. This graph shows that the distribution of the total health claims the company must pay is somewhat bell-shaped.

12.8.3 The Cumulative Distribution of the Output Cell

The Cumulative % column in Figure 12.13 shows the percentage of observations in the sample that are less than or equal to the values listed in column A. For example, cell C2 indicates that 0.33% of the 300 observations are less than $30,008,655, cell C3 indicates that 1% of the 300 observations are less than or equal to $30,764,826, and so on. These cumulative frequencies are also plotted on the graph shown in Figure 12.13.

Cumulative frequencies are helpful in answering a number of questions that might arise. For example, suppose that the chief financial officer (CFO) for Hungry Dawg Restaurants would rather accrue an excess of money to pay health claims than not accrue enough money. The CFO might want to know what amount the company should accrue so that there is only about a 10% chance of coming up short of funds at the end of the year. The value in cell C13 of Figure 12.13 indicates that 87% of the 300 observations are less than or equal to $38,326,535 and the value in cell C14 indicates that 92% of the observations are less than or equal to $39,082,706. Thus, assuming that our sample is representative of the actual distribution of total health costs the company might incur, we could tell the CFO that if Hungry Dawg budgets $39 million for health claims, there is roughly a 10% chance of the company not accruing enough funds.

Another way of answering the CFO's question is to sort the 300 observations in our sample in ascending order. (This can be done easily by selecting the range B3 through B302 in the Simulation worksheet and clicking the Ascending Sort button [A to Z] on the toolbar.) The 270th largest number represents the 90th percentile of the distribution because 90% of the values in the sample are less than this value (and only 10% are greater than this value). For the 300 values represented in Figure 12.13, the 270th largest number in the sample is $38,747,460, which is fairly close to the recommendation of $39 million suggested by our previous analysis.

One final point underscores the value of simulation. How could we answer the CFO's question using best-case/worst-case analysis or what-if analysis? The fact is, we could not answer the question with any degree of accuracy without using simulation to obtain the cumulative frequencies shown in Figure 12.13.

Software Tip

The **COUNTIF( )**function is often very useful in estimating probabilities from simulation results. For example, the proportion of sample observations in Figure 12.13 which are less than $39 million can be computed as =COUNTIF(B3:B302,"<$39,000,000")/300.

12.9 THE UNCERTAINTY OF SAMPLING

To this point, we have used simulation to generate 300 observations on our bottom-line performance measure and then calculated various statistics to describe the characteristics and behavior of the performance measure. For example, Figure 12.11 indicates that the mean of our sample is approximately $36.0 million. Using the results in Figure 12.13, we estimate that approximately a 90% chance exists for the performance measure assuming a value less than $39 million. But what if we repeat this process and generate another 300 observations? Would the sample mean for the new 300 observations also be $36.0 million? Or would exactly 90% of the observations in the new sample be less than $39 million?

The answer to both of these questions is "probably not." The sample of 300 observations used in our analysis was taken from a population of values that is theoretically infinite in size. That is, if we had enough time and our computer had enough memory, we could generate an infinite number of values for our bottom-line performance measure. Theoretically, we could then analyze this infinite population of values to determine its true mean value, its true standard deviation, and the true probability of the performance measure being less than $39 million. Unfortunately, we do not have the time or computer resources to determine these true characteristics (or parameters) of the population. The best we can do is take a sample from this population and, based on our sample, make estimates about the true characteristics of the underlying population. Our estimates will differ depending on the sample we choose and the size of the sample.

The mean of the sample we take is probably not equal to the true mean we would observe if we could analyze the entire population of values for our performance measure. The sample mean we calculate is just an estimate of the true population mean. In our example problem, we estimated that a 90% chance exists for our output variable to assume a value less than $39 million. However, this most likely is not equal to the true probability we would calculate if we could analyze the entire population. Thus, there is some element of uncertainty surrounding the statistical estimates resulting from simulation because we are using a sample to make inferences about the population. Fortunately, there are ways of measuring and describing the amount of uncertainty present in some of the estimates we make about the population under study. This is typically done by constructing confidence intervals for the population parameters being estimated.

12.9.1 Constructing a Confidence Interval for the True Population Mean

Constructing a confidence interval for the true population mean is a simple process. If $\bar{y}$ and s represent, respectively, the mean and standard deviation of a sample size n from any population, then assuming n is sufficiently large ($n \geq 30$), the Central Limit Theorem tells us that the lower and upper limits of a 95% confidence interval for the true mean of the population are represented by:

$$\text{95\% Lower Confidence Limit} = \bar{y} - 1.96 \times \frac{s}{\sqrt{n}}$$

$$\text{95\% Upper Confidence Limit} = \bar{y} + 1.96 \times \frac{s}{\sqrt{n}}$$

Although we can be fairly certain that the sample mean $\bar{y}$ we calculate from our sample data is not equal to the true population mean, we can be 95% confident that the true mean of the population falls somewhere between the lower and upper limits given above. If we want a 90% or 99% confidence interval, we must change the value 1.96 in the above equations to 1.645 or 2.575, respectively. (The values 1.645 and 2.575 represent the 95 and 99.5 percentiles of the standard normal distribution, respectively.)

For our example, the lower and upper limits of a 95% confidence interval for the true mean of the population of total company cost values can be calculated easily, as shown in cells E20 and E21 in Figure 12.14. The formulas for these cells are:

Formula for cell E20: =E4–1.96*E8/SQRT(E16)

Formula for cell E21: =E4+1.96*E8/SQRT(E16)

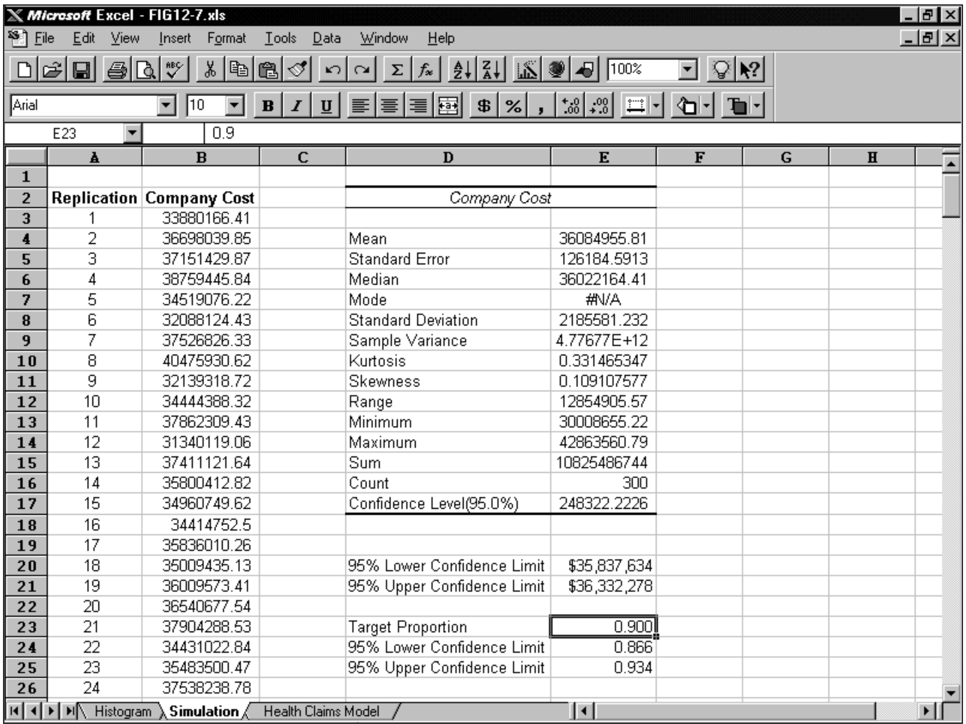
Thus, we can be 95% confident that the true mean of the population of total company cost values falls somewhere in the interval from $35,837,634 to $36,332,278.

Software Tip

In Figure 12.14, the value in cell E17 labeled Confidence Level (95%) corresponds to the half-width of a 95% confidence interval for the true population mean (that is, E17 ≈ 1.96*E8/SQRT(E16)).

12.9.2 Constructing a Confidence Interval for a Population Proportion

In our example we estimated that 90% of the population of total company cost values fall below $39 million based on our sample of 300 observations. However, if we could evaluate the entire population of total cost values, we might find that only 80% of these values fall below $39 million. Or we might find that 99% of the entire population fall below this mark. It would be helpful to determine how accurate the 90%



	A	B	C	D	E
1					
2	Replication	Company Cost		*Company Cost*	
3	1	33880166.41			
4	2	36698039.85		Mean	36084955.81
5	3	37151429.87		Standard Error	126184.5913
6	4	38759445.84		Median	36022164.41
7	5	34519076.22		Mode	#N/A
8	6	32088124.43		Standard Deviation	2185581.232
9	7	37526826.33		Sample Variance	4.77677E+12
10	8	40475930.62		Kurtosis	0.331465347
11	9	32139318.72		Skewness	0.109107577
12	10	34444388.32		Range	12854905.57
13	11	37862309.43		Minimum	30008655.22
14	12	31340119.06		Maximum	42863560.79
15	13	37411121.64		Sum	10825486744
16	14	35800412.82		Count	300
17	15	34960749.62		Confidence Level(95.0%)	248322.2226
18	16	34414752.5			
19	17	35836010.26			
20	18	35009435.13		95% Lower Confidence Limit	$35,837,634
21	19	36009573.41		95% Upper Confidence Limit	$36,332,278
22	20	36540677.54			
23	21	37904288.53		Target Proportion	0.900
24	22	34431022.84		95% Lower Confidence Limit	0.866
25	23	35483500.47		95% Upper Confidence Limit	0.934
26	24	37538238.78			

Figure 12.14 Confidence intervals for the population mean and population proportion.

Key Cell Formulas

Cell	*Formula*	*Copied to*
E20	=E4-1.96*E8/SQRT(E16)	—
E21	=E4+1.96*E8/SQRT(E16)	—
E24	=E23-1.96*SQRT(E23*(1-E23)/E16)	—
E25	=E23+1.96*SQRT(E23*(1-E23)/E16)	—

value is. So, at times we might want to construct a confidence interval for the true proportion of a population that falls below (or above) some value, say Y_p.

To see how this is done, let $\overline{p}$ denote the proportion of observations in a sample of size n that falls below some value Y_p. Assuming that n is sufficiently large ($n \geq 30$), the Central Limit Theorem tells us that the lower and upper limits of a 95% confidence interval for the true proportion of the population falling below Y_p are represented by:

$$\text{95\% Lower Confidence Limit} = \overline{p} - 1.96 \times \sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$$

$$\text{95\% Upper Confidence Limit} = \overline{p} + 1.96 \times \sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$$

Although we can be fairly certain that the proportion of observations falling below Y_p in our sample is not equal to the true proportion of the population falling below Y_p, we can be 95% confident that the true proportion of the population falling below Y_p is contained within the lower and upper limits given by the previous equations. Again, if we want a 90% or 99% confidence interval, we must change the value 1.96 in the above equation to 1.645 or 2.575, respectively.

Using these formulas, we can calculate the lower and upper limits of a 95% confidence interval for the true proportion of the population falling below $39 million. From our simulation results, we know that approximately 90% of the observations in our sample are less than $39 million. Thus, our estimated value of $\overline{p}$ is 0.90. This value was entered into cell E23 in Figure 12.14. The upper and lower limits of a 95% confidence interval for the true proportion of the population falling below $39 million are calculated in cells E24 and E25 of Figure 12.14 using the following formulas:

Formula for cell E24: =E23–1.96*SQRT(E23*(1–E23)/E16)

Formula for cell E25: =E23+1.96*SQRT(E23*(1–E23)/E16)

We can be 95% confident that the true proportion of the population of total cost values falling below $39 million is between 0.866 and 0.934. Because this interval is fairly tight around the value 0.90, we can be reasonably certain that the $39 million figure quoted to the CFO has approximately only a 10% chance of being exceeded.

12.9.3 Sample Sizes and Confidence Interval Widths

The formulas for the confidence intervals in the previous section depend directly on the number of replications (n) in the simulation. As the number of replications (n) increases, the width of the confidence interval decreases (or becomes more precise). Thus, for a given level of confidence (for example, 95%), the only way to make the upper and lower limits of the interval closer together (or tighter) is to make n larger—that is, use a larger sample size. A larger sample should provide more information about the population and, therefore, allow us to be more accurate in estimating the true parameters of the population.

12.10 THE BENEFITS OF SIMULATION

What have we accomplished through simulation? Are we really better off than if we had just used the results of the original model proposed in Figure 12.2? The estimated value for the expected total cost to the company in Figure 12.2 is comparable to that obtained through simulation (although this will not always be the case). But remember that the goal of modeling is to give us greater insight into a problem to help us make more informed decisions.

The results of our simulation analysis do give us greater insight into the example problem. In particular, we now have some idea of the best- and worst-case total cost outcomes for the company. We have a better idea of the distribution and variability of the possible outcomes and a more precise idea about where the mean of the distribution is located. We also now have a way of determining how likely it is for the actual outcome to fall above or below some value. Thus, in addition to our greater insight and understanding of the problem, we also have solid empirical evidence (the facts and figures) to support our recommendations.

12.11 ADDITIONAL USES OF SIMULATION

Earlier we indicated that simulation is a technique that *describes* the behavior or characteristics of a bottom-line performance measure. The next two examples show how describing the behavior of a performance measure gives a manager a useful tool in determining the optimal value for one or more controllable parameters in a decision problem. These examples reinforce the mechanics of using simulation and also demonstrate some additional simulation modeling techniques.

12.12 AN INVENTORY CONTROL EXAMPLE

According to *The Wall Street Journal* (7/18/94), U.S. businesses had a combined inventory worth $884.77 billion dollars as of the end of May 1994. Because so much money is tied up in inventories, businesses face many important decisions regarding the management of these assets. Frequently asked questions regarding inventory include:

- What's the best level of inventory for a business to maintain?
- When should goods be reordered (or manufactured)?
- How much safety stock should be held in inventory?

The study of inventory control principles is split into two distinct areas—one assumes that demand is known (or deterministic), and the other assumes that demand is random (or stochastic). If demand is known, various formulas can be derived that provide answers to the previous questions. (An example of one such formula is given in the discussion of the EOQ model in Chapter 8.) However, when demand for a product is uncertain or random, answers to the previous questions cannot be expressed in terms of a simple formula. In these situations, the technique of simulation proves to be a useful tool, as illustrated in the following example.

Laura Tanner is the owner of Computers of Tomorrow (COT), a retail computer store in Austin, Texas. Competition in retail computer sales is fierce—both in terms of price and service. Laura is concerned about the number of stockouts occurring on a popular type of computer monitor. This monitor is priced competitively and generates a marginal profit of $45 per unit sold. Stockouts are very costly to the business because when customers cannot buy this item at COT, they simply buy it from a competing store and COT loses the sale (there are no back-orders). Laura measures the effects of stockouts on her business in terms of service level, or the percentage of total demand that can be satisfied from inventory.

Laura has been following the policy of ordering 50 monitors whenever her daily ending inventory position (defined as ending inventory on hand plus outstanding orders) falls below her reorder point of 28 units. Laura places the order at the beginning of the next day. Orders are delivered at the beginning of the day and, therefore, can be used to satisfy demand on that day. For example, if the ending inventory position on day 2 is less than 28, Laura places the order at the beginning of day 3. If the actual time between order and delivery, or lead time, turns out to be four days, then the order arrives at the start of day 7. The current level of on-hand inventory is 50 units and no orders are pending.

COT sells an average of six monitors per day. However, the actual number sold on any given day can vary. By reviewing her sales records for the past several months, Laura determined that the daily demand for this monitor is a random variable that can be described by the following probability distribution:

Units Demanded:	0	1	2	3	4	5	6	7	8	9	10
Probability:	0.01	0.02	0.04	0.06	0.09	0.14	0.18	0.22	0.16	0.06	0.02

The manufacturer of this computer monitor is located in California. Although it takes an average of four days for COT to receive an order from this company, Laura has determined that the lead time of a shipment of monitors is also a random variable that can be described by the following probability distribution:

Lead Time (days):	3	4	5
Probability:	0.2	0.6	0.2

One way to guard against stockouts and improve the service level is to increase the reorder point for the item so that more inventory is on hand to meet the demand occurring during the lead time. Laura wants to determine the reorder point that results in an average service level of 99%.

12.12.1 Creating the RNGs

To solve this problem, we need to build a model to represent the inventory of computer monitors during an average month of 30 days. This model must account for the random daily demands that can occur and the random lead times encountered when orders are placed. Both variables are examples of *general, discrete* random variables because the possible outcomes they assume consist solely of integers, and the probabilities associated with each outcome are not equal (or not uniform). Thus, we will use the RNGDiscrete( ) function described in Figure 12.4 to model these variables.

To ensure that we understand what the RNGDiscrete( ) functions does, let's consider how we'd simulate the random order lead times in this problem using the RAND( ) function. Here, we need an RNG that returns the value 3 with probability 0.2, the value 4 with probability 0.6, and the value 5 with probability 0.2. Recall that RAND( ) returns a uniformly distributed random number between 0 and 1. If we subdivide the interval from 0 to 1 into three mutually exclusive and exhaustive pieces with widths corresponding to the probabilities associated with each possible lead time, we get:

Lower Limit	*Upper Limit*	*Lead Time*
0.0	$0.19\overline{9}$	3
0.2	$0.79\overline{9}$	4
0.8	$0.99\overline{9}$	5

The numbers generated by the RAND( ) function fall in the first interval (from 0 to $0.19\overline{9}$) approximately 20% of the time, the second interval (from 0.2 to $0.79\overline{9}$) approximately 60% of the time, and the third interval (from 0.8 to $0.99\overline{9}$) approximately 20% of the time. The width of each of these intervals corresponds directly to the desired probability of each lead time value. So if we associate each interval with the indicated lead times, a lead time of three days has a 20% chance of occurring, a four-day lead time has a 60% chance of occurring, and a five-day lead time has a 20% chance of occurring. (We could use similar logic to break the interval from 0 to 1 into 11 mutually exclusive and exhaustive intervals to correspond to the different random demands that might occur.) The RNGDiscrete( ) function uses this same logic to generate general discrete random numbers.

The data describing the distributions of both of the random variables in this problem are entered in Excel as shown in Figure 12.15 (and in the file FIG12-15.xls on your data disk).

Given the data in Figure 12.15, we can use the following formulas to generate random order lead times and random daily demands that follow the appropriate probability distributions:

RNG for order lead time:	=RNGDiscrete(B6:B8,C6:C8)
RNG for daily demand:	=RNGDiscrete(E6:E16,F6:F16)

12.12.2 Implementing the Model

Now that we have a way of generating the random numbers needed in this problem, we can consider how the model should be built. As shown in Figure 12.16, we begin by creating a worksheet that lists the basic parameters for the model (or variables that are under the decision maker's control).

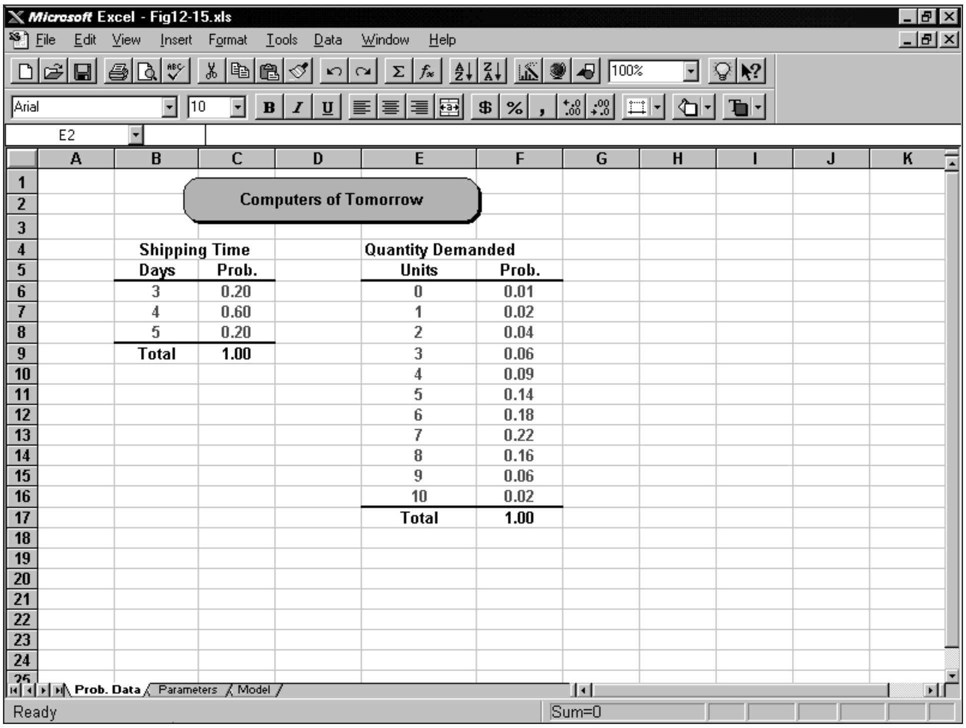


Figure 12.15 Probability distributions for shipping times and demand for COT's inventory problem.

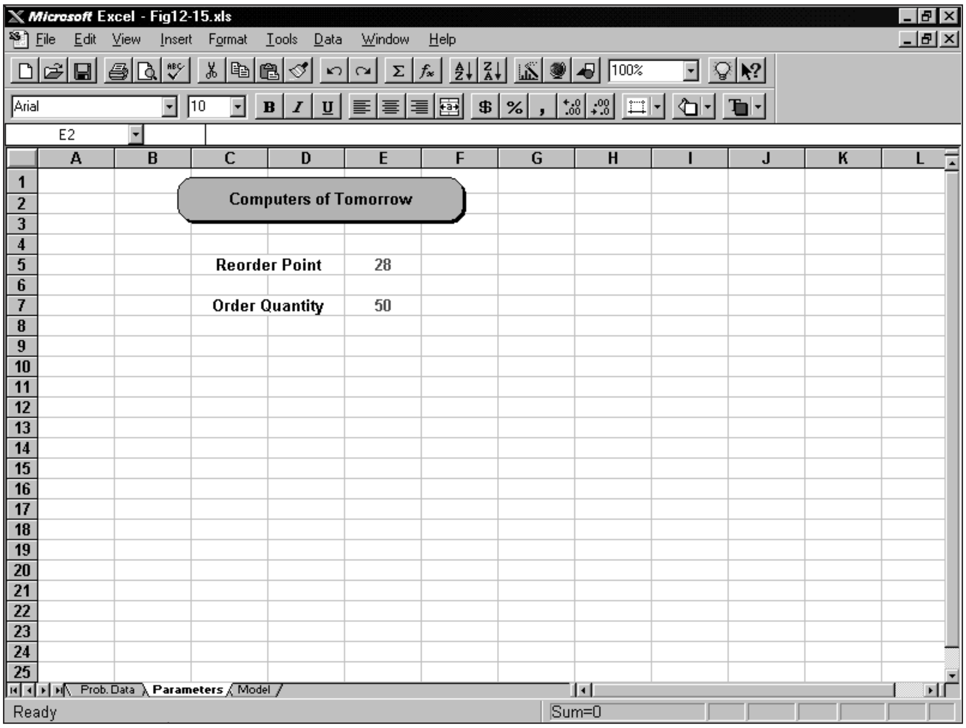


Figure 12.16 Parameters worksheet for COT's inventory problem.

Laura wants to determine the reorder point that results in an average service level of 99%. Cell E5 in Figure 12.16 is used to represent the reorder point. The order quantity for the problem is given in cell E7 so that Laura could also use this model to investigate the impact of changes in order quantity.

Figure 12.17 shows the model representing 30 days of inventory activity. In this spreadsheet, column B represents the inventory on hand at the beginning of each day, which is simply the ending inventory from the previous day. The formulas in column B are:

Formula in cell B6: =50

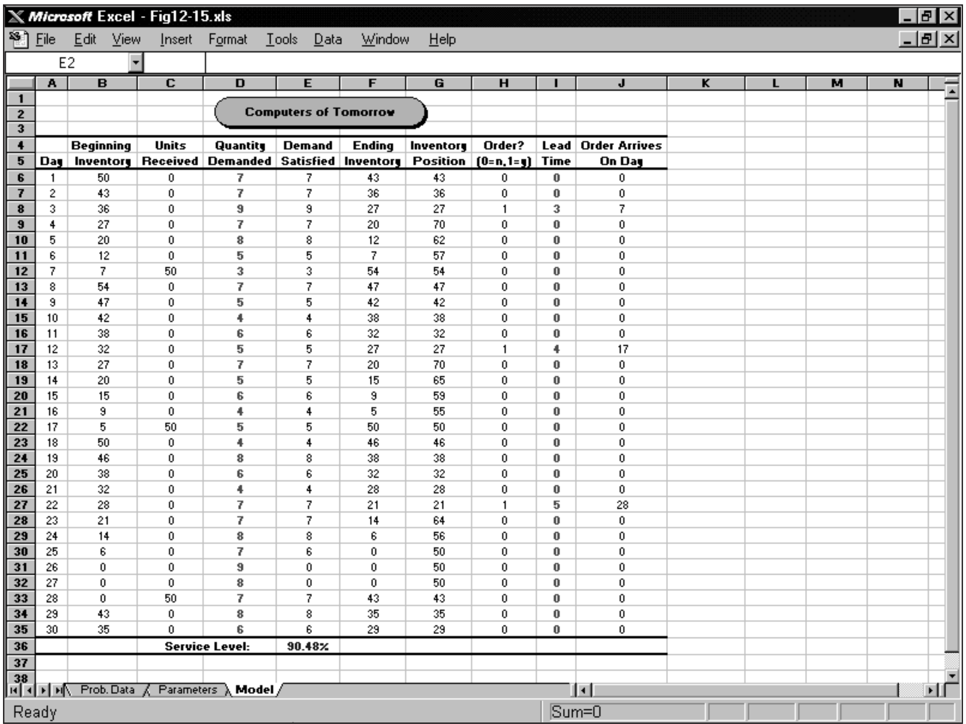
Formula in cell B7: =F6
(Copy to B8 through B35.)

Column C represents the number of units scheduled to be received each day. We'll discuss the formulas in column C after we discuss columns H, I, and J, which relate to ordering and order lead times.

In column D, we use the technique described earlier to generate random daily demands, as:

Formula for cell D6:=RNGDiscrete('Prob. Data'!$E$6:$E$16,'Prob. Data'!$F$6:$F$16)
(Copy to D7 through D35.)

Because it is possible for demand to exceed the available supply, column E indicates how much of the daily demand can be met. If the beginning inventory (in column B) plus the ordered units received (in column C) is greater than or equal to the actual demand, then all the demand can be satisfied; otherwise, COT can sell only as many units as are available. This condition is modeled as:



Day	Beginning Inventory	Units Received	Quantity Demanded	Demand Satisfied	Ending Inventory	Inventory Position	Order? (0=n,1=y)	Lead Time	Order Arrives On Day
1	50	0	7	7	43	43	0	0	0
2	43	0	7	7	36	36	0	0	0
3	36	0	9	9	27	27	1	3	7
4	27	0	7	7	20	70	0	0	0
5	20	0	8	8	12	62	0	0	0
6	12	0	5	5	7	57	0	0	0
7	7	50	3	3	54	54	0	0	0
8	54	0	7	7	47	47	0	0	0
9	47	0	5	5	42	42	0	0	0
10	42	0	4	4	38	38	0	0	0
11	38	0	6	6	32	32	0	0	0
12	32	0	5	5	27	27	1	4	17
13	27	0	7	7	20	70	0	0	0
14	20	0	5	5	15	65	0	0	0
15	15	0	6	6	9	59	0	0	0
16	9	0	4	4	5	55	0	0	0
17	5	50	5	5	50	50	0	0	0
18	50	0	4	4	46	46	0	0	0
19	46	0	8	8	38	38	0	0	0
20	38	0	6	6	32	32	0	0	0
21	32	0	4	4	28	28	0	0	0
22	28	0	7	7	21	21	1	5	28
23	21	0	7	7	14	64	0	0	0
24	14	0	8	8	6	56	0	0	0
25	6	0	7	6	0	50	0	0	0
26	0	0	9	0	0	50	0	0	0
27	0	0	8	0	0	50	0	0	0
28	0	50	7	7	43	43	0	0	0
29	43	0	8	8	35	35	0	0	0
30	35	0	6	6	29	29	0	0	0
		Service Level:		90.48%					

Figure 12.17 Spreadsheet representing a random month of inventory data.

Key Cell Formulas

Cell	*Formula*	*Copied to*
B7	=F6	B8:B35
C7	=COUNTIF($J$6:J6,A7)*Parameters!$E$7	C8:C35
D6	=RNGDiscrete('Prob. Data'!$E$6:$E$16,'Prob. Data'!$F$6:$F$16)	D7:D35
E6	=MIN(D6,B6+C6)	E7:E35
E36	=SUM(E6:E35)/SUM(D6:D35)	—
F6	=B6+C6-E6	F7:F35
G6	=F6	—
G7	=G6-E7+IF(H6=1,Parameters!$E$7,0)	G8:G35
H6	=IF(G6<Parameters!$E$5,1,0)	H7:H35
I6	=IF(H6=0,0,RNGDiscrete('Prob. Data'!$B$6:$B$8,'Prob. Data'!$C$6:$C$8))	I7:I35
J6	=IF(I6=0,0,A6+1+I6)	J7:J35

Formula for cell E6: =MIN(D6,B6+C6)
(Copy to E7 through E35.)

The values in column F represent the on-hand inventory at the end of each day, and are calculated as:

Formula for cell F6: =B6+C6-E6
(Copy to F7 through F35.)

To determine whether or not to place an order, we first must calculate the inventory position, which was defined earlier as the ending inventory plus any outstanding orders. This is implemented in column G as follows:

Formula for cell G6: =F6
Formula for cell G7: =G6-E7+IF(H6=1,Parameters!$E$7,0)
(Copy to G8 through G35.)

Column H indicates if an order should be placed based on inventory position and the reorder point, as:

Formula for cell H6: =IF(G6<Parameters!$E$5,1,0)
(Copy to H7 through H35.)

If an order is placed, we must generate the random lead time required to receive the order. This is done in column I as:

Formula for cell I6:=IF(H6=0,0,RNGDiscrete('Prob. Data'!$B$6:$B$8,'Prob. Data'!$C$6:$C$8))
(Copy to I7 through I35.)

This formula returns the value 0 if no order was placed (if H6 = 0); otherwise, it returns a random lead time value (if H6 = 1).

If an order is placed, column J indicates the day on which the order will be received based on its random lead time in column I. This is done as:

Formula for cell J6: =IF(I6=0,0,A6+1+I6)
(Copy to J7 through J35.)

The values in column C are coordinated with those in column J. The nonzero values in column J indicate the days on which orders will be received. For example, cell J8 indicates that an order will be received on day 7. The actual receipt of this order is reflected by the value of 50 in cell C12, which represents the receipt of an order at the beginning of day 7. The formula in cell C12 that achieves this is:

Formula for cell C12: =COUNTIF($J$6:J11,A12)*Parameters!$E$7

This formula counts how many times the value in cell A12 (representing day 7) appears as a scheduled receipt day between days 1 through 6 in column J. This represents the number of orders scheduled to be received on day 7. We then multiply this by the order quantity (50), given in cell E7 on the Parameters worksheet, to determine the total units to be received on day 7. So the values in column C are generated as:

Formula for cell C6: =0
Formula for cell C7: =COUNTIF($J$6:J6,A7)*Parameters!$E$7
(Copy to C8 through C35.)

The service level for the model is calculated in cell E36 using the values in columns D and E as:

Formula for cell E36: =SUM(E6:E35)/SUM(D6:D35)

Again, the service level represents the proportion of total demand that can be satisfied from inventory.

12.12.3 Replicating the Model

The model in Figure 12.17 indicates one possible scenario that could occur if Laura uses a reorder point of 28 units for the computer monitor. In the scenario shown, the value in cell E36 indicates that 90.48% of the total demand is satisfied. By replicating this model over and over, Laura could keep track of the service level occurring in each replication and average these values to determine the expected service level obtained with a reorder point of 28.

Using the simulation techniques described earlier, Laura could repeat this process with reorder points of 30 units, 32 units, and so on, until she finds the reorder point that achieved her goal of an average service level of 99%. However, there is another, easier way that Laura can perform the replications, as shown in Figure 12.18.

We prepared the spreadsheet in Figure 12.18 to replicate our model using a two-input (or two-way) data table. In cell B11 we entered the value 1 and used the Edit, Fill, Series command to fill the remainder of this column with the values 2, 3, and so on, up to 200. Cells C10 through G10 contain the values 28, 30, 32, 34, and 36, respectively, to represent a variety of reorder points that Laura might want to investigate to determine which reorder point will produce the desired service level. To track the service level associated with each replication of the model, the following formula is entered in cell B10:

Formula in cell B10: =Model!E36

We can now use the Data Table command to instruct Excel to substitute each value in the range C10 through G10 into cell E5 and recalculate the workbook 200 times, keeping track of the value that results in cell B10 for each replication. This is done as shown in the steps on page 518.

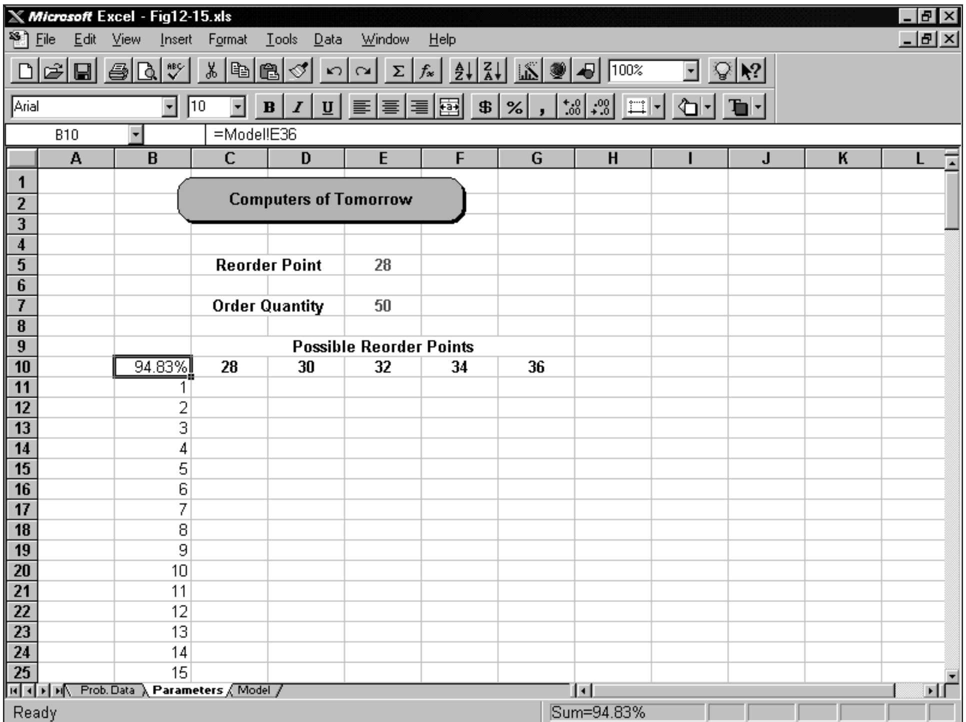


Figure 12.18 Spreadsheet prepared to simulate the service level at various reorder points using a two-way data table.

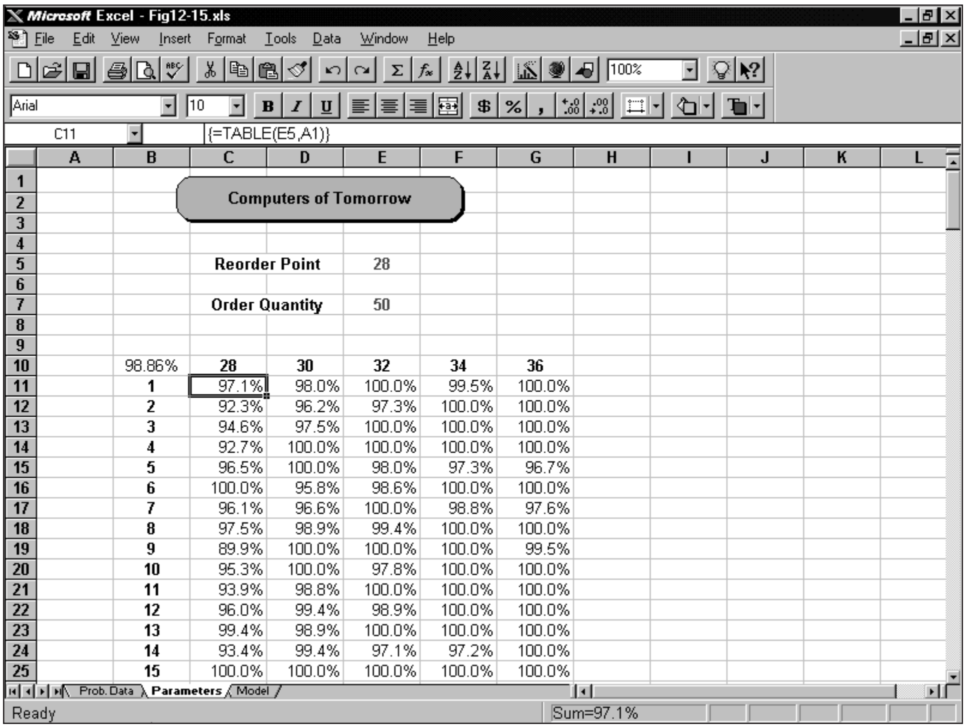
Key Cell Formulas

Cell	*Formula*	*Copied to*
B10	=Model!E36	—

> ### *Important Software Note*
>
> When using a two-way data table, the formula whose value you want to track must be placed in the upper-left corner of the data table range.

1. Select the range B10 through G210.
2. Click the Data menu.
3. Click Table.
4. In the Table dialog box, enter cell E5 for the Row input cell and enter cell A1 for the Column input cell.
5. Click OK. (This operation might take a minute or two if your computer is slow.)

Excel substitutes each value in the range C10 through G10 into cell E5. For each value substituted into cell E5, each value in the range B11 through B210 is substituted into cell A1 and the workbook is recalculated. The resulting value in cell B10 (representing the service level occurring in that replication) is then recorded in the appropriate cell in the data table. So for each reorder point value entered into cell E5, the workbook is replicated 200 times and the service level observed in each replication is recorded in the appropriate column of the data table. Because the model being recalculated here is fairly large, this could take several minutes depending on the speed of your computer. The results of this operation (after some simple formatting) are shown in Figure 12.19. (The random numbers you generate will *not* match those shown in Figure 12.19.)



	B	C	D	E	F	G
10	98.86%	28	30	32	34	36
11	1	97.1%	98.0%	100.0%	99.5%	100.0%
12	2	92.3%	96.2%	97.3%	100.0%	100.0%
13	3	94.6%	97.5%	100.0%	100.0%	100.0%
14	4	92.7%	100.0%	100.0%	100.0%	100.0%
15	5	96.5%	100.0%	98.0%	97.3%	96.7%
16	6	100.0%	95.8%	98.6%	100.0%	100.0%
17	7	96.1%	96.6%	100.0%	98.8%	97.6%
18	8	97.5%	98.9%	99.4%	100.0%	100.0%
19	9	89.9%	100.0%	100.0%	100.0%	99.5%
20	10	95.3%	100.0%	97.8%	100.0%	100.0%
21	11	93.9%	98.8%	100.0%	100.0%	100.0%
22	12	96.0%	99.4%	98.9%	100.0%	100.0%
23	13	99.4%	98.9%	100.0%	100.0%	100.0%
24	14	93.4%	99.4%	97.1%	97.2%	100.0%
25	15	100.0%	100.0%	100.0%	100.0%	100.0%

Figure 12.19 Results of the Data Table command for COT's inventory problem.

Each new entry Excel created starting in cell C11 contains the array formula {=TABLE(E5,A1)}. This is how the Data Table command performs the repeated substitution and recalculation. If we don't change these formulas into values, every time the spreadsheet is recalculated we'll have to wait (and wait) for this process to re-execute and we'll get a new sample of 200 replications for each reorder point listed at the top of the data table. To convert the contents of the data table into values:

1. Select the range C11 through G210.
2. Click the Edit menu.
3. Click Copy.
4. Click the Edit menu.
5. Click Paste Special.
6. Click Values.
7. Click OK.
8. Press [Enter].

The values in the data table are now numeric constants that will not change even if the spreadsheet is recalculated.

12.12.4 Data Analysis

The results of the simulation can be summarized easily using the Descriptive Statistics command described earlier. To do this:

1. Click the Tools menu.
2. Click Data Analysis.
3. Click Descriptive Statistics.
4. Complete the Descriptive Statistics dialog box, as shown in Figure 12.20.
5. Click OK. (This operation might take a minute or two if your computer is slow.)

The descriptive statistics about the sample data are placed on a new worksheet named Results. After deleting some of the extraneous information produced by the Descriptive Statistics command, and after some simple formatting, the results appear as shown in Figure 12.21.

Figure 12.21 indicates that COT's current reorder point of 28 units results in an average service level of approximately 95.92%. This implies that COT is unable to satisfy approximately 4.08% of the total demand using this reorder point. Cell C9 translates this into dollars as:

Formula for cell C9: =(1–C4)*6*30*45
(Copy to D9 through G9.)

Because the average daily demand for the monitor is for six units, and each monitor sold generates a marginal profit of $45, cell C9 indicates that in an average month COT loses about $330.55 in profit due to stockouts on this item.

However, our simulation results indicate that as the reorder point increases, the service level also increases and the percentage of lost sales decreases. Thus, if Laura increases the reorder point to 36, COT's average service level for the monitor would increase to 99.68% and the average monthly profit lost due to stockouts would decrease to approximately $25.75.

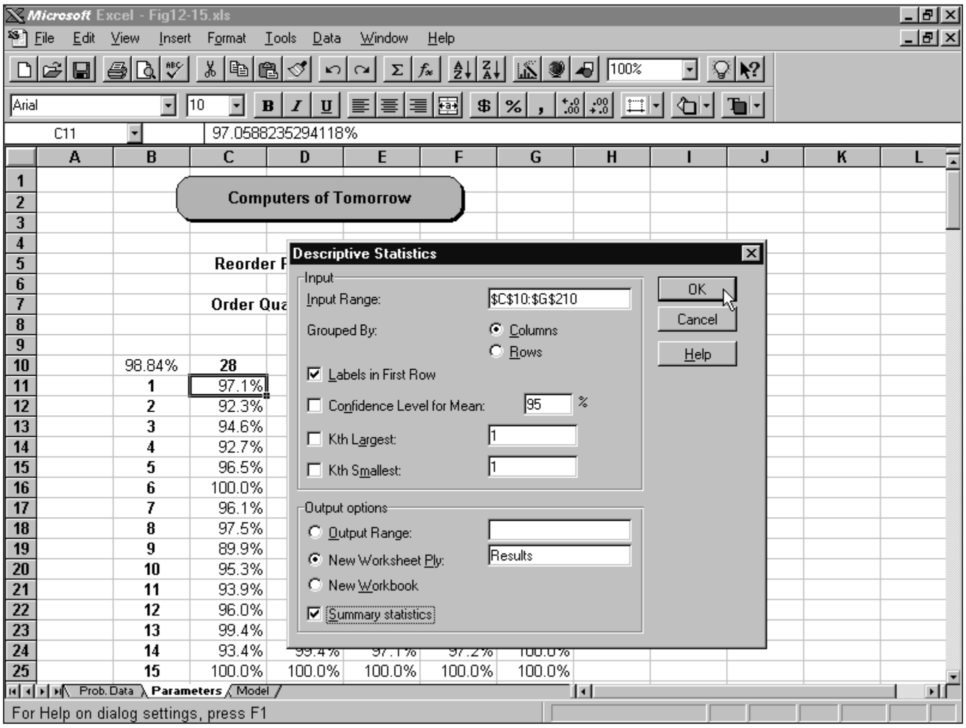


Figure 12.20 Descriptive Statistics dialog box for COT's inventory problem.

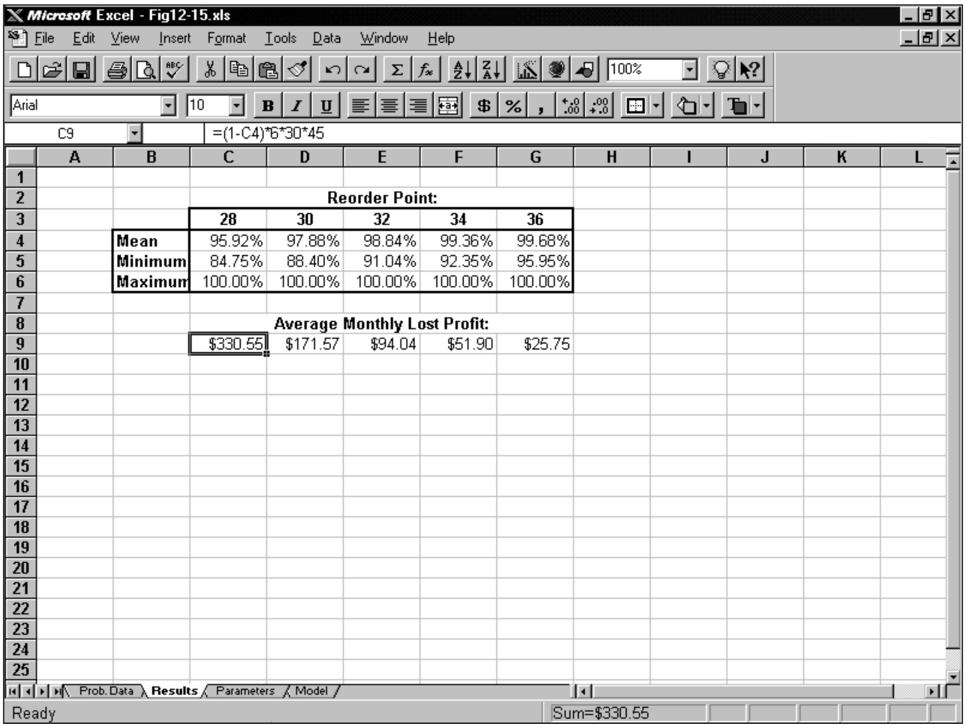
12.12.5 *A Final Comment on the Inventory Example*

Although increasing the reorder point decreases the percentage of lost sales, it also increases the average level of inventory held. Thus, another objective that might be considered in this problem involves weighing the costs of holding more inventory against the benefits of having fewer lost sales. We'll consider such an objective further in one of the problems at the end of this chapter.

12.13 *AN OVERBOOKING EXAMPLE*

Businesses that allow customers to make reservations for services (such as airlines, hotels, and car rental companies) know that some percentage of the reservations made will not be used for one reason or another, leaving these companies with a difficult decision problem. If they accept reservations for only the number of customers that can actually be served, then a portion of the company's assets will be underutilized when some customers with reservations fail to arrive. This results in an opportunity loss for the company. On the other hand, if they overbook (or accept more reservations than can be handled), then at times more customers will arrive than can actually be served. This typically results in additional financial costs to the company and often generates ill will among those customers who cannot be served. The following example illustrates how simulation might be used to help a company determine the optimal number of reservations to accept.

Marty Ford is an operations analyst for Piedmont Commuter Airlines (PCA). Recently, Marty was asked to make a recommendation on how many reservations



	Reorder Point:				
	28	30	32	34	36
Mean	95.92%	97.88%	98.84%	99.36%	99.68%
Minimum	84.75%	88.40%	91.04%	92.35%	95.95%
Maximum	100.00%	100.00%	100.00%	100.00%	100.00%

	Average Monthly Lost Profit:				
	$330.55	$171.57	$94.04	$51.90	$25.75

Figure 12.21 Simulation results for COT's inventory problem.

Key Cell Formulas

Cell	*Formula*	*Copied to*
C9	=(1-C4)*6*30*45	D9:G9

PCA should book on Flight 343—a flight in high demand from a small regional airport to a major hub airport. Historical data show that PCA frequently has seats left on Flight 343 if it accepts only 19 reservations (the plane's capacity). Industry statistics show that for every ticket sold for a commuter flight, a 0.10 probability exists that the ticket holder will not be on the flight. PCA sells non-refundable tickets for Flight 343 for $85 per seat. Thus, every empty seat on this flight represents an opportunity cost—even if a no-show customer had purchased a ticket for the seat—because the seat could be filled by another passenger paying $85. On the other hand, if PCA overbooks this flight and more than 19 passengers show up, some of them will have to be bumped to a later flight. To compensate for the inconvenience of being bumped, PCA gives these passengers vouchers for a free meal, a free flight at a later date, and sometimes also pays for them to stay overnight in a hotel near the airport. PCA pays an average of $155 for each passenger who gets bumped. Marty wants to determine if PCA can increase profits by overbooking this flight and, if so, how many reservations should be accepted to produce the maximum average profit. Marty wants to set up a model that allows him to investigate the consequences of accepting up to 24 reservations.

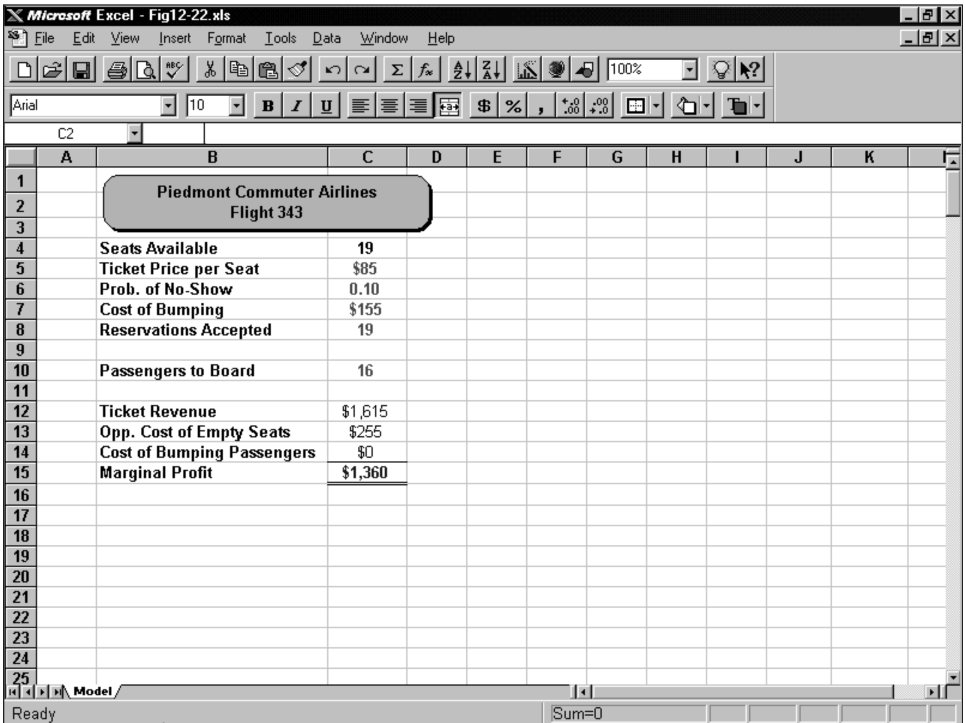
12.13.1 Implementing the Model

A spreadsheet model for this problem is shown in Figure 12.22 (and in the file FIG12-22.xls on your data disk).

This spreadsheet begins by listing the relevant data from the problem, including the number of seats available on the plane, the price PCA charges for each seat, the probability of a no-show (or a ticketed passenger not arriving in time for the flight), the cost of bumping passengers, and the number of reservations that will be accepted.

The uncertain, or random, element of this problem is the number of passengers arriving to board the plane, represented in cell C10 in Figure 12.22. If n tickets are sold and each ticket holder has a $p = 0.1$ probability of not showing up (or $1 - p = 0.9$ probability of showing up), then the number of passengers arriving to board the flight is a random variable that follows the binomial probability distribution—or a binomial random variable. Thus, the following formula for cell C10 generates the random number of ticketed passengers who arrive for each flight:

Figure 12.22 Spreadsheet model for PCA's overbooking problem.



Key Cell Formulas

Cell	*Formula*	*Copied to*
C10	=RNGBinomial(C8,1-C6)	—
C12	=C8*C5	—
C13	=C5*MAX(C4-C10,0)	—
C14	=C7*MAX(C10-C4,0)	—
C15	=C12-C13-C14	—

Formula for cell C10: =RNGBinomial(C8,1-C6)

Cell C12 represents the ticket revenue PCA earns based on the number of tickets they sell for each flight. The formula for this cell is:

Formula for cell C12: =C8*C5

Cell C13 computes the opportunity loss PCA incurs on each empty seat on a flight.

Formula for cell C13: =C5*MAX(C4-C10,0)

Cell C14 computes the costs PCA incurs when passengers must be bumped (i.e., when the number of passengers wanting to board exceeds the number of available seats).

Formula for cell C14: =C7*MAX(C10-C4,0)

Finally, cell C15 computes the marginal profit PCA earns on each flight.

Formula for cell C15: =C12-C13-C14

12.13.2 Replicating the Model

Marty wants to determine the number of reservations to accept that, on average, will result in the highest marginal profit. Figure 12.23 shows the spreadsheet in the format of a two-way data table.

In cell F5, we entered the value 1 and used the Edit, Fill, Series command to fill the remainder of this column with the values 2, 3, and so on, up to 200. Cells G4 through L4 contain the values 19, 20, 21, 22, 23, and 24, respectively, to represent different numbers of reservations that PCA might accept. The formula in cell F4 refers back to the marginal profit in cell C15 because this is the bottom-line performance measure we want to track in this model.

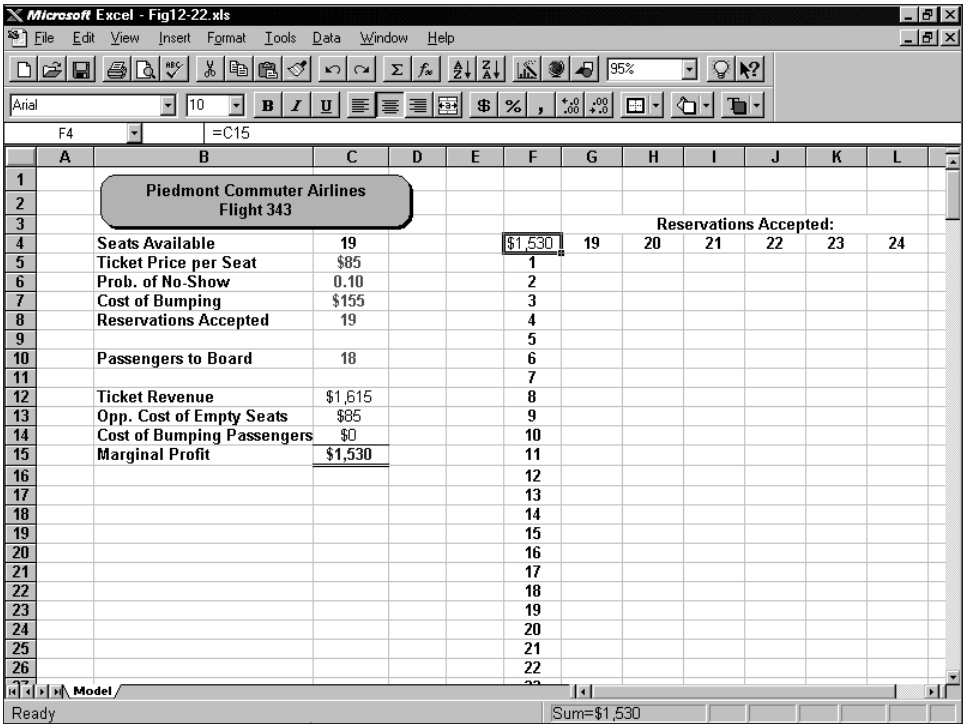
Formula for cell F4: =C15

We can now use the Data Table command to instruct Excel to substitute each value in the range G4 through L4 into cell C8 and recalculate the spreadsheet 200 times, keeping track of the value that results in cell F4 for each replication. This is done as follows:

1. Select the range F4 through L204.
2. Click the Data menu.
3. Click Table.
4. In the Table dialog box, enter cell C8 for the Row input cell and enter cell A1 for the Column input cell.
5. Click OK.

Excel substitutes each value in the range G4 through L4 into cell C8. For each value substituted into cell C8, each value in the range F5 through F204 is substituted into cell A1 and the workbook is recalculated. The resulting value in cell F4 (representing the marginal profit earned in that replication) is then recorded in the appropriate cell in the data table. So for each possible number of reservations entered into cell C8, the spreadsheet is replicated 200 times and the marginal profit observed in each replication is recorded in the appropriate column of the data table. Again, this could take several seconds to several minutes depending on the speed of your computer.

Figure 12.23
Spreadsheet prepared to simulate the average profit achieved by accepting different numbers of reservations using a two-way data table.



Key Cell Formulas

Cell	*Formula*	*Copied to*
F4	=C15	—

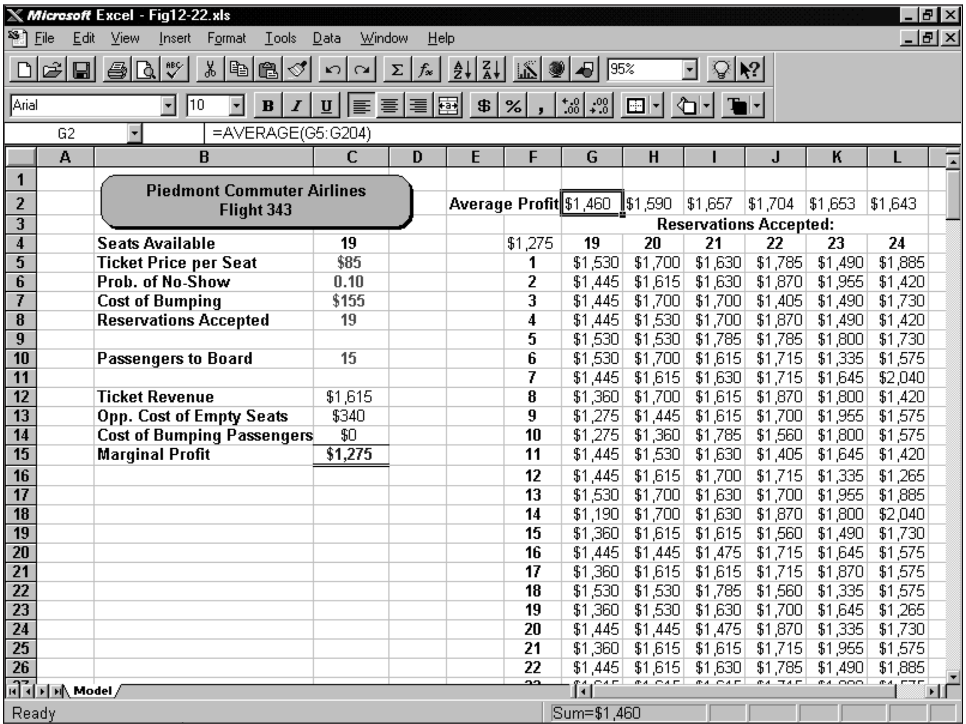
When Excel is finished running the replications, we need to convert the formula entries in the data table into values by completing the following:

1. Select the range F5 through L204.
2. Click the Edit menu.
3. Click Copy.
4. Click the Edit menu.
5. Click Paste Special.
6. Click Values.
7. Click OK.
8. Press [Enter].

Figure 12.24 shows the results of the simulation and the average profit associated with each number of reservations accepted. (Your results will *not* match those shown in Figure 12.24.) These averages are calculated as:

Formula for cell G2: =AVERAGE(G5:G204)
(Copy to H2 through L2.)

The averages in Figure 12.24 indicate that if PCA accepts only 19 reservations, its average marginal profit on this flight will be approximately $1,460 (assuming vacant seats are counted as an opportunity cost at $85 per seat). As the number of reservations



	A	B	C	D	E	F	G	H	I	J	K	L
1												
2		Piedmont Commuter Airlines Flight 343			Average Profit		$1,460	$1,590	$1,657	$1,704	$1,653	$1,643
3							Reservations Accepted:					
4		Seats Available	19			$1,275	19	20	21	22	23	24
5		Ticket Price per Seat	$85			1	$1,530	$1,700	$1,630	$1,785	$1,490	$1,885
6		Prob. of No-Show	0.10			2	$1,445	$1,615	$1,630	$1,870	$1,955	$1,420
7		Cost of Bumping	$155			3	$1,445	$1,700	$1,700	$1,405	$1,490	$1,730
8		Reservations Accepted	19			4	$1,445	$1,530	$1,700	$1,870	$1,490	$1,420
9						5	$1,530	$1,530	$1,785	$1,785	$1,800	$1,730
10		Passengers to Board	15			6	$1,530	$1,700	$1,615	$1,715	$1,335	$1,575
11						7	$1,445	$1,615	$1,630	$1,715	$1,645	$2,040
12		Ticket Revenue	$1,615			8	$1,360	$1,700	$1,615	$1,870	$1,800	$1,420
13		Opp. Cost of Empty Seats	$340			9	$1,275	$1,445	$1,615	$1,700	$1,955	$1,575
14		Cost of Bumping Passengers	$0			10	$1,275	$1,360	$1,785	$1,560	$1,800	$1,575
15		Marginal Profit	$1,275			11	$1,445	$1,530	$1,630	$1,405	$1,645	$1,420
16						12	$1,445	$1,615	$1,700	$1,715	$1,335	$1,265
17						13	$1,530	$1,700	$1,630	$1,700	$1,955	$1,885
18						14	$1,190	$1,700	$1,630	$1,870	$1,800	$2,040
19						15	$1,360	$1,615	$1,615	$1,560	$1,490	$1,730
20						16	$1,445	$1,445	$1,475	$1,715	$1,645	$1,575
21						17	$1,360	$1,615	$1,615	$1,715	$1,870	$1,575
22						18	$1,530	$1,530	$1,785	$1,560	$1,335	$1,575
23						19	$1,360	$1,530	$1,630	$1,700	$1,645	$1,265
24						20	$1,445	$1,445	$1,475	$1,870	$1,335	$1,730
25						21	$1,360	$1,615	$1,615	$1,715	$1,955	$1,575
26						22	$1,445	$1,615	$1,630	$1,785	$1,490	$1,885

Key Cell Formulas

Cell	*Formula*	*Copied to*
G2	=AVERAGE(G5:G204)	H2:L2

Figure 12.24
Simulation results for PCA's overbooking problem.

accepted increases, the average profit also increases and reaches a maximum of $1,704 at 22 reservations. Thus, it appears that the optimal number of reservations for PCA to accept is 22.

12.13.3 A Final Comment on the Overbooking Example

This problem assumed that all the reservations available for Flight 343 would always be taken—or that there would never be unused reservations. However, if PCA accepts up to 22 reservations, there might be times when only 16 or 17 seats would be demanded. Thus, the demand for seats on this flight could be modeled more accurately as a random variable. We'll explore this issue more fully in one of the problems at the end of this chapter.

SUMMARY

This chapter introduced the concept of risk analysis and simulation. Many of the input cells in a spreadsheet represent random variables whose values cannot be determined with certainty. Any uncertainty in the input cells flows through the spreadsheet model to create a related uncertainty in the value of the output cell(s). Decisions made on the basis of these uncertain values involve some degree of risk.

THE WORLD OF MANAGEMENT SCIENCE

The U.S. Postal Service Moves to the Fast Lane

Mail flows into the U.S. Postal Service at the rate of 500 million pieces per day, and it comes in many forms. There are standard-sized letters with nine-digit ZIP codes (with or without imprinted bar codes), five-digit ZIP codes, typed addresses that can be read by optical character readers, handwritten addresses that are barely decipherable. Christmas cards in red envelopes addressed in red ink and so on. The enormous task of sorting all these pieces at the sending post office and at the destination has caused postal management to consider and adopt many new forms of technology. These include operator-assisted mechanized sorters, optical character readers (last-line and multiple-line), and bar code sorters. Implementation of new technology brings with it associated policy decisions, such as rate discounts for bar coding by the customer, finer sorting at the origin, and so on.

A simulation model called META (model for evaluating technology alternatives) assists management in evaluating new technologies, configurations, and operating plans. Using distributions based on experience or projections of the effects of new policies, META simulates a random stream of mail of different types, routes the mail through the system configuration being tested, and prints reports detailing total pieces handled, capacity utilization, work hours required, space requirements, and cost.

META has been used on several projects associated with the Postal Service corporate automation plan. These include facilities planning, benefits of alternate sorting plans, justification of efforts to enhance address readability, planning studies for reducing the time carriers spend sorting vs. delivering, and identification of mail types that offer the greatest potential for cost savings.

According to the Associate Postmaster General, "... META became the vehicle to help steer our organization on an entirely new course at a speed we had never before experienced."

Source: Michael E. Cebry, Anura H. deSilva and Fred J. DiLisio, "Management Science in Automating Postal Operations: Facility and Equipment Planning in the United States Postal Service," *Interfaces*, vol. 22, no. 1, January-February 1992, pages 110–130.

Various methods of risk analysis are available, including best-case/worst-case analysis, what-if analysis, and simulation. Of these three methods, simulation is the only technique that provides hard evidence (facts and figures) that can be used objectively in making decisions. To simulate a model, RNGs are used to select representative values for each uncertain independent variable in the model. This process is repeated over and over to generate a sample of representative values for the dependent variable(s) in the model. The variability and distribution of the sample values for the dependent variable(s) can then be analyzed to gain insight into the possible outcomes that might occur. This technique can also be used to test different configurations of controllable parameters in the model in an attempt to determine the optimal values for these parameters.

QUESTIONS AND PROBLEMS

1. Under what condition(s) is it appropriate to use simulation to analyze a model? That is, what characteristics should a model possess in order for simulation to be used?

2. The graph of the probability distribution of a normally distributed random variable with a mean of 20 and standard deviation of 3 is shown in Figure 12.6.

   a. Use the RNGNormal( ) function to generate 100 sample values from this distribution.
   b. Produce a histogram of the 100 sample values you generated. Does your histogram look like the graph for this distribution in Figure 12.6?
   c. Repeat this experiment, only this time sample 1,000 values.
   d. Produce a histogram for the 1,000 sample values you generated. Does the histogram now more closely resemble the graph in Figure 12.6 for this distribution?
   e. Why does your second histogram look more "normal" than the first one?

3. Refer to the Hungry Dawg Restaurant example presented in section 12.4 of this chapter. Health claim costs actually tend to be seasonal, with higher levels of claims occurring during the summer months (when kids are out of school and more likely to injure themselves) and during December (when people schedule elective procedures before the next year's deductible must be paid). The following table summarizes the seasonal adjustment factors that apply to RNGs for average claims in the Hungry Dawg problem. For instance, the average claim for month 6 should be multiplied by 115% and those for month 1 should be multiplied by 80%.

Month:	1	2	3	4	5	6	7	8	9	10	11	12
Seasonal Factor:	0.80	0.85	0.87	0.92	0.93	1.15	1.20	1.18	1.03	0.95	0.98	1.14

   Suppose the company maintains an account from which it pays health insurance claims. Assume there is $2.5 million in the account at the beginning of month 1. Each month, employee contributions are deposited into this account and claims are paid from the account.

   a. Modify the spreadsheet shown in Figure 12.7 to include the cash flows in this account. If the company deposits $3 million in this account every month, what is the probability that the account will have insufficient funds to pay claims at some point during the year? Use 300 replications. (HINT: You can use the COUNTIF( ) function to count the number of months in a year where the ending balance in the account is below 0.)
   b. If the company wants to deposit an equal amount of money in this account each month, what should this amount be if it wants there to only be a 5% chance of having insufficient funds?

4. One of the examples in section 12.11 of this chapter dealt with determining the optimal reorder point for a computer monitor sold by Computers of Tomorrow (COT) in Austin, Texas. In this example we found that increasing the reorder point decreased the number of lost sales. However, this also increased the average amount of inventory carried. Suppose that it costs COT $0.30 per day in holding

costs for each monitor in beginning inventory, and it costs $20 to place an order. Each monitor sold generates a profit of $45, and each lost sale results in an opportunity cost of $65 (including the lost profit of $45 and $20 in lost goodwill). Modify the spreadsheet shown in Figure 12.15 to determine the reorder point that maximizes the average monthly profit associated with this monitor.

5. One of the examples in section 12.13 of this chapter dealt with determining the optimal number of reservations for Piedmont Commuter Airlines (PCA) to accept for one of its flights that uses a plane with 19 seats. The model discussed in the chapter assumes that all the reservations would be used, but some customers would not show up for the flight. It is probably more realistic to assume that the demand for these reservations is somewhat random. For example, suppose that the demand for reservations on this flight is given by the following discrete probability distribution:

Reservations Demanded:	12	13	14	15	16	17	18	19	20	21	22	23	24
Probability:	0.01	0.03	0.04	0.07	0.09	0.11	0.15	0.18	0.14	0.08	0.05	0.03	0.02

The number of passengers receiving reservations depends on the number of reservations PCA accepts and the demand for these reservations. For each passenger receiving a reservation, a 0.10 probability exists that he will not arrive at the gate to board the plane. Modify the spreadsheet shown in Figure 12.22 to determine the number of reservations PCA should accept to maximize the average profit associated with this flight.

6. A debate recently erupted about the optimal strategy for playing a game on the TV show called "Let's Make a Deal." In one of the games on this show, the contestant would be given the choice of prizes behind three closed doors. A valuable prize was behind one door and worthless prizes were behind the other two doors. After the contestant selected a door, the host would open one of the two remaining doors to reveal one of the worthless prizes. Then, before opening the selected door, the host would give the contestant the opportunity to switch his or her selection to the other door that had not been opened. The question is, should the contestant switch?

   a. Suppose a contestant is allowed to play this game 500 times, always picks door number 1, and never switches when given the option. If the valuable prize is equally likely to be behind each door at the beginning of each play, how many times would the contestant win the valuable prize? Use simulation to answer this question.
   b. Now suppose the contestant is allowed to play this game another 500 times. This time the player always selects door number 1 initially and switches when given the option. Using simulation, how many times would the contestant win the valuable prize?
   c. If you were a contestant on this show, what would you do if given the option of switching doors?

7. The monthly demand for the latest computer at Newland Computers follows a normal distribution with a mean of 350 and standard deviation of 75. Newland purchases these computers for $1,200 and sells them for $2,300. It costs the company $100 to place an order and $12 for every computer held in inventory at the end of each month. Currently the company places an order for 1,000 computers

whenever the inventory at the end of a month falls below 100 units. Assume unmet demand in any month is lost to competitors and that orders placed at the end of one month arrive at the beginning of the next month.

a. Create a spreadsheet model to simulate the profit the company will earn on this product over the next two years. Use 200 replications. What is the average level of profit the company will earn?
b. Suppose the company wants to determine the optimum reorder point and order quantity. Specifically, for every 100-unit increase in the reorder point, they will reduce the order quantity by 100. Which combination of reorder point and order quantity will provide the highest average profit over the next two years?

8. The manager of Moore's Catalog Showroom is trying to predict how much revenue will be generated by each major department in the store during 1998. The manager has estimated the minimum and maximum growth rates possible for revenues in each department. The manager believes that any of the possible growth rates between the minimum and maximum values are equally likely to occur. These estimates are summarized in the following table:

		Growth Rates	
Department	*1997 Revenues*	*Minimum*	*Maximum*
Electronics	$6,342,213	2%	10%
Garden Supplies	$1,203,231	–4%	5%
Jewelry	$4,367,342	–2%	6%
Sporting Goods	$3,543,532	–1%	8%
Toys	$4,342,132	4%	15%

Create a spreadsheet to simulate the total revenues that could occur in the coming year. Run 500 replications of the model and do the following:

a. Construct a 95% confidence interval for the average level of revenues the manager could expect for 1998.
b. According to your model, what are the chances that total revenues in 1998 will be more than 5% larger than those in 1997?

9. The Harriet Hotel in downtown Boston has 100 rooms that rent for $125 per night. It costs the hotel $30 per room in variable costs (cleaning, bathroom items, etc.) each night a room is occupied. For each reservation accepted, there is a 5% chance that the guest will not arrive. If the hotel overbooks, it costs $200 to compensate guests whose reservations cannot be honored.

   How many reservations should the hotel accept if it wishes to maximize the average daily profit? Use 500 simulations for each reservation level you consider.

10. Lynn Price recently completed her MBA and accepted a job with an electronics manufacturing company. Although she likes her job, she is also looking forward to retiring one day. To ensure that her retirement is comfortable, Lynn intends to invest $3,000 of her salary into a tax-sheltered retirement fund at the end of each year. Lynn is not certain what rate of return this investment will earn each year, but she expects each year's rate of return could be modeled appropriately as a normally distributed random variable with a mean of 12% and standard deviation of 2%.

a. If Lynn is 30 years old, how much money should she expect to have in her retirement fund at age 60? (Use 500 replications.)
b. Construct a 95% confidence interval for the average amount Lynn will have at age 60.
c. What is the probability that Lynn will have more than $1 million in her retirement fund when she reaches age 60?
d. How much should Lynn invest each year if she wants there to be a 90% chance of having at least $1 million in her retirement fund at age 60?

11. Employees of Georgia-Atlantic are permitted to contribute a portion of their earnings (in increments of $500) to a flexible spending account from which they can pay medical expenses not covered by the company's health insurance program. Contributions to an employee's "flex" account are not subject to income taxes. However, the employee forfeits any amount contributed to the "flex" account that is not spent during the year. Suppose Greg Davis makes $60,000 per year from Georgia-Atlantic and pays a marginal tax rate of 28%. Greg and his wife estimate that in the coming year their normal medical expenses not covered by the health insurance program could be as small as $500, as large as $5,000 and most likely about $1,300. However, Greg also believes there is a 5% change that an abnormal medical event could occur which might add $10,000 to the normal expenses paid from their flex account. If their uncovered medical claims exceed their contribution to their "flex" account, they will have to cover these expenses with the after-tax money Greg brings home.

    Use simulation to determine the amount of money Greg should contribute to his flexible spending account in the coming year if he wants to maximize his disposable income (after taxes and all medical expenses are paid). Use 500 replications for each level of "flex" account contribution you consider.

12. Acme Equipment Company is considering the development of a new machine that would be marketed to tire manufacturers. Research and development costs for the project are expected to be about $4 million but could vary between $3 and $6 million. The market life for the product is estimated to be three to eight years with all intervening possibilities being equally likely. The company thinks it will sell 250 units per year, but acknowledges that this figure could be as low as 50 or as high as 350. The company will sell the machine for about $23,000. Finally, the cost of manufacturing the machine is expected to be $14,000 but could be as low as $12,000 or as high as $18,000. The company's cost of capital is 15%.

    a. Use appropriate RNGs to create a spreadsheet to calculate the possible net present values (NPVs) that could result from taking on this project.
    b. Replicate the model 500 times. What is the expected NPV for this project?
    c. What is the probability of this project generating a positive NPV for the company?

13. Representatives from the American Heart Association are planning to go door-to-door throughout a community, soliciting contributions. From past experience they know that when someone answers the door, 80% of the time it is a female and 20% of the time it is a male. They also know that 70% of the females who answer the door make a donation, whereas only 40% of the males who answer the door make donations. The amount of money that females contribute follows a normal distribution with a mean of $20 and standard deviation of $3. The amount of money males contribute follows a normal distribution with a mean of $10 and standard deviation of $2.

a. Create a spreadsheet model that simulates what might happen whenever a representative of the American Heart Association knocks on a door and someone answers.
b. Replicate your model 500 times. What is the average contribution the Heart Association can expect to receive when someone answers the door?
c. Suppose that the Heart Association plans to visit 300 homes on a given Saturday. If no one is home at 25% of the residences, what is the total amount that the Heart Association can expect to receive in donations?

14. After spending 10 years as an assistant manager for a large restaurant chain, Ray Clark has decided to become his own boss. The owner of a local submarine sandwich store wants to sell the store to Ray for $65,000 to be paid in installments of $13,000 in each of the next five years. According to the current owner, the store brings in revenue of about $110,000 per year and incurs operating costs of about 63% of sales. Thus, once the store is paid for, Ray should make about $35,000–$40,000 per year before taxes. Until the store is paid for, he will make substantially less—but he will be his own boss. Realizing that some uncertainty is involved in this decision, Ray wants to simulate what level of net income he can expect to earn during the next five years as he operates and pays for the store. In particular, he wants to see what could happen if sales are allowed to vary uniformly between $90,000 and $120,000, and if operating costs are allowed to vary uniformly between 60% and 65% of sales. Assume that Ray's payments for the store are not deductible for tax purposes and that he is in the 28% tax bracket.

a. Create a spreadsheet model to simulate the annual net income Ray would receive during each of the next five years if he decides to buy the store.
b. Given the money he has in savings, Ray thinks he can get by for the next five years if he can make at least $12,000 from the store each year. Replicate the model 500 times and track: 1) the minimum amount of money Ray makes over the five-year period represented by each replication, and 2) the total amount Ray makes during the five-year period represented by each replication.
c. What is the probability that Ray will make at least $12,000 in each of the next five years?
d. What is the probability that Ray will make at least $60,000 total over the next five years?

15. Bob Davidson owns a newsstand outside the Waterstone office building complex in Atlanta, near Hartsfield International Airport. He buys his papers wholesale at $0.50 per paper and sells them for $0.75. Bob wonders what is the optimal number of papers to order each day. Based on history, he has found that demand (even though it is discrete) can be modeled by a normal distribution with a mean of 50 and standard deviation of 5. When he has more papers than customers, he can recycle all the extra papers the next day and receive $0.05 per paper. On the other hand, if he has more customers than papers, he loses some goodwill in addition to the lost profit on the potential sale of $0.25. Bob estimates the incremental lost goodwill costs five day's worth of business (that is, dissatisfied customers will go to a competitor the next week, but come back to him the week after that).

a. Create a spreadsheet model to determine the optimal number of papers to order each day. Use 250 replications and round the demand values generated by the normal RNG to the closest integer value.
b. Construct a 95% confidence interval for the expected payoff from the optimal decision.

16. Vinton Auto Insurance is trying to decide how much money to keep in liquid assets to cover insurance claims. In the past, the company held some of the premiums it received in interest-bearing checking accounts and put the rest into investments that are not quite as liquid, but tend to generate a higher investment return. The company wants to study cash flows to determine how much money it should keep in liquid assets to pay claims. After reviewing historical data, the company determined that the average repair bill per claim is normally distributed with a mean of $1,700 and standard deviation of $400. It also determined that the number of repair claims filed each week is a random variable that follows the probability distribution given below:

Number of Repair Claims	*Probability*
1	0.05
2	0.06
3	0.10
4	0.17
5	0.28
6	0.14
7	0.08
8	0.07
9	0.05

In addition to repair claims, the company also receives claims for cars that have been "totaled" and cannot be repaired. A 20% chance of receiving this type of claim exists in any week. These claims for "totaled" cars typically cost anywhere from $2,000 to $35,000, with $13,000 being the most common cost.

a. Create a spreadsheet model of the total claims cost incurred by the company in any week.
b. Replicate the model 500 times and create a histogram of the distribution of total cost values that were generated.
c. What is the average cost the company should expect to pay each week?
d. Suppose that the company decides to keep $20,000 cash on hand to pay claims. What is the probability that this amount would not be adequate to cover claims in any week?
e. Create a 95% confidence interval for the true probability of claims exceeding $20,000 in a given week.

17. The owner of a local car dealership has just received a call from a regional distributor stating that a $5,000 bonus will be awarded if the owner's dealership sells at least 10 new cars next Saturday. On an average Saturday, this dealership has 75 potential customers look at new cars, but there is no way to determine exactly how many customers will come this particular Saturday. The owner is fairly certain that the number would not be less than 40, but also thinks it would be unrealistic to expect more than 120 (which is the largest number of customers to ever show up in one day).

    The owner determined that, on average, about 1 out of 10 customers who look at cars at the dealership actually purchase a car—or, a 0.10 probability (or 10% chance) exists that any given customer will buy a new car.

a. Create a spreadsheet model for this problem and generate 1,000 random outcomes for the number of cars the dealership might sell next Saturday.

b. What is the probability that the dealership will earn the $5,000 bonus?
c. If you were this dealer, what is the maximum amount of money you'd be willing to spend on sales incentives to try to earn this bonus?

18. Dr. Sarah Benson is an ophthalmologist who, in addition to prescribing glasses and contact lenses, performs optical laser surgery to correct nearsightedness. This surgery is fairly easy and inexpensive to perform. Thus, it represents a potential gold mine for her practice. To inform the public about this procedure, Dr. Benson advertises in the local paper and holds information sessions in her office one night a week at which she shows a videotape about the procedure and answers any questions potential patients might have. The room where these meetings are held can seat 10 people, and reservations are required. The number of people attending each session varies from week to week. Dr. Benson cancels the meeting if two or fewer people have made reservations. Using data from the previous year, Dr. Benson determined that the distribution of reservations is as follows:

Number of Reservations:	0	1	2	3	4	5	6	7	8	9	10
Probability:	0.02	0.05	0.08	0.16	0.26	0.18	0.11	0.07	0.05	0.01	0.01

Using data from the past year, Dr. Benson determined that each person who attends an information has a 0.25 probability of electing to have the surgery. Of those who do not, most cite the cost of the procedure—$2,000—as their major concern.

a. On average, how much revenue does Dr. Benson's practice in laser surgery generate each week? (Use 500 replications.)
b. On average, how much revenue would the laser surgery generate each week if Dr. Benson did not cancel sessions with two or fewer reservations?
c. Dr. Benson believes that 40% of the people attending the information sessions would have the surgery if she reduced the price to $1,500. Under this scenario, how much revenue could Dr. Benson expect to realize per week from laser surgery?

19. Michael Abrams runs a specialty clothing store that sells collegiate sports apparel. One of his primary business opportunities involves selling custom screen-printed sweatshirts for college football bowl games. He is trying to determine how many sweatshirts to produce for the upcoming Tangerine Bowl game. During the month before the game, Michael plans to sell his sweatshirts for $25 a piece. At this price, he believes the demand for sweatshirts will be triangularly distributed with a minimum demand of 10,000, maximum demand of 30,000, and a most likely demand of 18,000. During the month after the game, Michael plans to sell any remaining sweatshirts for $12 a piece. At this price, he believes the demand for sweatshirts will be triangularly distributed with a minimum demand of 2,000, maximum demand of 7,000, and a most likely demand of 5,000. Two months after the game, Michael plans to sell any remaining sweatshirts to a surplus store which has agreed to buy up to 2,000 sweatshirts for a price of $3 per shirt. Michael can order custom screen-printed sweatshirts for $8 a piece in lot sizes of 3,000.

a. On average, how much profit would Michael earn if he orders 18,000 sweatshirts? Use 500 replications.
b. How many sweatshirts should he order if he wants to maximize his expected profit? Again use 500 replications in each simulation you perform.

20. The Major Motors Corporation is trying to decide whether or not to introduce a new midsize car. The directors of the company want to produce the car only if it has at least an 80% chance of generating a positive net present value over the next 10 years. If the company decides to produce the car, it will have to pay an uncertain initial start-up cost that is estimated to follow a triangular distribution with a minimum value of $300 million, maximum value of $600 million, and a most likely value of $450 million. In the first year the company would produce 100,000 units. Demand during the first year is uncertain but expected to be normally distributed with a mean of 95,000 and standard deviation of 7,000. For any year in which the demand exceeds production, production will be increased by 5% in the following year. For any year in which the production exceeds demand, production will be decreased by 5% in the next year and the excess cars will be sold to a rental car company at a 20% discount. After the first year, the demand in any year will be modeled as a normally distributed random variable with a mean equal to the actual demand in the previous year and standard deviation of 7,000. In the first year, the sales price of the car will be $13,000 and the total variable cost per car is expected to be $7,500. Both the selling price and variable cost is expected to increase each year at the rate of inflation which is assumed to be uniformly distributed between 2% and 7%. The company uses a discount rate of 9% to discount future cash flows.

    a. Create a spreadsheet model for this problem and replicate it 300 times. What is the minimum, average, and maximum NPV Major Motors can expect if they decide to produce this car? (HINT: Consider using the NPV( ) function to discount the profits Major Motors would earn each year.)
    b. What is the probability of Major Motors earning a positive NPV over the next 10 years?
    c. Should they produce this car?

21. Each year, the Schriber Corporation must determine how much to contribute to the company's pension plan. The company uses a 10-year planning horizon to determine the contribution which, if made annually in each of the next 10 years, would allow for only a 10% chance of the fund running short of money. The company then makes that contribution in the current year (year 1) and repeats this process in each subsequent year to determine the specific amount to contribute each year. (Last year the company contributed $43 million to the plan.) The pension plan covers two types of employees: hourly and salaried. In the current year, there will be 6,000 former hourly employees and 3,000 former salaried employees receiving benefits from the plan. The change in the number of retired hourly employees from one year to the next is expected to vary according to a normal distribution with mean of 4% and standard deviation of 1%. The change in the number of retired salaried employees from one year to the next is expected to vary between 1% and 4% according to a truncated normal distribution with mean of 2% and standard deviation of 1%. Currently, hourly retirees receive an average benefit of $15,000 per year, while salaried retirees receive an average annual benefit of $40,000. Both of these averages are expected to increase annually with the rate of inflation, which is assumed to vary annually between 2% and 7% according to a triangular distribution with a most likely value of 3.5%. The current balance in the company's pension fund is $1.5 billion. Investments in this fund earn an annual return that is assumed to be normally distributed with a mean of 12% and standard deviation of 2% each year. Create a spreadsheet model for this problem and use

simulation to determine the pension fund contribution the company should make in the current year. Assume benefits are paid throughout the year and the company contribution is made at the end of the year. What is your recommendation?

CASE 12.1 THE SOUND'S ALIVE COMPANY

Contributed by Jack Yurkiewicz, Lubin School of Business, Pace University.

Marissa Jones is the president and CEO of Sound's Alive, a company that manufactures and sells a line of speakers, CD players, receivers, high-definition televisions, and other items geared for the home entertainment market. Respected throughout the industry for bringing many high-quality, innovative products to market, Marissa is considering adding a speaker system to her product line.

The speaker market has changed dramatically during the last several years. Originally, high-fidelity aficionados knew that to reproduce sound covering the fullest range of frequencies—from the lowest kettle drum to the highest violin—a speaker system had to be large and heavy. The speaker had various drivers: a woofer to reproduce the low notes, a tweeter for the high notes, and a midrange driver for the broad spectrum of frequencies in between. Many speaker systems had a minimum of three drivers, but some had even more. The trouble was that such a system was too large for anything but the biggest rooms, and consumers were reluctant to spend thousands of dollars and give up valuable wall space to get the excellent sound these speakers could reproduce.

The trend has changed during the past several years. Consumers still want good sound, but they want it from smaller boxes. Therefore, the satellite system became popular. Consisting of two small boxes that house either one driver (to cover the midrange and high frequencies) or two (a midrange and tweeter), a satellite system can easily be mounted on walls or shelves. To reproduce the low notes, a separate subwoofer that is approximately the size of a cube 18 inches on a side is also needed. This subwoofer can be placed anywhere in the room. Taking up less space than a typical large speaker system and sounding almost as good, yet costing hundreds of dollars less, these satellite systems are hot items in the high-fidelity market.

Recently the separate wings of home entertainment—high-fidelity (receivers, speakers, CD players, CDs, cassettes, and so on), television (large-screen monitors, video cassette recorders, laser players), and computers (games with sounds, virtual reality software, and so on)—have merged into the home theater concept. To simulate the movie environment, a home theater system requires the traditional stereo speaker system plus additional speakers placed in the rear of the room so that viewers are literally surrounded with sound. Although the rear speakers do not have to match the high quality of the front speakers and, therefore, can be less expensive, most consumers choose a system in which the front and rear speakers are of equal quality, reproducing the full range of frequencies with equal fidelity.

It is this speaker market that Marissa wants to enter. She is considering having Sound's Alive manufacture and sell a home theater system that consists of seven speakers. Three small speakers—each with one dome tweeter that could reproduce the frequency range of 200 Hertz to 20,000 Hertz (upper-low frequencies to the highest frequencies)—would be placed in front, and three similar speakers would be placed strategically around the sides and back of the room. To reproduce the lowest frequencies (from 35 Hertz to 200 Hertz), a single subwoofer would also be part of the system. This sub-woofer is revolutionary because it is smaller than the ordinary sub-woofer, only 10 inches per side, and it has a built-in amplifier to power it. Consumer and critics are thrilled with the music from early prototype systems, claiming that these speakers have the best balance of sound and size. Marissa is extremely encouraged by these early reviews, and although her company has never produced a product with its house label on it (having always sold systems from established high-fidelity companies), she believes that Sound's Alive should enter the home theater market with this product.

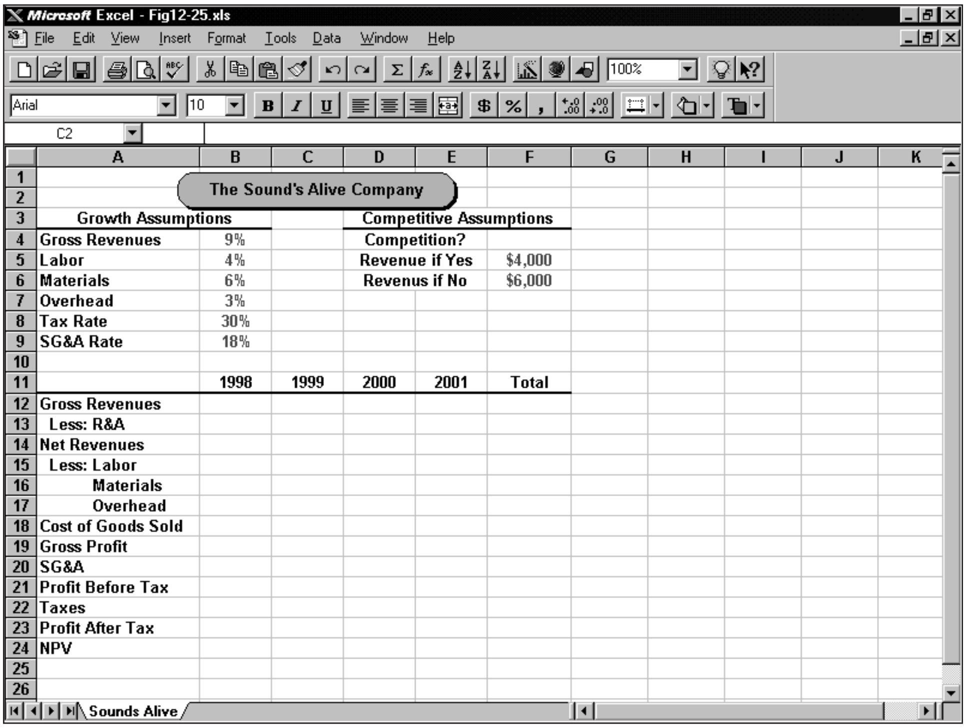
Phase One: Projecting Profits

Marissa decides to create a spreadsheet that will project profits over the next several years. After consulting with economists, market analysts, employees in her own company, and employees from other companies that sell house brand components. Marissa is confident that the gross revenues for these speakers in 1998 would be around $6 million. She must also figure that a small percentage of speakers will be damaged in transit, or some will be returned by dissatisfied customers shortly after the sales. These returns and allowances (R&As) are usually calculated as 2% of the gross revenues. Hence, the net revenues are simply the gross revenues minus the R&As. Marissa believes that the 1998 labor costs for these speakers will be $995,100. The cost of materials (including boxes to ship the speakers) should be $915,350 for 1998. Finally, her overhead costs (rent, lighting, heating in winter, air conditioning in summer, security, and so on) for 1998 should be $1,536,120. Thus, the cost of goods sold is the sum of labor, material, and overhead costs. Marissa figures the gross profit as the difference between the net revenues and the cost of goods sold. In addition, she must consider the selling, general, and administrative (SG&A) expenses. These expenses are more difficult to estimate, but the standard industry practice is to use 18% of the net revenues as the nominal percentage value for these expenses. Therefore, Marissa's profit *before taxes* is the gross profit minus the SG&A value. To calculate taxes, Marissa multiplies her profits before taxes times the tax rate, currently 30%. If her company is operating at a loss, however, no taxes would have to be paid. Finally, Marissa's net (or after tax) profit is simply the difference between the profit before taxes and the actual taxes paid.

To determine the numbers for 1999 through 2001, Marissa assumes that gross revenues, labor costs, material costs, and overhead costs will increase over the years. Although the rates of increase for these items are difficult to estimate, Marissa figures that gross revenues will increase by 9% per year, labor costs will increase by 4% per year, material costs will increase by 6% per year, and overhead costs will increase by 3% per year. She figures that the tax rate will not change from the 30% mark, and she assumes that the SG&A value will remain at 18%.

The basic layout of the spreadsheet that Marissa creates is shown in the following figure (and in the file FIG12-25.xls on your data disk). (Ignore the Competitive Assumptions section for now; we'll consider it later.) Construct the spreadsheet and determine the values for the years 1998 through 2001, then determine the totals for the four years.

Marissa not only wants to determine her net profits for 1998 through 2001, she also must justify her decisions to the company's Board of Trustees. Should she even consider entering this market, from a financial point of view? One way to answer this question is to find the net present value (NPV) of the net profits for 1998 through



Microsoft Excel - Fig12-25.xls

	A	B	C	D	E	F
1		The Sound's Alive Company				
2						
3	Growth Assumptions			Competitive Assumptions		
4	Gross Revenues	9%		Competition?		
5	Labor	4%		Revenue if Yes		$4,000
6	Materials	6%		Revenus if No		$6,000
7	Overhead	3%				
8	Tax Rate	30%				
9	SG&A Rate	18%				
10						
11		1998	1999	2000	2001	Total
12	Gross Revenues					
13	Less: R&A					
14	Net Revenues					
15	Less: Labor					
16	Materials					
17	Overhead					
18	Cost of Goods Sold					
19	Gross Profit					
20	SG&A					
21	Profit Before Tax					
22	Taxes					
23	Profit After Tax					
24	NPV					
25						
26						

Sounds Alive

Figure 12.25 Spreadsheet template for the Sound's Alive case.

2001. Use Excel's NPV capability to find the NPV, at the current interest rate of 5%, of the profit values for 1998 through 2001.

To avoid large values in the spreadsheet, enter all dollar calculations in thousands. For example, enter labor costs as 995.10 and overhead costs as 1536.12.

Phase Two: Bringing Competition Into the Model

With her spreadsheet complete, Marissa is confident that entering the home theater speaker market would be lucrative for Sound's Alive. However, she has not considered one factor in her calculations—competition. The current market leader and company she is most concerned about is the Bose Corporation. Bose pioneered the concept of a satellite speaker system, and its AMT series is very successful. Marissa is concerned that Bose will enter the home market, cutting into her gross revenues. If Bose does enter the market, Marissa believes that Sound's Alive would still make money; however, she would have to revise her gross revenues estimate from $6 million to $4 million for 1998.

To account for the competition factor, Marissa revises her spreadsheet by adding a Competition Assumptions section. Cell F4 will contain either a 0 (no competition) or a 1 (if Bose enters the market). Cells F5 and F6 provide the gross revenue estimates (in thousands of dollars) for the two possibilities. Modify your spreadsheet to take these options into account. Use the IF( ) function for the gross revenues for 1998 (cell B12). If Bose does enter the market, not only would Marissa's gross revenues be lower, but the labor, materials, and overhead costs would also be lower because Sound's Alive would be making and selling fewer speakers. Marissa thinks that if Bose enters the market, her 1998 labor costs would be $859,170, 1998 material costs would be

$702,950, and 1998 overhead costs would be $1,288,750. She believes that her growth rate assumptions would stay the same whether or not Bose enters the market. Add these possible values to your spreadsheet using the IF( ) function in the appropriate cells.

Look at the net profits for 1998 through 2001. In particular, examine the NPV for the two scenarios: Bose does or does not enter the home theater speaker market.

Phase Three: Bringing Uncertainty Into the Model

Jim Allison, the chief of operations at Sound's Alive and a quantitative methods specialist, plays a key role in providing Marissa with estimates for the various revenues and costs. He is uneasy about the basic estimates for the growth rates. For example, although market research indicates that a 9% gross revenue increase per year is reasonable, Jim knows that if this value is 7%, for example, the profit values and the NPV would be quite different. Even more troublesome is a potential tax increase, which would hit Sound's Alive hard. Jim believes that the tax rate could vary around the expected 30% figure. Finally, Jim is uncomfortable with the industry's standard estimate of 18% for the SG&A rate. Jim thinks that this value could be higher or even lower.

The Sound's Alive problem is too complicated for solving with what-if analysis because seven assumed values could change: the growth rates for gross revenues, labor, materials, overhead costs, tax rate, SG&A percent, and whether or not Bose enters the market. Jim believes that a Monte Carlo simulation would be a better approach. Jim thinks that the behavior of these variables can be modeled as follows:

Gross Revenues (%): normally distributed, mean = 9.9, std dev = 1.4
Labor Growth (%): normally distributed, mean = 3.45, std dev = 1.0

Materials (%)	*Probability*
4	0.10
5	0.15
6	0.15
7	0.25
8	0.25
9	0.10

Overhead (%)	*Probability*
2	0.20
3	0.35
4	0.25
5	0.20

Tax Rate (%)	*Probability*
30	0.15
32	0.30
34	0.30
36	0.25

SG&A (%)	*Probability*
15	0.05
16	0.10
17	0.20
18	0.25
19	0.20
20	0.20

Finally, Jim and Marissa agree that there is a 50/50 chance that Bose will enter the market.

Use simulation to analyze the Sound's Alive problem. Based on your results, what is the expected net profit for the years 1998 through 2001, and what is the expected NPV for this business venture?

The Board of Trustees told Marissa that the stockholders would feel comfortable with this business venture if its NPV is at least $5 million. What are the chances that Sound's Alive home theater venture will result in an NPV of $5 million or more?

CASE 12.2 THE FOXRIDGE INVESTMENT GROUP

Inspired by a case written by MBA students Fred Hirsch and Ray Rogers for Professor Larry Weatherford at the University of Wyoming.

The Foxridge Investment Group buys and sells rental income properties in Southwest Virginia. Bill Hunter, president of Foxridge, has asked for your assistance in analyzing a small apartment building the group is interested in purchasing.

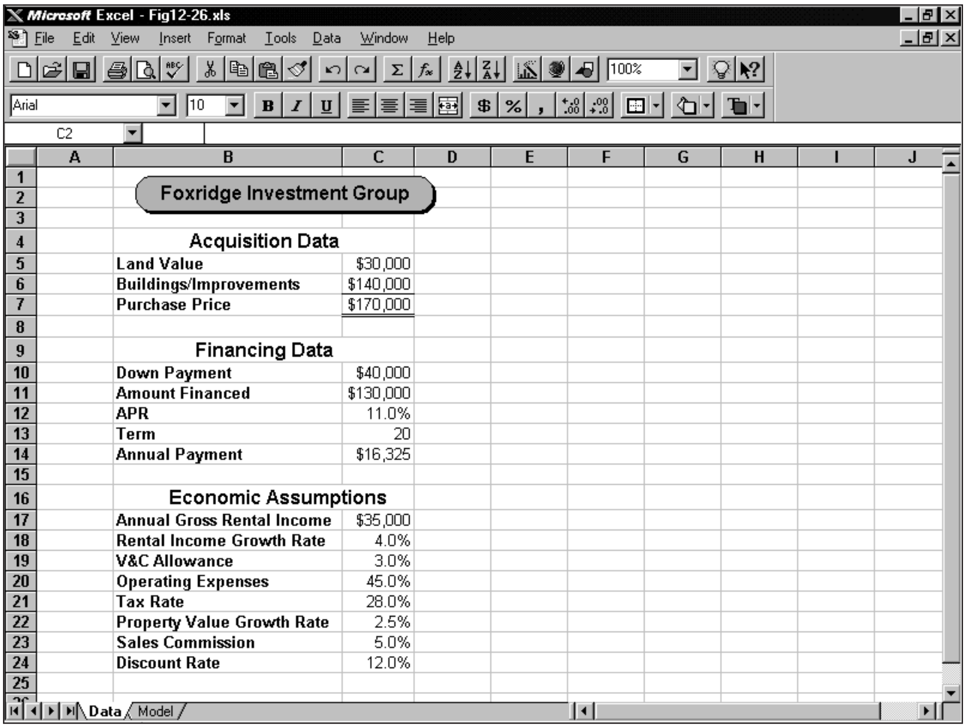
The property in question is a small two-story structure with three rental units on each floor. The purchase price of the property is $170,000, representing $30,000 in land value and $140,000 in buildings and improvements. Foxridge will depreciate the buildings and improvements value on a straight-line basis over 27.5 years. The Foxridge Group will make a down payment of $40,000 to acquire the property and finance the remainder of the purchase price over 20 years with an 11% fixed-rate loan with payments due annually. Figure 12.26 (and the file FIG12-26.xls on your data disk) summarizes this and other pertinent information.

If all units are fully occupied, Mr. Hunter expects the property to generate rental income of $35,000 in the first year and expects to increase the rent at the rate of inflation (currently 4%). Because vacancies occur and some residents may not always be able to pay their rent, Mr. Hunter factors in a 3% vacancy & collection (V&C) allowance against rental income. Operating expenses are expected to be approximately 45% of rental income. The group's marginal tax rate is 28%.

If the group decides to purchase this property, their plan is to hold it for five years and then sell it to another investor. Presently, property values in this area are increasing at a rate of approximately 2.5% per year. The group will have to pay a sales commission of 5% of the gross selling price when they sell the property.

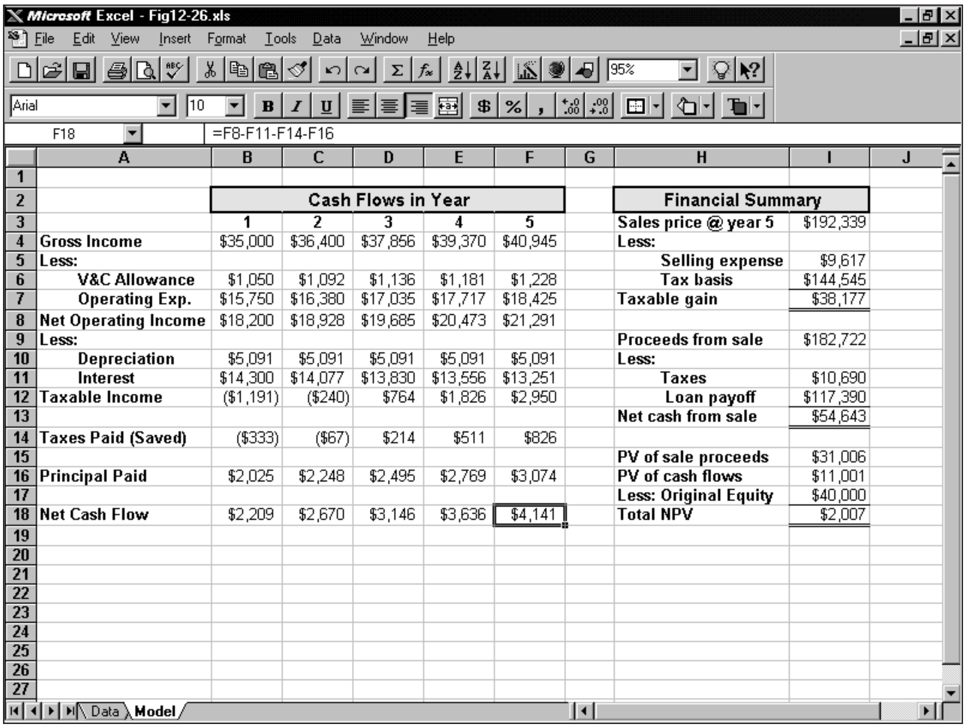
Figure 12.27 shows a spreadsheet model Mr. Hunter developed to analyze this problem. This model first uses the data and assumptions given in Figure 12.26 to generate the expected net cash flows in each of the next five years. It then provides a final summary of the proceeds expected from selling the property at the end of five years. The total net present value (NPV) of the project is then calculated in cell I18 using the discount rate of 12% in cell C24 of Figure 12.26. Thus, after discounting all the

Figure 12.26
Assumptions for the Foxridge Investment Group case.



Foxridge Investment Group	
Acquisition Data	
Land Value	$30,000
Buildings/Improvements	$140,000
Purchase Price	$170,000
Financing Data	
Down Payment	$40,000
Amount Financed	$130,000
APR	11.0%
Term	20
Annual Payment	$16,325
Economic Assumptions	
Annual Gross Rental Income	$35,000
Rental Income Growth Rate	4.0%
V&C Allowance	3.0%
Operating Expenses	45.0%
Tax Rate	28.0%
Property Value Growth Rate	2.5%
Sales Commission	5.0%
Discount Rate	12.0%

Figure 12.27
Cash flow and financial summary for the Foxridge Investment Group case.



	Cash Flows in Year				
	1	2	3	4	5
Gross Income	$35,000	$36,400	$37,856	$39,370	$40,945
Less:					
V&C Allowance	$1,050	$1,092	$1,136	$1,181	$1,228
Operating Exp.	$15,750	$16,380	$17,035	$17,717	$18,425
Net Operating Income	$18,200	$18,928	$19,685	$20,473	$21,291
Less:					
Depreciation	$5,091	$5,091	$5,091	$5,091	$5,091
Interest	$14,300	$14,077	$13,830	$13,556	$13,251
Taxable Income	($1,191)	($240)	$764	$1,826	$2,950
Taxes Paid (Saved)	($333)	($67)	$214	$511	$826
Principal Paid	$2,025	$2,248	$2,495	$2,769	$3,074
Net Cash Flow	$2,209	$2,670	$3,146	$3,636	$4,141

Financial Summary	
Sales price @ year 5	$192,339
Less:	
Selling expense	$9,617
Tax basis	$144,545
Taxable gain	$38,177
Proceeds from sale	$182,722
Less:	
Taxes	$10,690
Loan payoff	$117,390
Net cash from sale	$54,643
PV of sale proceeds	$31,006
PV of cash flows	$11,001
Less: Original Equity	$40,000
Total NPV	$2,007

future cash flows associated with this investment by 12% per year, the investment still generates an NPV of $2,007.

While the group has been using this type of analysis for many years to make investment decisions, one of Mr. Hunter's investment partners recently read an article in *The Wall Street Journal* about risk analysis and simulation using spreadsheets. As a result, the partner realizes there is quite a bit of uncertainty associated with many of the economic assumptions shown in Figure 12.26. After explaining the potential problem to Mr. Hunter, the two have decided to apply simulation to this model before making a decision. Since neither of them know how to do simulation, they have asked for your assistance.

To model the uncertainty in this decision problem, Mr. Hunter and his partner have decided that the growth in rental income from one year to the next could vary uniformly from 2% to 6% years 2 through 5. Similarly, they believe the V&C allowance in any year could be as low as 1% in any year and as high as 5%, with 3% being the most likely outcome. They think the operating expenses in each year should be normally distributed with a mean of 45% and standard deviation of 2% but should never be less than 40% and never greater than 50% of gross income. Finally, they believe the property value growth rate could be as small as 1% or as large as 5%, with 2.5% being the most likely outcome.

1. Revise the spreadsheets shown in Figures 12.26 and 12.27 to reflect the uncertainty outlined above.
2. Construct a 95% confidence interval for the average total NPV the Foxridge Investment Group can expect if they undertake this project. (Use 500 replications.) Interpret this confidence interval.
3. Based on your analysis, what is the probability of this project generating a positive total NPV if the group uses a 12% discount rate?
4. Suppose the investors are willing to buy the property if the expected total NPV is greater then zero. Based on your analysis, should they buy this property?
5. Assume the investors decide to increase the discount rate to 14% and repeat questions 2, 3, and 4.

REFERENCES

Banks, J. and J. Carson. *Discrete-Event Simulation.* Englewood Cliffs, NJ: Prentice Hall, 1984.

Hamzawi, S. "Management and Planning of Airport Gate Capacity: A Microcomputer-Based Gate Assignment Simulation Model." *Transportation Planning and Technology,* vol. 11, 1986.

Kaplan, A. and S. Frazza. "Empirical Inventory Simulation: A Case Study." *Decision Sciences,* vol. 14, January 1983.

Khoshnevis, B. *Discrete Systems Simulation.* New York: McGraw-Hill, 1994.

Law, A. and W. Kelton. *Simulation Modeling and Analysis.* New York: McGraw-Hill, 1990.

Marcus, A. "The Magellan Fund and Market Efficiency." *Journal of Portfolio Management,* Fall 1990.

Russell, R. and R. Hickle. "Simulation of a CD Portfolio." *Interfaces,* vol. 16, no. 3, 1986.

Vollman, T., W. Berry and C. Whybark. *Manufacturing Planning and Control Systems.* Homewood, IL: Irwin, 1987.

Watson, H. *Computer Simulation in Business.* New York: Wiley, 1981.