

Examples of Performing Simulations on Spreadsheets with Analytic Solver Platform

Section 20.6 introduced the subject of how to perform simulations on a spreadsheet while using the powerful Excel add-in, Analytic Solver Platform, developed by Frontline Systems, Inc. This chapter elaborates considerably on the same subject.

Section 20.6 included a complete example in the area of inventory management (Freddie the newsboy's problem). Sections 28.1–28.5 present five additional examples that further illustrate how to formulate simulation models on a spreadsheet for a variety of important applications while applying Analytic Solver Platform for Education (ASPE). Section 28.6 focuses on how to choose the right probability distribution as inputs for a simulation. Section 28.7 then describes how parameter analysis reports and trend charts can be constructed and applied to make a decision about the problem being formulated.

28.1 BIDDING FOR A CONSTRUCTION PROJECT

Managers frequently must make decisions whose outcomes will be greatly affected by the corresponding decisions being made by the management of competitor firms. For example, marketing decisions often fall into this category. To illustrate, consider the case in which a manager must determine the price for a new product being brought to market. How well this decision works out will depend greatly on the pricing decisions being made nearly simultaneously by other firms marketing competitive new products. Similarly, the success of a decision on how soon to market a product under development will be determined largely by whether this product reaches the market before competitive products are released by other firms.

When a decision must be made before learning the corresponding decisions being made by competitors, the analysis needs to take into account the uncertainty surrounding what competitors' decisions will be. Simulation provides a natural way of doing this by using uncertain variable cells to represent competitors' decisions.

The following example illustrates this process by considering a situation where the decision being made is the bid to submit on a construction project while three other companies are simultaneously preparing their own bids.

The Reliable Construction Co. Bidding Problem

The prototype example carried throughout Chap. 22 involves the Reliable Construction Co. and its project to construct a new plant for a major manufacturer. That chapter describes how the project manager (David Perty) made extensive use of PERT/CPM models to help guide his management of the project.

As the opening sentence of Sec. 22.1 indicates, the example in that chapter begins as the company has just made the winning bid of \$5.4 million to do this project. We now will back up in time to describe how the company's management used simulation with ASPE to guide its choice of \$5.4 million as its bid for the project. You will not need to review the presentation in Chap. 22 to follow the current example.

Reliable's first step in this process was to estimate what the company's total cost would be if it were to undertake the project. This was determined to be \$4.55 million. (This amount excludes a penalty for missing the deadline for completion of the project, as well as a bonus for completion well before the deadline, since management considers either event to be relatively unlikely.) There also is an additional cost of approximately \$50,000 for preparing the bid, including estimating the project cost and analyzing the bidding strategies of the competition.

Three other construction companies also were invited to submit bids for this project. All three have been long-standing competitors of the Reliable Construction Co., so the company has had a great deal of experience in observing their bidding strategies. A veteran analyst in the bid preparation office has taken on the task of estimating what bid each of these competitors will submit. Since there is so much uncertainty in this process, the analyst has determined that each of these estimates needs to be in the form of a probability distribution. Competitor 1 is known to use a 30 percent profit margin above the total (direct) cost of a project in setting its bid. However, competitor 1 also is a particularly unpredictable bidder because of an inability to estimate the true costs of a project with much accuracy. Its actual profit margin on past bids has ranged from as low as minus 5 percent to as high as 60 percent. Competitor 2 uses a 25 percent profit margin and is somewhat more accurate than competitor 1 in estimating project costs, but it still has set bids in the past that have missed this profit margin by as much as 15 percent in either direction. On the other hand, competitor 3 is unusually accurate in estimating project costs (as is the Reliable Construction Co.). Competitor 3 also is adept at adjusting its bidding strategy, so it is equally likely to set its profit margin anywhere between 20 and 30 percent, depending on its assessment of the competition, its current backlog of work, and various other factors.

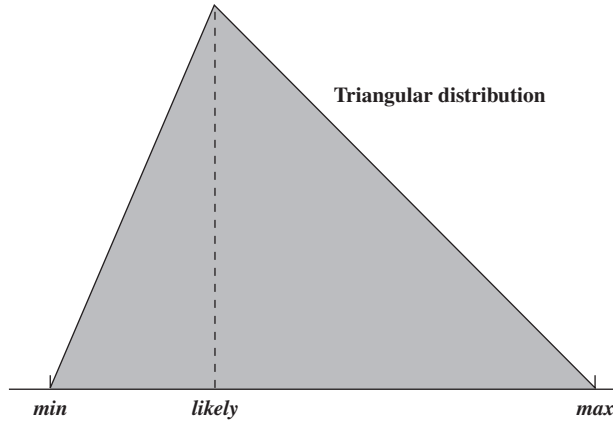
This information about the competitors is invaluable, but the analyst who developed it knew that her work wasn't quite done yet. Based on these numbers, she still needed to develop an estimate of the probability distribution of what the bid will be for each of the competitors.

This task is relatively straightforward in the case of competitor 3. Because the analyst estimates that this competitor is equally likely to set its profit margin anywhere between 20 and 30 percent, its bid then is equally likely to be anywhere between 120 and 130 percent of the total project cost. The probability distribution that fits this is the *uniform distribution* between 120 and 130 percent.

However, this task is not as easy when considering competitors 1 and 2. Fortunately, the analyst has been able to estimate three key numbers for each competitor—a *minimum* value, a *most likely* value, and a *maximum* value—for the profit margin and so (by adding 100 percent) for the bid as a percentage of the total project cost. For example, the analyst has estimated that the bid of competitor 1 (expressed as a percentage of total project cost) has a minimum value of 95 percent, a most likely value of 130 percent, and a maximum value of 160 percent. (The corresponding numbers for competitor 2 are 110 percent, 125 percent, and 140 percent, respectively.) There is a particularly convenient type of probability distribution

■ **FIGURE 28.1**

The shape of a triangular distribution and the location of its three parameters: (1) *min* (the minimum possible value), (2) *likely* (the most likely value), and (3) *max* (the maximum possible value).



called the *triangular distribution* that is based on these same three kinds of numbers. Figure 28.1 shows the shape of a triangular distribution. Its three parameters are *min* (the minimum value), *likely* (the most likely value), and *max* (the maximum value). (Figure 28.1 shows *likely* as being much closer to *min* than to *max*, but it actually can be anywhere between *min* and *max*.) These three parameters are a perfect fit for the distributions of the bids from competitors 1 and 2, so the analyst has chosen a triangular distribution as her best estimate of these distributions. (This is not surprising since triangular distributions are a particularly popular choice for performing simulations.)

In summary, the estimated probability distributions of the bids that the three competitors will submit, expressed as a percentage of Reliable's assessment of the total project cost (\$4.55 million), are as follows.

Competitor 1: A triangular distribution with a minimum value of 95 percent, a most likely value of 130 percent, and a maximum value of 160 percent.

Competitor 2: A triangular distribution with a minimum value of 110 percent, a most likely value of 125 percent, and a maximum value of 140 percent.

Competitor 3: A uniform distribution between 120 percent and 130 percent.

A Spreadsheet Model for Applying Simulation

Figure 28.2 shows the spreadsheet model that has been formulated to evaluate any possible bid that Reliable might submit. Since there is uncertainty about what the competitors' bids will be, this model needs *CompetitorBids* (C8:E8) to be *uncertain variable cells*, so the above probability distributions are entered into these cells. As described in Sec. 20.6, this is done by selecting each cell in turn, choosing the appropriate distribution from the *Distributions* menu on the *ASPE* ribbon (in this case under the *Common* submenu), which brings up the dialog box for that distribution. Figure 28.3 shows the *Triangular Distribution* dialog box that has been used to set the parameter values (*min*, *likely*, and *max*) for competitor 1, and competitor 2 would be handled similarly. These parameter values for competitor 1 come from cells C18:C20, where the parameters in percentage terms (cells C13:C15) have been converted to dollars by multiplying them by *OurProjectCost* (C4). The *Uniform Distribution* dialog box is used instead to set the parameter values for competitor 3 in cell E8.

MinimumCompetitorBid (C23) records the smallest of the competitors' bids for each trial of the simulation. The company wins the bid on a given trial only if the quantity entered into *OurBid* (C25) is less than the smallest of the competitors' bids. The *IF* function entered into *WinBid?* (C27) then returns a 1 if this occurs and a 0 otherwise.

	A	B	C	D	E
1	Reliable Construction Co. Contract Bidding				
2					
3		Data			
4		Our Project Cost (\$million)	4.550		
5		Our Bid Cost (\$million)	0.050		
6					
7		Competitor Bids	Competitor 1	Competitor 2	Competitor 3
8		Bid (\$million)	6.810	5.931	5.771
9					
10		Distribution	<i>Triangular</i>	<i>Triangular</i>	<i>Uniform</i>
11					
12		Competitor Distribution Parameters (Proportion of Our Project Cost)			
13		Minimum	95%	110%	120%
14		Most Likely	130%	125%	
15		Maximum	160%	140%	130%
16					
17		Competitor Distribution Parameters (\$millions)			
18		Minimum	4.323	5.005	5.460
19		Most Likely	5.915	5.688	
20		Maximum	7.280	6.370	5.915
21					
22		Minimum Competitor			
23		Bid (\$million)	5.771		
24					
25		Our Bid (\$million)	5.4		
26					
27		Win Bid?	1	(1=yes, 0=no)	
28					
29		Profit (\$million)	0.800		
30					
31		Mean Profit (\$million)	0.4872		

■ **FIGURE 28.2**

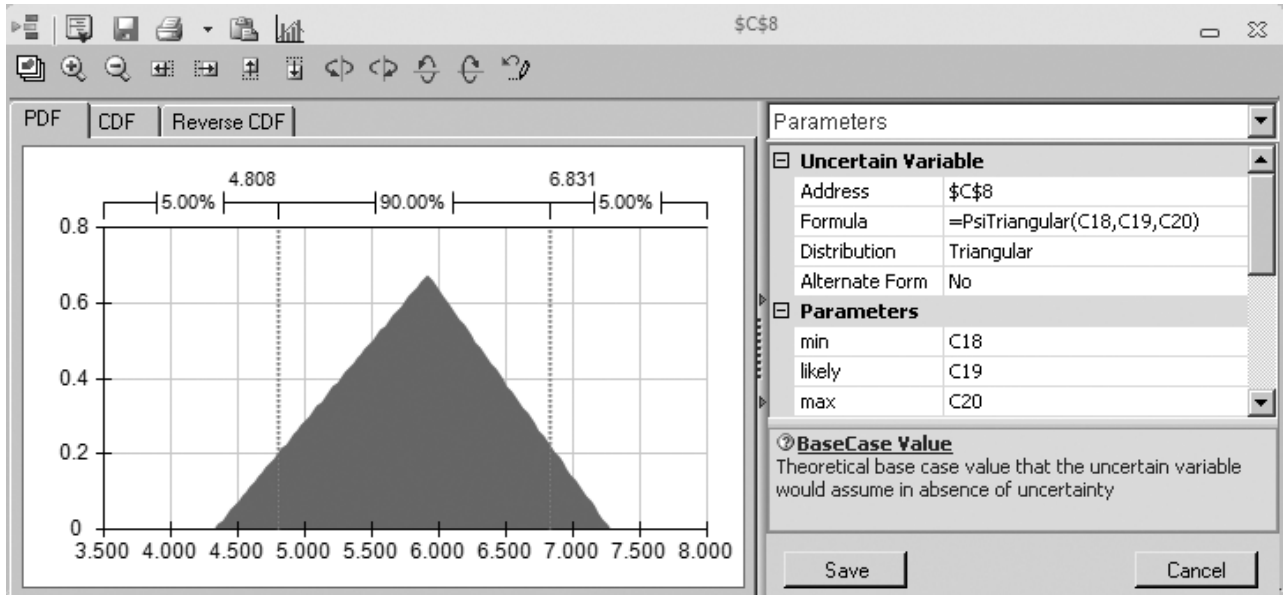
A spreadsheet model for applying simulation to the Reliable Construction Co.'s contract bidding problem. The uncertain variable cells are CompetitorBids (C8:E8), the results cell is Profit (C29), the statistic cell is MeanProfit (C31), and the decision variable is OurBid (C25).

Range Name	Cells
CompetitorBids	C8:E8
MeanProfit	C31
MinimumCompetitorBid	C23
OurBid	C25
OurBidCost	C5
OurProjectCost	C4
Profit	C29
WinBid?	C27

	B	C	D	E
7	Competitor Bids	Competitor 1	Competitor 2	Competitor 3
8	Bid (\$million)	=PsiTriangular(C18,C19,C20)	=PsiTriangular(D18,D19,D20)	=PsiUniform(E18,E20)

	B	C	D	E
18	Minimum	=OurProjectCost*C13	=OurProjectCost*D13	=OurProjectCost*E13
19	Most Likely	=OurProjectCost*C14	=OurProjectCost*D14	
20	Maximum	=OurProjectCost*C15	=OurProjectCost*D15	=OurProjectCost*E15

	B	C
22	Minimum Competitor	
23	Bid (\$million)	=MIN(C8:E8)
24		
25	Our Bid (\$million)	5.4
26		
27	Win Bid?	=IF(OurBid<MinimumCompetitorBid,1,0)
28		
29	Profit (\$million)	=WinBid?*(OurBid-OurProjectCost)-OurBidCost + PsiOutput()
30		
31	Mean Profit (\$million)	=PsiMean(C29)



■ **FIGURE 28.3**

The Triangular Distribution dialog box. It is being used here to enter a triangular distribution with the parameters min 5 C18 (4.323), likely 5 C19 (5.915), and max 5 C20 (7.280) into the uncertain variable cell C8 in the spreadsheet model in Fig. 28.2.

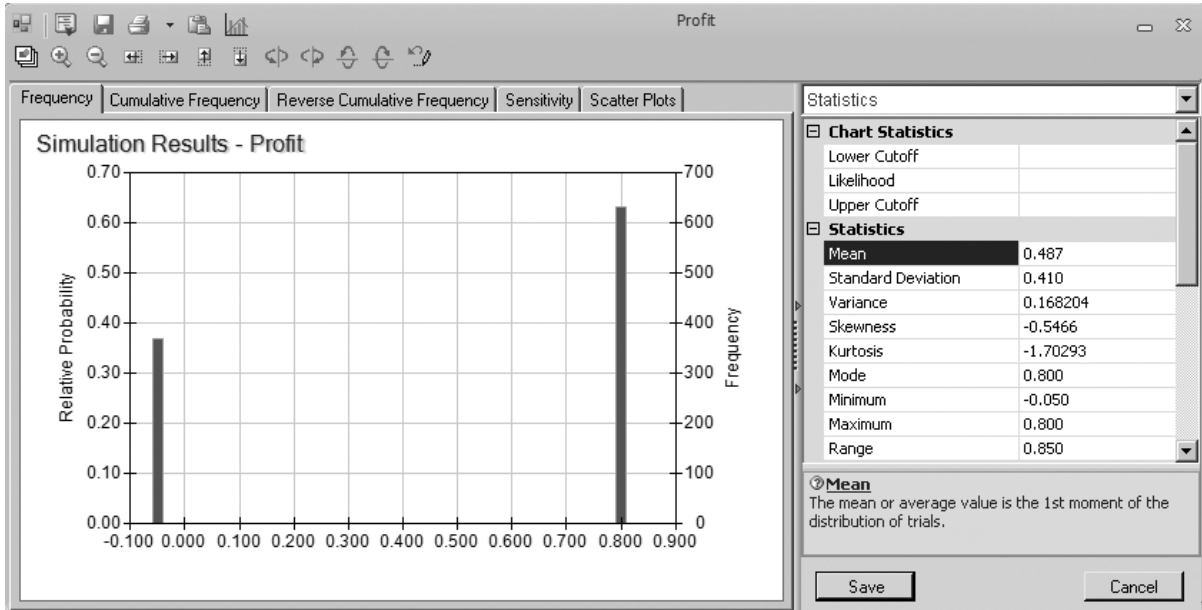
Since management wants to maximize the expected profit from the entire process of determining a bid (if the bid wins) and then doing the project, the results cell in this model is Profit (C29). The profit achieved on a given trial depends on whether the company wins the bid. If not, the profit actually is a loss of \$50,000 (the bid cost). However, if the bid wins, the profit is the amount by which the bid exceeds the sum of the project cost and the bid cost. The equation entered into Profit (C29) performs this calculation for whichever case applies. Profit (C29) is defined as a results cell by clicking on the cell and then choosing Output/In Cell from the Results menu on the ASPE ribbon. Finally, MeanProfit (C31) is defined as a statistic cell by selecting the Profit cell (C29), choosing Mean from the Statistic submenu of the Results menu, and then clicking in cell C31. This will show the mean value of the profit after the simulation is run.

Here is a summary of the key cells in this model.

Uncertain variable cells:	CompetitorBids (C8:E8)
Decision variable:	OurBid (C25)
Results cell:	Profit (C29)
Statistic cell:	MeanProfit (C31) (See Sec. 20.6 for the details regarding how to define these kinds of cells.)

The Simulation Results

To evaluate a possible bid of \$5.4 million entered into OurBid (C25), a simulation of this model ran for 1,000 trials. Figure 28.4 shows the results in the form of a frequency chart and a statistics table. Using units of millions of dollars, the profit on each trial has only two possible values, namely, a loss shown as -0.050 in these figures (if the bid loses) or a profit of 0.800 (if the bid wins). The frequency chart indicates that this loss of \$50,000 occurred on about 380 of the 1,000 trials whereas the profit of \$800,000 occurred on the other 620



■ **FIGURE 28.4**

The frequency chart and statistics table that summarize the results of running the simulation model in Fig. 28.2 for the Reliable Construction Co. contract bidding problem.

trials. This resulted in a mean profit of 0.487 (\$487,000) from all 1,000 trials, as well as the other statistics recorded in the statistics table.

By themselves, these results do not show that \$5.4 million is the best bid to submit. We still need to estimate with additional simulation runs whether a larger expected profit could be obtained with another bid value. Section 28.7 will describe how doing this with a parameter analysis report leads to choosing \$5.4 million as the bid. This turned out to be the winning bid for the Reliable Construction Co., which then led into the prototype example that was analyzed in Chap. 22.

■ 28.2 PROJECT MANAGEMENT

One of the most important responsibilities of a project manager is to meet the deadline that has been set for the project. Therefore, a skillful project manager will revise the plan for conducting the project as needed to ensure a strong likelihood of meeting the deadline. But how does the project manager estimate the probability of meeting the deadline with any particular plan? Section 22.4 described one method provided by PERT/CPM. We now will illustrate how simulation provides a better method.

This example illustrates a common role for simulation—refining the results from a preliminary analysis conducted with approximate mathematical models. You also will get a first look at uncertain variable cells where the values shown are *times*. Another interesting feature of this example is its use of a special kind of ASPE chart called the *sensitivity chart*. This chart will provide a key insight into how the project plan should be revised.

The Problem Being Addressed

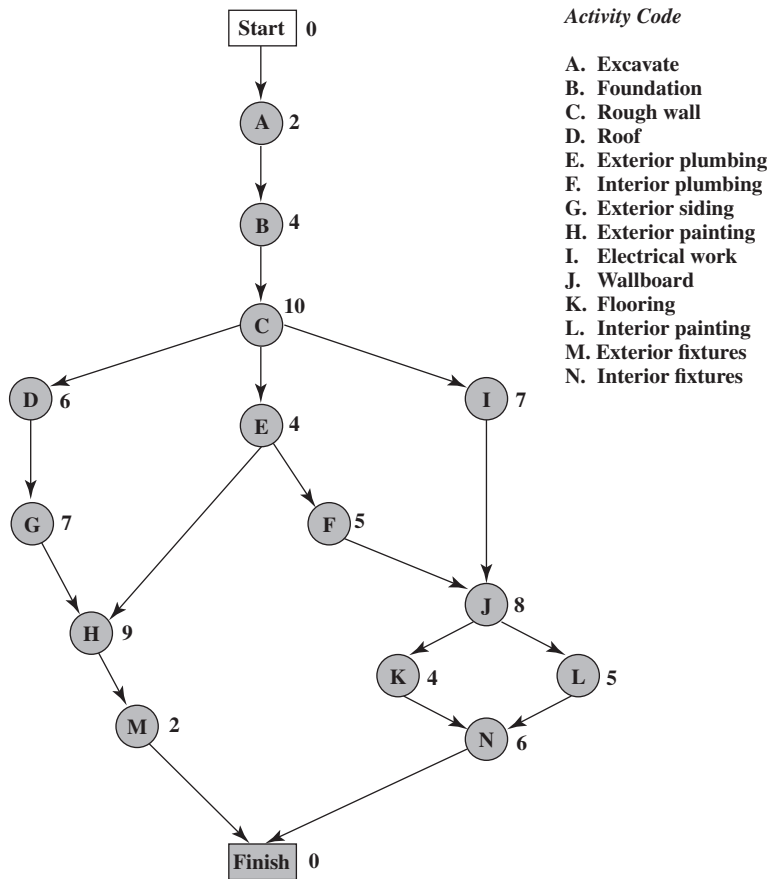
Like the example in the preceding section, this one also revolves around the story of the Reliable Construction Co. that was introduced in Sec. 22.1 and continued throughout Chap 22. However, rather than preceding the part of the story described in Chap. 22, this

example arises in the middle of that story. In particular, Sec. 22.4 discussed how a PERT/CPM procedure was used to obtain a rough approximation of the probability of meeting the deadline for the Reliable Construction Co. project. It then was pointed out that simulation could be used to obtain a better approximation. We now are in a position to describe how this is done.

Here are the essential facts that are needed for the current example. (There is no need for you to refer to Chap. 22 for further details.) The Reliable Construction Company has just made the winning bid to construct a new plant for a major manufacturer. However, the contract includes a large penalty of \$300,000 if construction is not completed by the deadline 47 weeks from now. Therefore, a key element in evaluating alternative construction plans is the *probability of meeting this deadline* under each plan. There are 14 major activities involved in carrying out this construction project, as listed on the right-hand side of Figure 28.5 (which repeats Fig. 22.1 for your convenience). The project network in this figure depicts the precedence relationships between the activities. Thus, there are six sequences of activities (paths through the network), all of which must be completed to finish the project. These six sequences are listed below.

- Path 1:** Start → A → B → C → D → G → H → M → Finish
- Path 2:** Start → S → A → S → B → S → C → S → E → S → H → S → M → S → Finish
- Path 3:** Start → A → B → C → E → F → J → K → N → Finish
- Path 4:** Start → A → B → C → E → F → J → L → N → Finish
- Path 5:** Start → A → B → C → I → J → K → N → Finish
- Path 6:** Start → A → B → C → I → J → L → N → Finish

■ **FIGURE 28.5**
The project network for the Reliable Construction Co. project.



The numbers next to the activities in the project network represent the *estimates* of the number of weeks the activities will take if they are carried out in the normal manner with the usual crew sizes, and so forth. Adding these times over each of the paths (as was done in Table 22.2) reveals that path 4 is the *longest path*, requiring a total of 44 weeks. Since the project is finished as soon as its longest path is completed, this indicates that the project can be completed in 44 weeks, 3 weeks before the deadline.

Now we come to the crux of the problem. The times for the activities in Fig. 28.5 are only estimates, and there actually is considerable uncertainty about what the duration of each activity will be. Therefore, the duration of the entire project could well differ substantially from the estimate of 44 weeks, so there is a distinct possibility of missing the deadline of 47 weeks. What is the *probability* of missing this deadline? To estimate this probability, we need to learn more about the probability distribution of the duration of the project.

This is the reason for the PERT three-estimate approach described in Sec. 22.4. This approach involves obtaining three estimates—a *most likely estimate*, an *optimistic estimate*, and a *pessimistic estimate*—of the duration of each activity. (Table 22.4 lists these estimates for all 14 activities for the project under consideration.) These three quantities are intended to estimate the most likely duration, the minimum duration, and the maximum duration, respectively. Using these three quantities, PERT assumes (somewhat arbitrarily) that the form of the probability distribution of the duration of an activity is a *beta distribution*. By also making three simplifying approximations (described in Sec. 22.4), this leads to an analytical method for roughly approximating the probability of meeting the project deadline.

One key advantage of simulation is that it does not need to make most of the simplifying approximations that may be required by analytical methods. Another is that there is great flexibility about which probability distributions to use. It is not necessary to choose an analytically convenient one.

When dealing with the duration of an activity, simulations commonly use a *triangular distribution* as the distribution of this duration. The triangular distribution fits the PERT three-estimate approach very well because it has three parameters that correspond to the three estimates in a very natural way. Figure 28.1 shows the shape of this distribution and its three parameters—*min* (the minimum possible value), *likely* (the most likely value), and *max* (the maximum possible value). Thus, the duration of an activity is assumed to have a triangular distribution where min 5 optimistic estimate, likely 5 most likely estimate, and max 5 pessimistic estimate. For each *uncertain variable cell* containing this distribution, a Triangular Distribution dialog box (such as the one shown in Fig. 28.3) is used to enter the values of the three estimates by entering their respective cell references into the min, likely, and max boxes.

A Spreadsheet Model for Applying Simulation

Figure 28.6 shows a spreadsheet model for simulating the duration of the Reliable Construction Co. project. The values of *o*, *m*, and *p* in columns D, E, and F are obtained directly from Table 22.4 in Chap. 22. The equations entered into the cells in columns G and I give the start times and finish times for the respective activities. For each trial of the simulation, the maximum of the finish times for the last two activities (M and N) gives the duration of the project (in weeks), which goes into the results cell ProjectCompletion (I21).

Since the activity times generally are variable, the cells H6:H19 all need to be *uncertain variable cells*. Figure 28.7 shows the Triangular Distribution dialog box after it has been used to specify the parameters for the first uncertain variable cell, which records the time of activity A with a range name of ATime (H6). The right side of Fig. 28.7 notes that ASPE has automatically entered a formula (5PsiTriangular(D6,E6,F6)) into ATime (H6) to calculate a random value from this distribution. Rather than repeating this process for all the other uncertain variable cells, it is quicker to simply copy and paste. To copy the

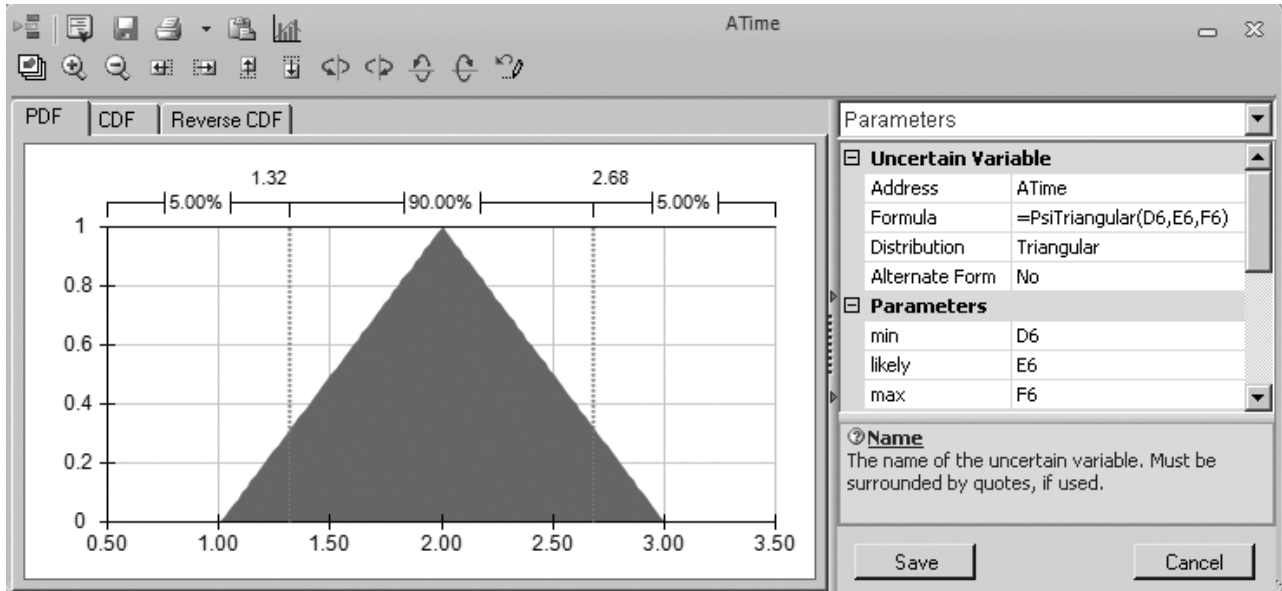
■ FIGURE 28.6

A spreadsheet model for applying simulation to the Reliable Construction Co. project scheduling problem. The uncertain variable cells are cells H6:H15. The results cell is ProjectCompletion (I21). The statistic cell is MeanProjectCompletion (I23).

	A	B	C	D	E	F	G	H	I
1	Simulation of Reliable Construction Co. Project								
2									
3								Activity	
4			Immediate	Time Estimates			Start	Time	Finish
5		Activity	Predecessor	<i>o</i>	<i>m</i>	<i>p</i>	Time	(triangular)	Time
6		A	—	1	2	3	0	2.28	2.28
7		B	A	2	3.5	8	2.28	3.52	5.80
8		C	B	6	9	18	5.80	15.37	21.17
9		D	C	4	5.5	10	21.17	6.86	28.04
10		E	C	1	4.5	5	21.17	4.18	25.35
11		F	E	4	4	10	25.35	5.44	30.80
12		G	D	5	6.5	11	28.04	6.80	34.83
13		H	E, G	5	8	17	34.83	7.33	42.16
14		I	C	3	7.5	9	21.17	3.55	24.72
15		J	F, I	3	9	9	30.80	6.64	37.44
16		K	J	4	4	4	37.44	4.00	41.44
17		L	J	1	5.5	7	37.44	5.45	42.89
18		M	H	1	2	3	42.16	2.47	44.63
19		N	K, L	5	5.5	9	42.89	5.77	48.67
20									
21							Project Completion	48.67	
22									
23							Mean Project Completion	46.26	

Range Name	Cell
AFinish	I6
AStart	G6
ATime	H6
BFinish	I7
BStart	G7
BTime	H7
CFinish	I8
CStart	G8
CTime	H8
DFinish	I9
DStart	G9
DTime	H9
EFinish	I10
EStart	G10
ETime	H10
FFinish	I11
FStart	G11
FTime	H11
GFinish	I12
GStart	G12
GTime	H12
HFinish	I13
HStart	G13
HTime	H13
IFinish	I14
IStart	G14
ITime	H14
JFinish	I15
JStart	G15
JTime	H15
KFinish	I16
KStart	G16
KTime	H16
LFinish	I17
LStart	G17
LTime	H17
MFinish	I18
MeanProjectCompletion	I23
MStart	G18
MTime	H18
NFinish	I19
NStart	G19
NTime	H19
ProjectCompletion	I21

	G	H	I
3		Activity	
4	Start	Time	Finish
5	Time	(triangular)	Time
6	0	=PsiTriangular(D6,E6,F6)	=AStart+ATime
7	=AFinish	=PsiTriangular(D7,E7,F7)	=BStart+BTime
8	=BFinish	=PsiTriangular(D8,E8,F8)	=CStart+CTime
9	=CFinish	=PsiTriangular(D9,E9,F9)	=DStart+DTime
10	=CFinish	=PsiTriangular(D10,E10,F10)	=EStart+ETime
11	=EFinish	=PsiTriangular(D11,E11,F11)	=FStart+FTime
12	=DFinish	=PsiTriangular(D12,E12,F12)	=GStart+GTime
13	=MAX(EFinish,GFinish)	=PsiTriangular(D13,E13,F13)	=HStart+HTime
14	=CFinish	=PsiTriangular(D14,E14,F14)	=IStart+ITime
15	=MAX(FFinish,IFinish)	=PsiTriangular(D15,E15,F15)	=JStart+JTime
16	=JFinish	=PsiTriangular(D16,E16,F16)	=KStart+KTime
17	=JFinish	=PsiTriangular(D17,E17,F17)	=LStart+LTime
18	=HFinish	=PsiTriangular(D18,E18,F18)	=MStart+MTime
19	=MAX(KFinish,LFinish)	=PsiTriangular(D19,E19,F19)	=NStart+NTime
20			
21		Project Completion	=MAX(MFinish,NFinish) + PsiOutput()
22			
23		Mean Project Completion	=PsiMean(I21)



■ **FIGURE 28.7**

A triangular distribution with parameters D6 (51), E6 (52), and F6 (53) is being entered into the first uncertain variable cell ATime (H6) in the spreadsheet model in Fig. 28.6.

formula in H6 down to H7 through H19, select cell H6 and drag the fill handle (the small box on the lower right corner of the cell cursor) down to cell H19. This copies the formula in H6 ($\text{PsiTriangular}(D6,E6,F6)$), used by ASPE to calculate a random value from the triangular distribution, with parameters min 5D6, likely 5E6, and max 5F6) into cells H7 through H19. Since the parameters in cell H6 (D6, E6, and F6) are relative references, the row numbers of the parameters will update appropriately to refer to the data in the correct rows during the copy-and-paste process. For example, the formula in H7 will update to $\text{PsiTriangular}(D7,E7,F7)$.

Here is a summary of the key cells in this model.

Uncertain variable cells: Cells H6:H15

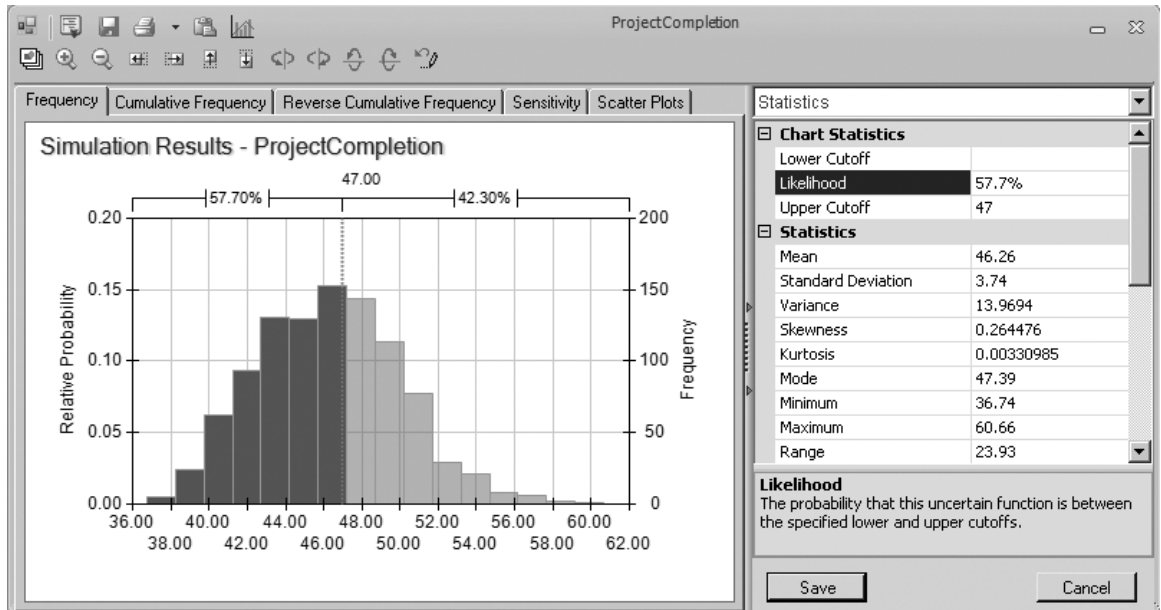
Results cell: ProjectCompletion (I21)

Statistic cell: MeanProjectCompletion (I23)

(See Sec. 20.6 for the details regarding how to define uncertain variable cells, results cells, and statistic cells.)

The Simulation Results

We now are ready to evaluate the simulation of the spreadsheet model in Fig. 28.6. After running a simulation of 1,000 trials, Fig. 28.8 shows the results in the form of a frequency chart and a statistics table. These results show a very wide range of possible project durations. Out of the 1,000 trials, the statistics table indicates that one trial had a duration as short as 36.74 weeks while another was as long as 60.66 weeks. The frequency chart indicates that the duration that occurred most frequently during the 1,000 trials is close to 47 weeks (the project deadline), but that many other durations up to a few weeks either shorter or longer than this also occurred with considerable frequency. The mean is 46.26 weeks, which is much too close to the deadline of 47 weeks to leave much margin for slippage in the project schedule.



■ **FIGURE 28.8**

The frequency chart and statistics table that summarize the results of running the simulation model in 28.6 for the Reliable Construction Co. project scheduling problem.

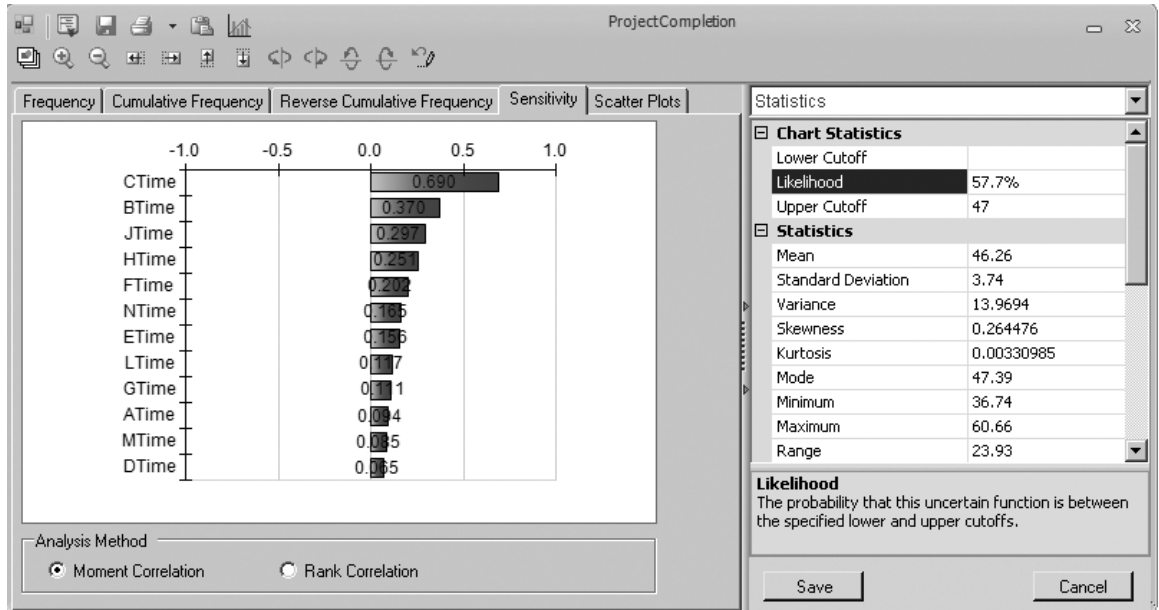
A statistic of special interest to Reliable's management is the probability of meeting the deadline of 47 weeks under the current project plan (Remember that the contract includes a severe penalty of \$300,000 for missing this deadline.). Figure 28.8 shows that all you need to do to identify the exact percentage is to type the deadline of 47 in the Upper Cutoff box. The Likelihood box then reveals that about 57.7 percent of the trials met the deadline.

If the simulation run were to be repeated with another 1,000 trials, this percentage probably would change a little. However, with such a large number of trials, the difference in the percentages should be slight. Therefore, the probability of 0.577 provided by the Likelihood box in Fig. 28.8 is a close estimate of the true probability of meeting the deadline under the assumptions of the spreadsheet model in Fig. 28.6. Note how much smaller this relatively precise estimate is than the rough estimate of 0.84 obtained by the PERT three-estimate approach in Sec. 22.4. Thus, the simulation estimate provides much better guidance to management in deciding whether the project plan should be changed to improve the chances of meeting the deadline. This illustrates how useful simulation can be in refining the results obtained by approximate analytical results.

A Key Insight Provided by the Sensitivity Chart

Given such a low probability (0.577) of meeting the project deadline, Reliable's project manager (David Perty) will want to revise the project plan to improve the probability substantially. ASPE has another tool, called the *sensitivity chart*, that provides strong guidance in identifying which revisions in the project plan would be most beneficial.

To view a sensitivity chart after running a simulation, click on the Sensitivity tab above the chart for the results cell. This reveals a sensitivity chart, as shown in Fig. 28.9. Using range names, the left side of the chart identifies various uncertain variable cells (activity times) in column H of the spreadsheet model in Figure 28.6.



■ **FIGURE 28.9**

This sensitivity chart shows how strongly various activity times in the Reliable Construction Co. project are influencing the project completion time.

The bars in the chart give the correlation coefficient (based on moment values) between each uncertain variable cell and the results cell. A correlation coefficient between two variables measures the strength of the relationship between those variables. Thus, each correlation coefficient in Fig. 28.9 measures how strongly that activity time is influencing the project completion time. The higher the correlation coefficient, the stronger is this influence. Therefore, the activities with the highest correlation coefficients are those where the greatest effort should be made to reduce their activity times.

Figure 28.9 indicates that CTime has a far higher correlation coefficient than the times for any of the other activities. An examination of Figs. 28.5 and 28.6 suggests why. Figure 28.5 shows that activity C precedes *all* the other activities except activities A and B, so any delay in completing activity C would delay the start time for all these other activities. Furthermore, cells D8:F8 in Fig. 28.6 indicate that CTime is highly variable, with an unusually large spread of 9 weeks between its most likely estimate and its pessimistic estimate, so long delays beyond the most likely estimate may well occur.

This very high correlation coefficient for CTime suggest that the best way to reduce the project completion time (and its variability) is to focus on reducing this activity time (and its variability). This can be accomplished by revising the project plan to assign activity C more personnel, better equipment, stronger supervision, and so forth. ASPE's sensitivity chart clearly highlights this insight into where the project plan needs to be revised.

■ 28.3 CASH FLOW MANAGEMENT

Many applications of simulation involve scenarios that evolve far into the future. Since nobody can predict the future with certainty, simulation is needed to take future uncertainties into account when making decisions. For example, businesses typically have great uncertainty about what their future cash flows will be. An attempt often is made to predict

these future cash flows as a first step toward making decisions about what should be done (e.g., arranging for loans) to meet cash flow needs. However, effective cash management requires going a step further to consider the effect of the uncertainty in the future cash flows. This is where simulation comes in, with uncertain variable cells being used for the cash flows in various future periods. This process is illustrated by the following example.

The Everglade Cash Flow Management Problem

The case study analyzed in Chap. 21 involves the Everglade Golden Years Company (which operates upscale retirement communities) and its efforts to manage its cash flow problems. In particular, because of both a temporary decline in business and some current or future construction costs, the company is facing some negative cash flows in the next few years as well as in some more distant years. As first provided in Table 21.1, Table 28.1 shows the projected net cash flows over the next 10 years (2014 to 2023). The company has some new retirement communities opening within the 10 years, so it is anticipated (or at least hoped) that a large positive cash flow will occur in 2023. Therefore, the problem confronting Everglade management is how to best arrange Everglade's financing to tide the company over until its investments in new retirement communities can start to pay off.

Chapter 21 describes how a decision was made to combine taking a long-term (10-year) loan now (the beginning of 2014) and a series of short-term (1-year) loans as needed to maintain a positive cash balance of at least \$500,000 (as dictated by company policy) throughout the 10 years. Assuming no deviation from the projected cash flows shown in Table 28.1, linear programming was used to optimize the size of both the long-term loan and the short-term loans so as to maximize the company's cash balance at the beginning of 2024 when all of the loans have been paid off. Figure 21.5 in Chap. 21 shows the complete spreadsheet model after using Solver to obtain the optimal solution. For your convenience, Figure 21.5 is repeated here as Fig. 28.10. The changing cells, LTLoan (D11) and STLoan (E11:E20), give the sizes of the long-term loan and the short-term loans at the beginning of the various years. The objective cell EndBalance (J21) indicates that the resulting cash balance at the end of the 10 years (the beginning of 2024) would be \$5.39 million. Since this is the cell that is being maximized, any other plan for the sizes of the loans would result in a smaller cash balance at the end of the 10 years.

Obtaining the "optimal" financing plan presented in Fig. 28.10 is an excellent first step in developing a final plan. However, the drawback of the spreadsheet model in Fig. 28.10 is that it makes no allowance for the inevitable deviations from the projected cash flows shown in Table 28.1. The actual cash flow for the first year (2014) probably will turn out to be quite close to the projection. However, it is difficult to predict the cash flows in even the second and third years with much accuracy, let alone up to 10 years into the future. Simulation is needed to assess the effect of these uncertainties.

■ **TABLE 28.1** Projected Net Cash Flows for the Everglade Golden Years Company over the Next 10 Years

Year	Projected Net Cash Flow (millions of dollars)
2014	-8
2015	-2
2016	-4
2017	3
2018	6
2019	3
2020	-4
2021	7
2022	-2
2023	10

FIGURE 28.10

The spreadsheet model that used linear programming in Chap. 21 (Fig. 21.5) to analyze the Everglade Golden Years Company cash flow management problem without taking the uncertainty in future cash flows into account.

A	B	C	D	E	F	G	H	I	J	K	L
1	Everglade Cash Flow Management Problem When Applying Linear Programming										
2											
3	LT Rate	5%									
4	ST Rate	7%									
5											
6	Start Balance	1									
7	Minimum Cash	0.5									
8											
9	Cash		LT	ST	LT	ST	LT	ST	LT	ST	Minimum
10	Flow		Loan	Loan	Interest	Interest	Payback	Payback	Balance	Balance	Balance
11	2014	-8	4.65	2.85	-0.23	-0.20	-2.85	0.50	≥	0.5	0.5
12	2015	-2		5.28	-0.23	-0.37	-5.28	0.50	≥	0.5	0.5
13	2016	-4		9.88	-0.23	-0.69	-9.88	0.50	≥	0.5	0.5
14	2017	3		7.81	-0.23	-0.55	-7.81	0.50	≥	0.5	0.5
15	2018	6		2.59	-0.23	-0.18	-2.59	0.50	≥	0.5	0.5
16	2019	3		0	-0.23	0	0	0.50	≥	0.5	0.5
17	2020	-4		4.23	-0.23	0	0	0.50	≥	0.5	0.5
18	2021	7		0	-0.23	-0.30	-4.23	2.74	≥	0.5	0.5
19	2022	-2		0	-0.23	0	0	0.51	≥	0.5	0.5
20	2023	10		0	-0.23	0	0	10.27	≥	0.5	0.5
21	2024			0	-0.23	0	-6.65	0	5.39	≥	0.5

(all cash figures in millions of dollars)

Range Name	Cells
CashFlow	C11:C20
EndBalance	J21
EndingBalance	J11:J21
LTLoss	D11
LTRate	C3
MinimumBalance	L11:L21
MinimumCash	C7
StartBalance	C6
STLoss	E11:E20
STRate	C4

Solver Parameters
Set Objective Cell: EndBalance
To: Max
By Changing Variable Cells: LTLoss, STLoss
Subject to the Constraints: EndingBalance >= MinimumBalance
Solver Options: Make Variables Nonnegative
 Solving Method: Simplex LP

	F	G	H	I	J	K	L
9	LT	ST	LT	ST	Ending	Minimum	
10	Interest	Interest	Payback	Payback	Balance	Balance	
11							
12	=LTRate*LTLoss	=STRate*E11		=E11	=StartBalance+SUM(C11:111)	≥	=MinimumCash
13	=LTRate*LTLoss	=STRate*E12		=E12	=J11+SUM(C12:112)	≥	=MinimumCash
14	=LTRate*LTLoss	=STRate*E13		=E13	=J12+SUM(C13:113)	≥	=MinimumCash
15	=LTRate*LTLoss	=STRate*E14		=E14	=J13+SUM(C14:114)	≥	=MinimumCash
16	=LTRate*LTLoss	=STRate*E15		=E15	=J14+SUM(C15:115)	≥	=MinimumCash
17	=LTRate*LTLoss	=STRate*E16		=E16	=J15+SUM(C16:116)	≥	=MinimumCash
18	=LTRate*LTLoss	=STRate*E17		=E17	=J16+SUM(C17:117)	≥	=MinimumCash
19	=LTRate*LTLoss	=STRate*E18		=E18	=J17+SUM(C18:118)	≥	=MinimumCash
20	=LTRate*LTLoss	=STRate*E19		=E19	=J18+SUM(C19:119)	≥	=MinimumCash
21	=LTRate*LTLoss	=STRate*E20	=-LTLoss	=E20	=J19+SUM(C20:120)	≥	=MinimumCash
					=J20+SUM(C21:121)	≥	=MinimumCash

A Spreadsheet Model for Applying Simulation

Figure 28.11 shows the modification of the spreadsheet model in Fig. 28.10 that is needed to apply simulation. One key difference is that the constants in CashFlow (C11:C20) in Fig. 28.10 have turned into random inputs in CashFlow (F12:F21) in Fig. 28.11. Thus, the latter cells, CashFlow (F12:F21), are *uncertain variable cells*. (The numbers appearing in these cells are just one possible random outcome—the last trial of the simulation run.) As indicated in cells D9:E9, the assumption has been made that each of the cash flows has a triangular distribution. Estimates have been made of the three parameters of this distribution (min, likely, and max) for each of the years, as presented in cells C12:E21.

The number 4.65 entered into LTLoan (G12) is the size of the long-term loan (in millions of dollars) that was obtained in Fig. 28.10. However, because of the variability in the cash flows, it no longer makes sense to lock in the sizes of the short-term loans that were obtained in STLoan (E11:E20) in Fig. 28.10. It is better to be flexible and adjust these sizes based on the actual cash flows that occur in the preceding years. If the balance at the beginning of a year (as calculated in BalanceBeforeSTLoan [L12:L22]) already exceeds the required minimum balance of \$0.50 million, then there is no need to take any short-term loan at that point. However, if the balance is not this large, then a sufficiently large short-term loan should be taken to bring the balance up to \$0.50 million. This is what is done by the equations entered into STLoan (M12:M22) that are shown at the bottom of Fig. 28.11.

The objective cell EndBalance (J21) in Fig. 28.10 becomes the results cell EndBalance (N22) in Fig. 28.11. A statistic cell, MeanEndBalance (N24), is defined to determine the mean value of EndBalance for the simulation run. On any trial of the simulation, if the simulated cash flows in CashFlow (F12:F21) in Fig. 28.11 are more favorable than the projected cash flows given in Table 28.1 (as is the case for the current numbers in Fig. 28.11), then EndBalance (N22) in Fig. 28.11 would be larger than EndBalance (J21) in Fig. 28.10. However, if the simulated cash flows are less favorable than the projections, then EndBalance (N22) in Fig. 28.11 might even be a negative number. For example, if all the simulated cash flows are close to the corresponding minimum values given in cells C12:C21, then the required short-term loans will become so large that paying off the last one at the beginning of 2024 (along with paying off the long-term loan then) will result in a very large negative number in EndBalance (N22). This would spell serious trouble for the company. Simulation will reveal the relative likelihood of this occurring versus a favorable outcome.

Here is a summary of the key cells in this model:

Uncertain variable cells: CashFlow (F12:F21)

Results cell: EndBalance (N22)

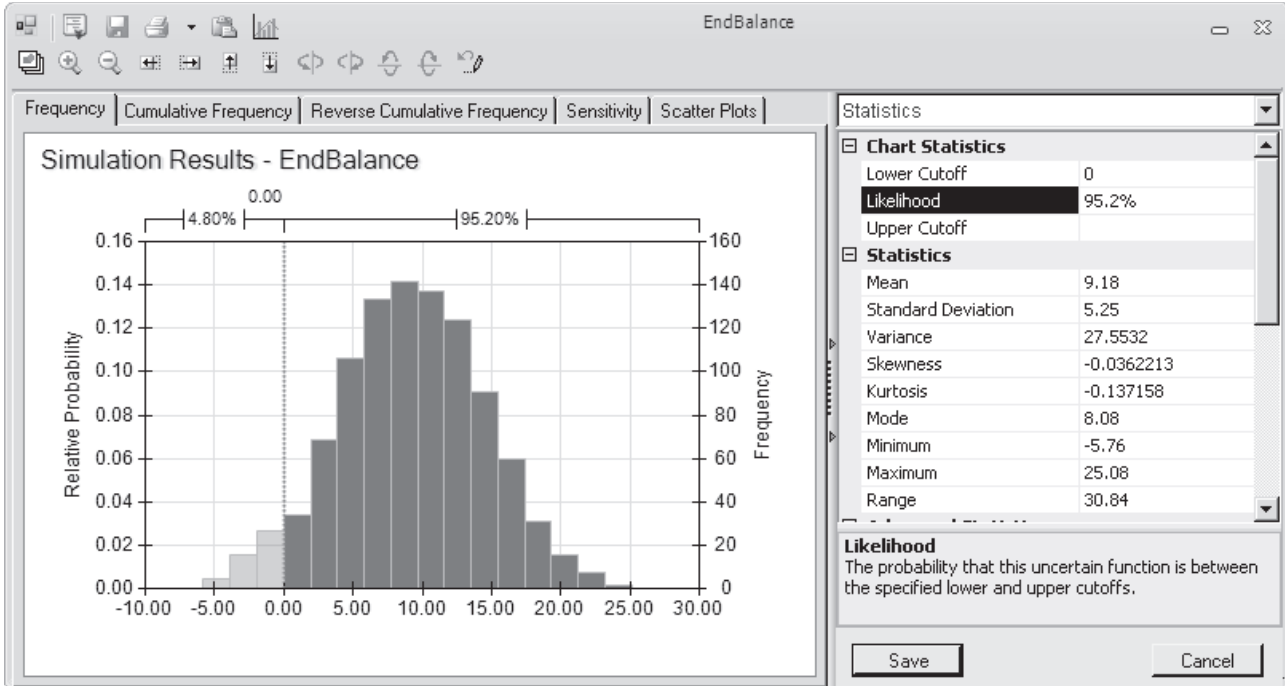
Statistic cell: MeanEndBalance (N24)

(See Sec. 20.6 for the details regarding how to define uncertain variable cells, results cells, and statistic cells.)

The Simulation Results

Figure 28.12 shows the results from applying simulation with 1,000 trials. Because Everglade management is particularly interested in learning how likely it is that the current financing plan would result in a positive cash balance at the end of the 10 years, the number 0 has been entered into the Lower Cutoff box in the statistics table. The Likelihood box then indicates that over 95 percent of the trials resulted in a positive cash balance at the end. Furthermore, the frequency chart shows that many of these positive cash balances are reasonably large, with many exceeding \$10 million. The overall mean is \$9.18 million.

On the other hand, it is worrisome that nearly 5 percent of the trials resulted in a negative cash balance at the end. Although huge losses were rare, some of these negative cash balances were quite significant, ranging up to \$5 million.



■ **FIGURE 28.12**

The frequency chart and statistics table that summarize the results of running the simulation model in Fig. 28.11 for the Everglade Golden Years cash flow management problem.

Conclusions

Everglade management is pleased that the simulation results indicate that the proposed financing plan is likely to lead to a favorable outcome at the end of the 10 years. At the same time, management feels that it would be prudent to take steps to reduce the 5 percent chance of an unfavorable outcome.

One possibility would be to increase the size of the long-term loan, since this would reduce the sizes of the higher interest short-term loans that would be needed in the later years if the cash flows are not as good as currently projected. This possibility is investigated in Problem 28.9.

The scenarios that would lead to a negative cash balance at the end of the 10 years are those where the company's retirement communities fail to achieve full occupancy because of overestimating the demand for this service. Therefore, Everglade management concludes that it should take a more cautious approach in moving forward with its current plans to build more retirement communities over the next 10 years. In each case, the final decisions regarding the start date for construction and the size of the retirement community should be made only after obtaining and carefully assessing a detailed forecast of the trends in the demand for this service.

After adopting this policy, Everglade management approves the financing plan that is incorporated into the spreadsheet model in Fig. 28.11. In particular, a 10-year loan of \$4.65 million will be taken now (the beginning of 2014). In addition, a one-year loan will be taken at the beginning of each year from 2014 to 2023 if it is needed to bring the cash balance for that year up to the level of \$500,000 required by company policy.

28.4 FINANCIAL RISK ANALYSIS

One of the earliest areas of application of simulation, dating back to the 1960s, was *financial risk analysis*. This continues today to be one of the most important areas of application.

When assessing any financial investment (or a portfolio of investments), the key trade-off is between the *return* from the investment and the *risk* associated with the investment. Of these two quantities, the less difficult one to determine is the return that would be obtained if everything evolves as currently projected. However, assessing the risk is relatively difficult. Fortunately, simulation is ideally suited to perform this risk analysis by obtaining a **risk profile**, namely, a *frequency distribution* of the return from the investment. The portion of the frequency distribution that reflects an unfavorable return clearly describes the risk associated with the investment.

The following example illustrates this approach in the context of real estate investments. Like the Everglade example in the preceding section, you will see simulation being used to refine prior analysis done by linear programming because this prior analysis was unable to take the uncertainty in future cash flows into account.

The Think-Big Financial Risk Analysis Problem

The Think-Big Development Co. is a major investor in commercial real estate development projects. It has been considering taking a share in three large construction projects—a high-rise office building, a hotel, and a shopping center. In each case, the partners in the project would spend three years with the construction, then retain ownership for three years while establishing the property, and then sell the property in the seventh year. By using estimates of expected cash flows, as well as constraints on the amounts of investment capital available both now and over the next three years, linear programming has been applied to obtain the following proposal for how big a share Think-Big should take in each of these projects:

Proposal

Do not take any share of the high-rise building project.

Take a 16.50 percent share of the hotel project.

Take a 13.11 percent share of the shopping center project.

This proposal is estimated to return a *net present value* (NPV) of \$18.11 million to Think-Big.

However, Think-Big management understands very well that such decisions should not be made without taking risk into account. These are very risky projects since it is unclear how well these properties will compete in the marketplace when they go into operation in a few years. Although the construction costs during the first three years can be estimated fairly roughly, the net incomes during the following three years of operation are very uncertain. Consequently, there is an extremely wide range of possible values for each sale price in year 7. Therefore, management wants *risk analysis* to be performed in the usual way (with simulation) to obtain a *risk profile* of what the total NPV might actually turn out to be with this proposal.

To perform this risk analysis, Think-Big staff now has devoted considerable time to estimating the amount of uncertainty in the cash flows for each project over the next seven years. These data are summarized in Table 28.2 (in units of millions of dollars) for a 100 percent share of each project. Thus, when taking a smaller percentage share of a project, the numbers in the table should be reduced proportionally to obtain the relevant numbers

for Think-Big. In years 1 through 6 for each project, the probability distribution of cash flow is assumed to be a *normal distribution*, where the first number shown is the estimated *mean* and the second number is the estimated *standard deviation* of the distribution. In year 7, the income from the sale of the property is assumed to have a *uniform distribution* over the range from the first number shown to the second number shown.

To compute NPV, a cost of capital of 10 percent per annum is being used. Thus, the cash flow in year n is divided by $(1.1)^n$ before adding these discounted cash flows to obtain NPV.

A Spreadsheet Model for Applying Simulation

A spreadsheet model has been formulated for this problem in Fig. 28.13. There is no uncertainty about the immediate (year 0) cash flows appearing in cells D6 and D16, so these are data cells. However, because of the uncertainty for years 1–7, cells D7:D13 and D17:D23 containing the simulated cash flows for these years need to be uncertain variable cells. (The numbers in these cells in Fig. 28.13 represent one possible random outcome—the last trial of the simulation run.) Table 28.2 specifies the probability distributions and their parameters that have been estimated for these cash flows, so the form of the distributions has been recorded in cells E7:E13 and E17:E23 while entering the corresponding parameters in cells F7:G13 and F17:G23. Figure 28.14 shows the Normal Distribution dialog box that is used to enter the parameters (mean and standard deviation) for the normal distribution into the first uncertain variable cell D7 by referencing cells F7 and G7. The formula in D7 is then copied and pasted into cells D8:D12 and D17:D22 to define these uncertain variable cells. The Uniform Distribution dialog box (like the similar one displayed earlier in Fig. 20.9 for an integer uniform distribution) is used in a similar way to enter the parameters (minimum and maximum) for this kind of distribution into the uncertain variable cells D13 and D23.

The simulated cash flows in cells D6:D13 and D16:D23 are for 100 percent of the hotel project and the shopping center project, respectively, so Think-Big's share of these cash flows needs to be reduced proportionally based on its shares in these projects. The proposal being analyzed is to take the shares shown in cells H28:H29. The equations entered into cells D28:D35 (see the bottom of Fig. 28.13) then gives Think-Big's total cash flow in the respective years for its share of the two projects.

Think-Big's management wants to obtain a risk profile of what the total net present value (NPV) might be with this proposal. Therefore, the results cell is NetPresent Value (D37). To show the mean NPV over the simulation run, MeanNPV (D39) is defined as a statistic cell.

■ TABLE 28.2 The Estimated Cash Flows for 100 Percent of the Hotel and Shopping Center Projects

Hotel Project		Shopping Center Project	
Year	Cash Flow (\$1,000,000s)	Year	Cash Flow (\$1,000,000s)
0	–80	0	–90
1	Normal (–80, 5)	1	Normal (–50, 5)
2	Normal (–80, 10)	2	Normal (–20, 5)
3	Normal (–70, 15)	3	Normal (–60, 10)
4	Normal (+30, 20)	4	Normal (+15, 15)
5	Normal (+40, 20)	5	Normal (+25, 15)
6	Normal (+50, 20)	6	Normal (+40, 15)
7	Uniform (+200, 844)	7	Uniform (160, 600)

	A	B	C	D	E	F	G	H
1	Simulation of Think-Big Development Co. Problem							
2								
3				Project Simulated				
4				Cash Flow				
5	Hotel Project:			(\$millions)				
6	Construction Costs:	Year 0		-80				
7		Year 1		-79.057	Normal	-80	5	(mean, st. dev.)
8		Year 2		-80.343	Normal	-80	10	(mean, st. dev.)
9		Year 3		-73.063	Normal	-70	15	(mean, st. dev.)
10	Revenue per Share	Year 4		14.059	Normal	30	20	(mean, st. dev.)
11		Year 5		29.746	Normal	40	20	(mean, st. dev.)
12		Year 6		81.373	Normal	50	20	(mean, st. dev.)
13	Selling Price per Share	Year 7		395.247	Uniform	200	844	(lower,upper)
14								
15	Shopping Center Project							
16	Construction Costs:	Year 0		-90				
17		Year 1		-42.329	Normal	-50	5	(mean, st. dev.)
18		Year 2		-15.124	Normal	-20	5	(mean, st. dev.)
19		Year 3		-54.653	Normal	-60	10	(mean, st. dev.)
20	Revenue per Share	Year 4		21.923	Normal	15	15	(mean, st. dev.)
21		Year 5		10.122	Normal	25	15	(mean, st. dev.)
22		Year 6		14.780	Normal	40	15	(mean, st. dev.)
23	Selling Price per Share	Year 7		494.378	Uniform	160	615	(lower,upper)
24								
25				Think-Big's				
26				Simulated Cash Flow				
27				(\$millions)				Share
28		Year 0		-24.999			Hotel	16.50%
29		Year 1		-18.594			Shopping Center	13.11%
30		Year 2		-15.239				
31		Year 3		-19.221			Cost of Capital	10%
32		Year 4		5.194				
33		Year 5		6.235				
34		Year 6		15.364				
35		Year 7		130.029				
36								
37	Net Present Value (\$millions)			13.879				
38								
39	MeanNPV (\$millions)			18.120				

	C	D
3		Project Simulated
4		Cash Flow
5		(\$millions)
6	Year 0	-80
7	Year 1	=PsiNormal(F7,G7)
8	Year 2	=PsiNormal(F8,G8)
9	Year 3	=PsiNormal(F9,G9)
10	Year 4	=PsiNormal(F10,G10)
11	Year 5	=PsiNormal(F11,G11)
12	Year 6	=PsiNormal(F12,G12)
13	Year 7	=PsiUniform(F13,G13)
14		
15		
16	Year 0	-90
17	Year 1	=PsiNormal(F17,G17)
18	Year 2	=PsiNormal(F18,G18)
19	Year 3	=PsiNormal(F19,G19)
20	Year 4	=PsiNormal(F20,G20)
21	Year 5	=PsiNormal(F21,G21)
22	Year 6	=PsiNormal(F22,G22)
23	Year 7	=PsiUniform(F23,G23)

	C	D
25		Think Big's
26		Simulated Cash Flow
27		(\$millions)
28	Year 0	=HotelShare*D6+ShoppingCenterShare*D16
29	Year 1	=HotelShare*D7+ShoppingCenterShare*D17
30	Year 2	=HotelShare*D8+ShoppingCenterShare*D18
31	Year 3	=HotelShare*D9+ShoppingCenterShare*D19
32	Year 4	=HotelShare*D10+ShoppingCenterShare*D20
33	Year 5	=HotelShare*D11+ShoppingCenterShare*D21
34	Year 6	=HotelShare*D12+ShoppingCenterShare*D22
35	Year 7	=HotelShare*D13+ShoppingCenterShare*D23
36		
37	Net Present Value (\$millions)	=CashFlowYear0+NPV(CostOfCapital,CashFlowYear1To7) +PsiOutput()
38		
39	Mean NPV (\$millions)	=PsiMean(D37)

Range Name	Cells
CashFlowYear0	D28
CashFlowYear1To7	D29:D35
CostOfCapital	H31
HotelShare	H28
MeanNPV	D39
NetPresentValue	D37
ShoppingCenterShare	H29

■ **FIGURE 28.13**
 A spreadsheet model for applying simulation to the Think-Big Development Co. financial risk analysis problem. The uncertain variable cells are cells D7:D13 and D17:D23, the results cell is NetPresentValue (D37), the statistic cell is MeanNPV (D39), and the decision variables are HotelShare (H28) and ShoppingCenterShare (H29).

Here is a summary of the key cells in this model:

Uncertain variable cells: Cells D7:D13 and D17:D23

Decision variables: HotelShare (H28) and ShoppingCenterShare (H29)

Results cell: NetPresentValue (D37)

Statistic cell: MeanNPV (D39)

(See Sec. 20.6 for the details regarding how to define uncertain variable cells, results cells, and statistic cells.)

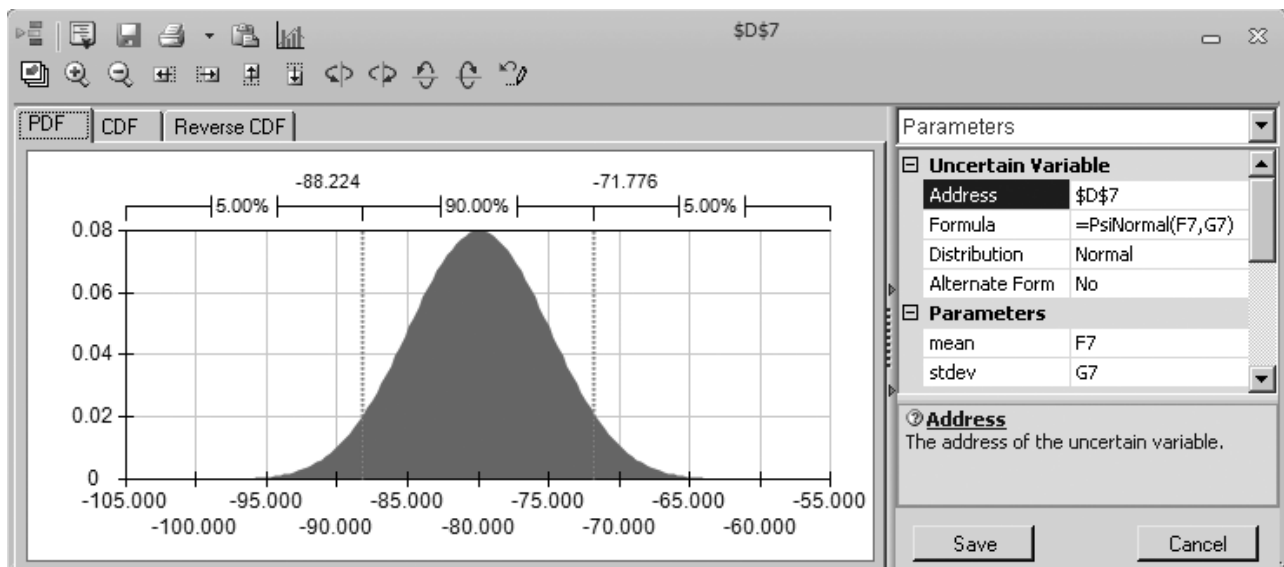
The Simulation Results

Using the Simulation Options dialog box to specify 1,000 trials, Fig. 28.15 shows the results of applying simulation to the spreadsheet model in Fig. 28.13. The frequency chart in Fig. 28.15 provides the risk profile for the proposal since it shows the relative likelihood of the various values of NPV, including those where NPV is negative. The mean is \$18.120 million, which is very attractive. However, the 1,000 trials generated an extremely wide range of NPV values, all the way from about $-\$28$ million to over $\$62$ million. Thus, there is a significant chance of incurring a huge loss. By entering 0 into the box in the Upper Cutoff box of the statistics table, the Likelihood box indicates that 81 percent of the trials resulted in a profit (a positive value of NPV). This also gives the bad news that there is roughly a 19 percent chance of incurring a loss of some size. The lightly shaded portion of the chart to the left of 0 shows that most of the trials with losses involved losses up to about $\$10$ million, but that quite a few trials had losses that ranged from $\$10$ million to nearly $\$30$ million.

Armed with all this information, a managerial decision now can be made about whether the likelihood of a sizable profit justifies the significant risk of incurring a loss and perhaps even a very substantial loss. Thus, the role of simulation is to provide the information needed for making a sound decision, but it is management that uses its best judgment to make the decision.

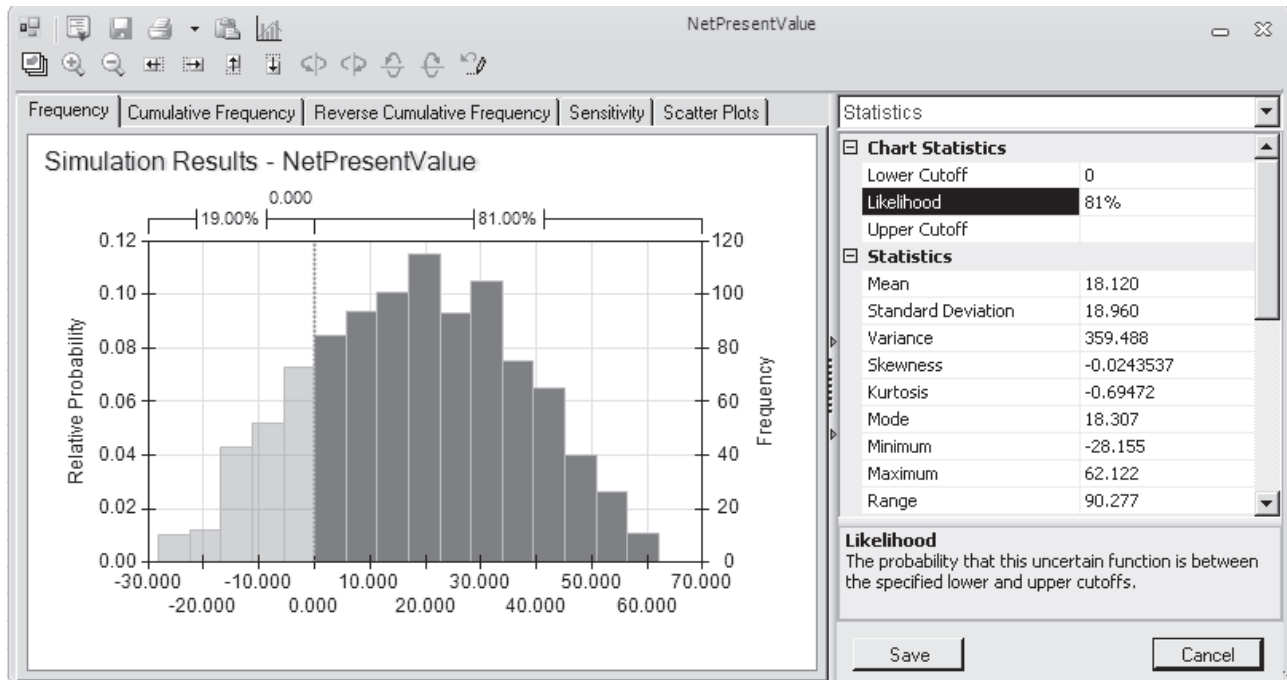
■ FIGURE 28.14

A normal distribution with parameters F7 (5280) and G7 (55) is being entered into the first uncertain variable cell D7 in the spreadsheet model in Fig. 28.13.



■ FIGURE 28.15

The frequency chart and statistics table that summarize the results of running the simulation model in Figure 28.13 for the Think-Big Development Co. financial risk analysis problem. The Likelihood box in the statistics table reveals that 81 percent of the trials resulted in a positive net present value.



■ 28.5 REVENUE MANAGEMENT IN THE TRAVEL INDUSTRY

As described in Sec. 18.8, one of the most prominent areas for the application of operations research in recent years has been in improving *revenue management* in the travel industry. Revenue management refers to the various ways of increasing the flow of revenues through such devices as setting up different fare classes for different categories of customers. The objective is to maximize total income by setting fares that are at the upper edge of what the different market segments are willing to pay and then allocating seats appropriately to the various fare classes.

As the example in this section will illustrate, one key area of revenue management is *overbooking*, that is, accepting a slightly larger number of reservations than the number of seats available. There usually are a small number of no-shows, so overbooking will increase revenue by essentially filling the available seating. However, there also are costs incurred if the number of arriving customers exceeds the number of available seats. Therefore, the amount of overbooking needs to be set carefully so as to achieve an appropriate trade-off between filling seats and avoiding the need to turn away customers who have a reservation.

American Airlines was the pioneer in making extensive use of operations research for improving its revenue management. The guiding motto was “selling the right seats to the right customers at the right time.” This work won the 1991 Franz Edelman Award as that year’s best application of operations research and management science anywhere throughout the world. This application was credited with increasing annual revenues for American Airlines by over \$500 million. Nearly half of these increased revenues came from the use of a new overbooking model.

Following this breakthrough at American Airlines, other airlines quickly stepped up their use of operations research in similar ways. These applications to revenue management then spread to other segments of the travel industry (train travel, cruise lines, rental cars, hotels, etc.) around the world. Our example below involves overbooking by an airline company.

The Transcontinental Airlines Overbooking Problem

Transcontinental Airlines has a daily flight (excluding weekends) from San Francisco to Chicago that is mainly used by business travelers. There are 150 seats available in the single cabin. The average fare per seat is \$300. This is a nonrefundable fare, so no-shows forfeit the entire fare. The fixed cost for operating the flight is \$30,000, so more than 100 reservations are needed to make a profit on any particular day.

For most of these flights, the number of requests for reservations considerably exceeds the number of seats available. The company's OR group has been compiling data on the number of reservation requests per flight for the past several months. The average number has been 195, but with considerable variation from flight to flight on both sides of this average. Plotting a frequency chart for these data suggests that they roughly follow a bell-shaped curve. Therefore, the group estimates that the number of reservation requests per flight has a *normal distribution* with a mean of 195. A calculation from the data estimates that the standard deviation is 30.

The company's policy is to accept 10 percent more reservations than the number of seats available on nearly all its flights, since roughly 10 percent of all its customers making reservations end up being no-shows. However, if its experience with a particular flight is much different from this, then an exception is made and the OR group is called in to analyze what the overbooking policy should be for that particular flight. This is what has just happened regarding the daily flight from San Francisco to Chicago. Even when the full quota of 165 reservations has been reached (which happens for most of the flights), there usually are a significant number of empty seats. While gathering its data, the OR group has discovered the reason why. On the average, only 80 percent of the customers who make reservations for this flight actually show up to take the flight. The other 20 percent forfeit the fare (or, in most cases, allow their company to do so) because their plans have changed.

Now that the data have been gathered, the OR group decides to begin its analysis by investigating the option of increasing the number of reservations to accept for this flight to 190, since 80 percent of 190 = 152, which is very close to the number of seats of available (150). If the number of reservation requests for a particular day actually reaches this level, then this number should be large enough to avoid many, if any, empty seats. Furthermore, this number should be small enough that there will not be many occasions when a significant number of customers need to be bumped from the flight because the number of arrivals exceeds the number of seats available. Thus, 190 appears to be a good first guess for an appropriate trade-off between avoiding many empty seats and avoiding bumping many customers.

When a customer is bumped from this flight, Transcontinental Airlines arranges to put the customer on the next available flight to Chicago on another airline. The company's average cost for doing this is \$150. In addition, the company gives the customer a voucher worth \$200 for use on a future flight. The company also feels that an additional \$100 should be assessed for the intangible cost of a loss of goodwill on the part of the bumped customer. Therefore, the total cost of bumping a customer is estimated to be \$450.

The OR group now wants to investigate the option of accepting 190 reservations by using simulation to generate frequency charts for the following three measures of performance for each day's flight:

1. The profit.
2. The number of filled seats.
3. The number of customers denied boarding.

A Spreadsheet Model for Applying Simulation

Figure 28.16 shows a spreadsheet model for this problem. Because there are three measures of interest here, the spreadsheet model needs three results cells. These results cells are Profit (F23), NumberOfFilledSeats (C20), and NumberDeniedBoarding (C21). In addition, three statistic cells are defined in cells C23:C25 to measure the mean value of each of the results cells for the simulation run. The decision variable ReservationsToAccept (C13) has been set at 190 for investigating this current option. Some basic data have been entered near the top of the spreadsheet in cells C4:C7.

Each trial of the simulation will correspond to one day's flight. There are two random inputs associated with each flight, namely, the number of customers requesting reservations (abbreviated as Ticket Demand in cell B10) and the number of customers who actually arrive to take the flight (abbreviated as Number That Show in cell B17). Thus, the two uncertain variable cells in this model are SimulatedTicketDemand (C10) and NumberThatShow (C17).

Since the OR group has estimated that the number of customers requesting reservations has a normal distribution with a mean of 195 and a standard deviation of 30, this information has been entered into cells D10:F10. The Normal Distribution dialog box (shown earlier in Fig. 28.14) then has been used to enter this distribution with these parameters into SimulatedTicketDemand (C10). Because the normal distribution is a continuous distribution, whereas the number of reservations must have an integer value, Demand (C11) uses Excel's ROUND function to round the number in SimulatedTicketDemand (C10) to the nearest integer.

The random input for the second uncertain variable cell NumberThatShow (C17) depends on two key quantities. One is TicketsPurchased (E17), which is the minimum of Demand (C11) and ReservationsToAccept (C13). The other key quantity is the probability that an individual making a reservation actually will show up to take the flight. This probability has been set at 80 percent in ProbabilityToShowUp (F17) since this is the *average* percentage of those who have shown up for the flight in recent months.

However, the *actual* percentage of those who show up on any particular day may vary somewhat on either side of this average percentage. Therefore, even though NumberThatShow (C17) would be expected to be fairly close to the product of cells E17 and F17, there will be some variation according to some probability distribution. What is the appropriate distribution for this uncertain variable cell? Section 28.6 will describe the characteristics of various distributions. The one that has the characteristics to fit this uncertain variable cell turns out to be the *binomial distribution*.

As indicated in Sec. 28.6, the binomial distribution gives the distribution of the number of times a particular event occurs out of a certain number of opportunities. In this case, the *event* of interest is a passenger showing up to take the flight. The *opportunity* for this event to occur arises when a customer makes a reservation for the flight. These opportunities are conventionally referred to as *trials* (not to be confused with a trial of a simulation). The binomial distribution assumes that the trials are statistically independent and that, on each trial, there is a fixed probability (80 percent in this case) that the event will occur. The parameters of the distribution are this fixed probability and the number of trials.

Figure 28.17 displays the Binomial Distribution dialog box that enters this distribution into NumberThatShow (C17) by referencing the parameters TicketsPurchased (E17) and ProbabilityToShowUp (F17). The actual value in Trials for the binomial distribution will vary from simulation trial to simulation trial because it depends on the number of tickets purchased which in turn depends on the ticket demand which is random. ASPE therefore

■ **FIGURE 28.16**

A spreadsheet model for applying simulation to the Transcontinental Airlines overbooking problem. The uncertain variable cells are SimulatedTicketDemand (C10) and NumberThatShow (C17). The results cells are Profit (F23), NumberOfFilledSeats (C20), and NumberDeniedBoarding (C21). The statistic cells are MeanFilledSeats (C23), MeanDeniedBoarding (C24), and MeanProfit (C25). The decision variable is ReservationsToAccept (C13).

	A	B	C	D	E	F
1		Transcontinental Airlines overbooking				
2						
3			Data			
4		Available Seats	150			
5		Fixed Cost	\$30,000			
6		Avg. Fare / Seat	\$300			
7		Cost of Bumping	\$450			
8						
9					Mean	Standard Dev.
10		Ticket Demand	179.74	<i>Normal</i>	195	30
11		Demand (rounded)	180			
12						
13		Reservations to Accept	190			
14						
15					Tickets	Probability
16					Purchased	to Show Up
17		Number That Show	153	<i>Binomial</i>	180	80%
18						
19						
20		Number of Filled Seats	150		Ticket Revenue	\$45,000
21		Number Denied Boarding	3		Bumping Cost	\$1,350
22					Fixed Cost	\$30,000
23		Mean Filled Seats	142.27		Profit	\$13,650
24		Mean Denied Boarding	2.02			
25		Mean Profit	\$11,775			

	B	C
11	Demand (rounded)	=ROUND(SimulatedTicketDemand,0)

	E
15	Tickets
16	Purchased
17	=MIN(Demand,ReservationsToAccept)

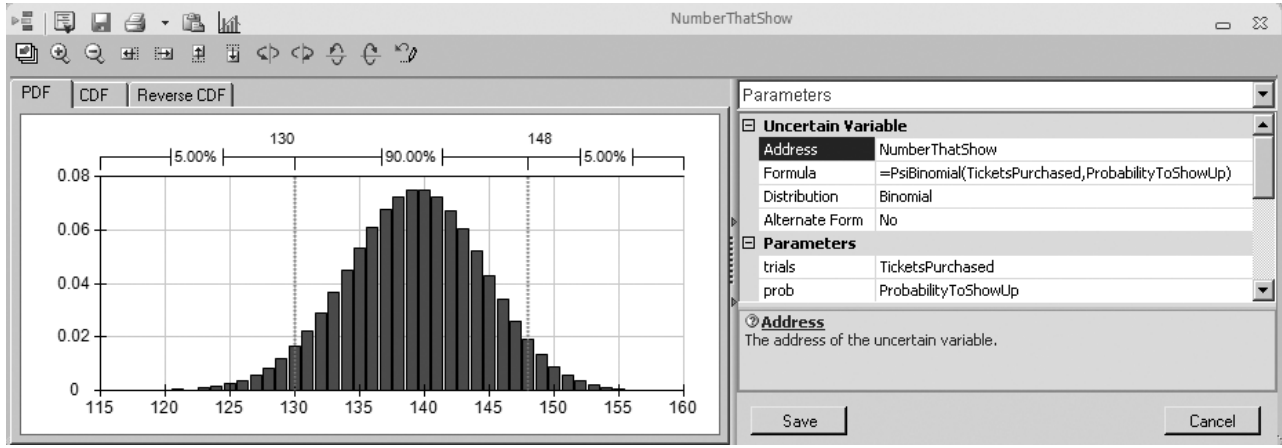
Range Name	Cell
AvailableSeats	C4
AverageFare	C6
BumpingCost	F21
CostOfBumping	C7
Demand	C11
FixedCost	C5
MeanDeniedBoarding	C24
MeanFilledSeats	C23
MeanProfit	C25
NumberDeniedBoarding	C21
NumberOfFilledSeats	C20
NumberThatShow	C17
ProbabilityToShowUp	F17
Profit	F23
ReservationsToAccept	C13
SimulatedTicketDemand	C10
TicketRevenue	F20
TicketsPurchased	E17

	B	C
20	Number of Filled Seats	=MIN(AvailableSeats,NumberThatShow) + PsiOutput()
21	Number Denied Boarding	=MAX(0,NumberThatShow - AvailableSeats) + PsiOutput()
22		
23	Mean Filled Seats	=PsiMean(C20)
24	Mean Denied Boarding	=PsiMean(C21)
25	Mean Profit	=PsiMean(F23)

	E	F
20	Ticket Revenue	=AverageFare*NumberOfFilledSeats
21	Bumping Cost	=CostOfBumping*NumberDeniedBoarding
22	Fixed Cost	=FixedCost
23	Profit	=TicketRevenue - BumpingCost - FixedCost + PsiOutput()

■ **FIGURE 28.17**

A binomial distribution with parameters TicketsPurchased (E17) and ProbabilityToShowUp (F17) is being entered into the uncertain variable cell NumberThatShow (C17).



must determine the value for TicketsPurchased (E17) *before* it can randomly generate NumberThatShow (C17). Fortunately, ASPE automatically takes care of the order in which to generate the various uncertain variable cells so that this is not a problem.

The equations entered into all the output cells, results cells, and statistic cells are given at the bottom of Fig. 28.16.

Here is a summary of the key cells in this model:

Uncertain variable cells: SimulatedTicketDemand (C10) and NumberThatShow (C17)

Decision variable: ReservationsToAccept (C13)

Results cells: Profit (F23), NumberOfFilledSeats (C20), and
NumberDeniedBoarding (C21)

Statistic cells: MeanFilledSeats (C23), MeanDeniedBoarding (C24), MeanProfit (C25)

(See Sec. 20.6 for the details regarding how to define uncertain variable cells, results cells, and statistic cells.)

The Simulation Results

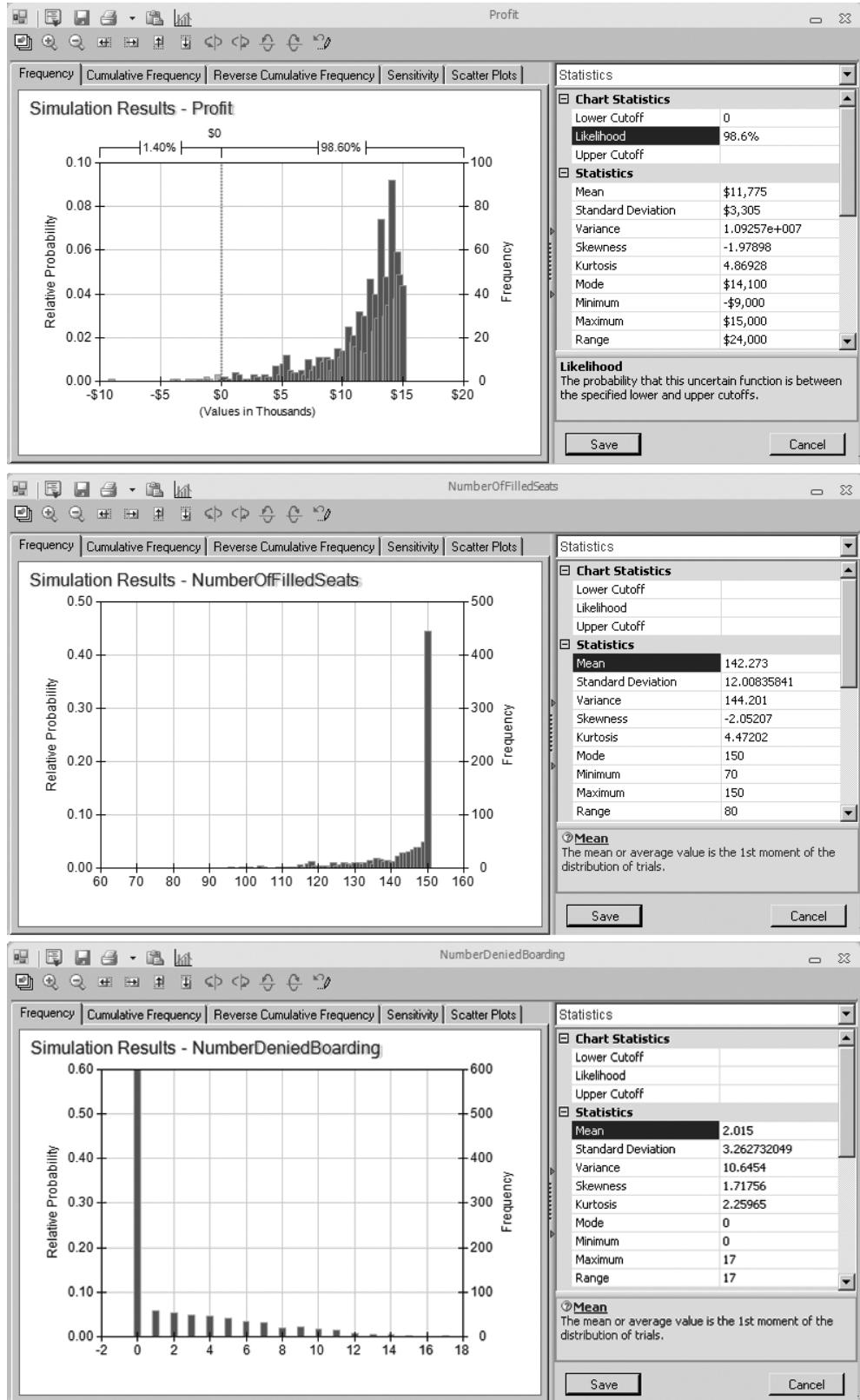
Figure 28.18 shows the frequency chart obtained for each of the three results cells after applying simulation for 1,000 trials to the spreadsheet model in Fig. 28.16, with ReservationsToAccept (C13) set at 190.

The profit results estimate that the mean profit per flight would be \$11,775. However, this mean is a little less than the profits that had the highest frequencies. The reason is that a small number of trials had profits far below the mean, including even a few that incurred losses, which dragged the mean down somewhat. By entering 0 into the Lower Cutoff box, the Likelihood box reports that 98.6 percent of the trials resulted in a profit for that day's flight.

The frequency chart for NumberOfFilledSeats (C20) indicates that almost half of the 1,000 trials resulted in all 150 seats being filled. Furthermore, most of the remaining trials had at least 130 seats filled. The fact that the mean of 142.273 is so close to 150 shows that a policy of accepting 190 reservations would do an excellent job of filling seats.

The price that would be paid for filling seats so well is that a few customers would need to be bumped from some of the flights. The frequency chart for NumberDeniedBoarding (C21) indicates that this occurred about 40 percent of the time. On nearly all of

FIGURE 28.18
 The frequency charts and statistics tables that summarize the results for the respective results cells—Profit (F23), NumberOfFilledSeats (C20), and NumberDeniedBoarding (C21)—from running the simulation model in Fig. 28.16 for the Transcontinental Airlines overbooking problem. The Likelihood box in the first statistics table reveals that 98.6 percent of the trials resulted in a positive profit.



these trials, the number ranged between 1 and 10. Considering that no customers were denied boarding for 60 percent of the trials, the mean number is only 2.015.

Although these results suggest that a policy of accepting 190 reservations would be an attractive option for the most part, they do not demonstrate that this is necessarily the best option. Additional simulation runs are needed with other numbers entered in Reservations To Accept (C13) to pin down the optimal value of this decision variable. This can be done fairly easily with trial-and-error. We also will demonstrate how to do this efficiently with the help of a parameter analysis table in Sec. 28.7.

28.6 CHOOSING THE RIGHT DISTRIBUTION

As mentioned in Sec. 20.6, ASPE's Distributions menu provides a wealth of choices. Any of 46 probability distributions can be selected as the one to be entered into any uncertain variable cell. In the preceding sections, we have illustrated the use of five of these distributions (the integer uniform, uniform, triangular, normal, and binomial distributions). However, not much was said about *why* any particular distribution was chosen.

In this section, we focus on the issue of how to choose the right distribution. We begin by surveying the characteristics of many of the 46 distributions and how these characteristics help to identify the best choice. We next describe a special feature of ASPE for creating one of the 7 available custom distributions when none of the other 39 choices in the Distributions menu will do. We then return to the example analyzed in Sec. 20.6 to illustrate another special feature of ASPE. When historical data are available, this feature will identify which of the available distributions provides the best fit to these data while also estimating the parameters of this distribution. If you do not like this choice, it will even identify which of the distributions provides the second best fit, the third best fit, and so on.

Characteristics of the Available Distributions

The probability distribution of any random variable describes the relative likelihood of the possible values of the random variable. A *continuous* distribution is used if *any* values are possible, including both integer and fractional numbers, over the entire range of possible values. A *discrete* distribution is used if only certain specific values (e.g., only the integer numbers over some range) are possible. However, if the only possible values are integer numbers over a relatively broad range, a continuous distribution may be used as an approximation by rounding any fractional value to the nearest integer. (This approximation was used in cells C10:C11 of the spreadsheet model in Fig. 28.16.) ASPE's Distributions menu includes both continuous and discrete distributions. We will begin by looking at the continuous distributions.

The right-hand side of Fig. 28.19 shows the dialog box for three popular continuous distributions from the Common submenu of the Distributions menu. The dark figure in each dialog box displays a typical *probability density function* for that distribution. The height of the probability density function at the various points shows the relative likelihood of the corresponding values along the horizontal axis. Each of these distributions has a most likely value where the probability density function reaches a peak. Furthermore, all the other relatively high points are near the peak. This indicates that there is a tendency for one of the central values located near the most likely value to be the one that occurs. Therefore, these distributions are referred as *central-tendency distributions*. The characteristics of each of these distributions are listed on the left-hand side of Fig. 28.19.

The Normal Distribution

The normal distribution is widely used by both OR professionals and others because it describes so many natural phenomena. (Because of its importance, Appendix 5 provides a

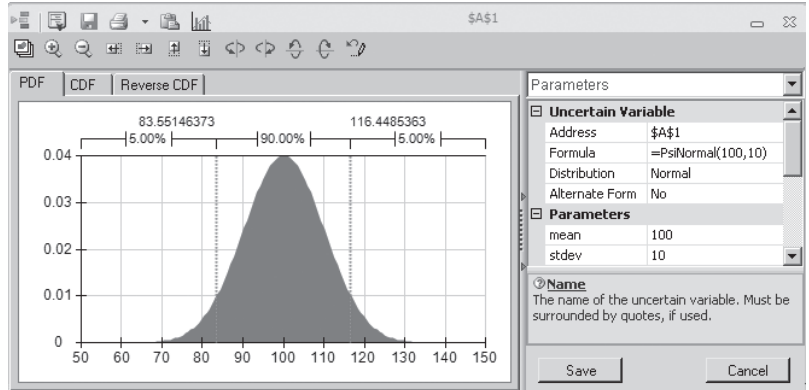
■ **FIGURE 28.19**

The characteristics and dialog boxes for three popular central-tendency distributions in ASPE's Common submenu of the Distributions menu: (1) the normal distribution, (2) the triangular distribution, and (3) the lognormal distribution.

Popular Central-Tendency Distributions

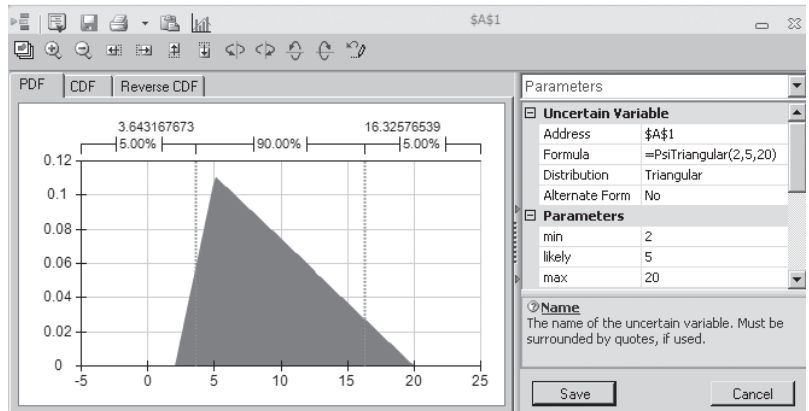
Normal Distribution:

- Some value most likely (the mean)
- Values close to mean more likely
- Symmetric (as likely above as below mean)
- Extreme values possible, but rare



Triangular Distribution:

- Some value most likely
- Values close to most likely value more common
- Can be asymmetric
- Fixed upper and lower bound



Lognormal Distribution:

- Some value most likely
- Positively skewed (below mean more likely)
- Values cannot fall below zero
- Extreme values (high end only) possible, but rare

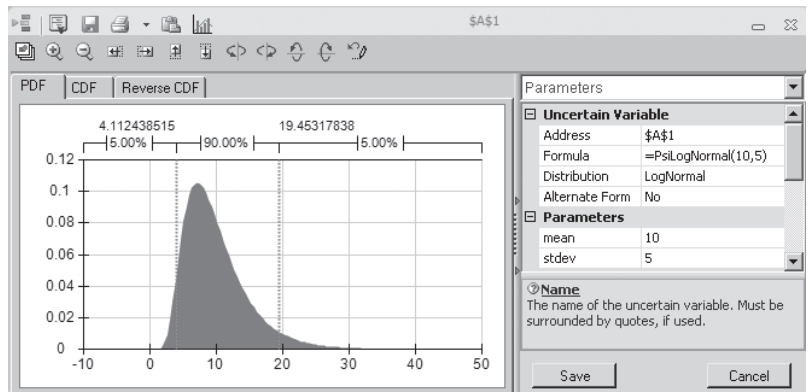


table for this distribution.) One reason that it arises so frequently is that the sum of many random variables tends to have a normal distribution (approximately) even when the individual random variables do not. Using this distribution requires estimating the mean and the standard deviation. The mean coincides with the most likely value because this is a

symmetric distribution. Thus, the mean is a very intuitive quantity that can be readily estimated, but the standard deviation is not. About two-thirds of the distribution lies within one standard deviation of the mean. Therefore, if historical data are not available for calculating an estimate of the standard deviation, a rough estimate can be elicited from a knowledgeable individual by asking for an amount such that the random value will be within that amount of the mean about two-thirds of the time.

One danger with using the normal distribution for some applications is that it can give negative values even when such values actually are impossible. Fortunately, it can give negative values with significant frequency only if the mean is less than three standard deviations. For example, consider the situation where a normal distribution was entered into an uncertain variable cell in Fig. 28.16 to represent the number of customers requesting a reservation. A negative number would make no sense in this case, but this was no problem since the mean (195) was much larger than three standard deviations (3 3 30 5 90) so a negative value essentially could never occur. (When normal distributions were entered into uncertain variable cells in Fig. 28.13 to represent cash flows, the means were small or even negative, but this also was no problem since cash flows can be either negative or positive.)

The Triangular Distribution

A comparison of the shapes of the triangular and normal distributions in Fig. 28.19 reveals some key differences. One is that the triangular distribution has a fixed minimum value and a fixed maximum value, whereas the normal distribution allows rare extreme values far into the tails. Another is that the triangular distribution can be asymmetric (as shown in the figure), because the most likely value does not need to be midway between the bounds, whereas the normal distribution always is symmetric. This asymmetry provides additional flexibility to the triangular distribution. Another key difference is that all its parameters—min (the minimum value), likely (the most likely value), and max (the maximum value)—are intuitive ones, so they are relatively easy to estimate.

These advantages have made the triangular distribution a popular choice for simulations. They are the reason why this distribution was used in previous examples to represent competitors' bids for a construction contract (in Fig. 28.2), activity times (in Fig. 28.6), and cash flows (in Fig. 28.11).

However, the triangular distribution also has certain disadvantages. One is that, in many situations, rare extreme values far into the tails are possible, so it is quite artificial to have fixed minimum and maximum values. This also makes it difficult to develop meaningful estimates of the bounds. Still another disadvantage is that a curve with a gradually changing slope, such as the bell-shaped curve for the normal distribution, usually describes the true distribution more accurately than the straight line segments in the triangular distribution.

The Lognormal Distribution

The lognormal distribution shown at the bottom of Fig. 28.19 combines some of the advantages of the normal and triangular distributions. It has a curve with a gradually changing slope. It also allows rare extreme values on the high side. At the same time, it does not allow negative values, so it automatically fits situations where this is needed. This is particularly advantageous when the mean is less than three standard deviations and the normal distribution should not be used.

This distribution always is "positively skewed," meaning that the long tail always is to the right. This forces the most likely value to be toward the left side (so the mean is on its right), so this distribution is less flexible than the triangular distribution. Another disadvantage is that it has the same parameters as the normal distribution (the mean and the standard deviation), so the less intuitive one (the standard deviation) is difficult to estimate unless historical data are available.

When a positively skewed distribution that does not allow negative values is needed, the lognormal distribution provides an attractive option. That is why this distribution frequently is used to represent stock prices or real estate prices.

The Uniform and Integer Uniform Distributions

Although the preceding three distributions are all central-tendency distributions, the uniform distributions shown in Fig. 28.20 definitely are not. They have a fixed minimum value and a fixed maximum value. Otherwise, they say that no value between these bounds is any more likely than any other possible value. Therefore, these distributions have more variability than the central-tendency distributions with the same range of possible values (excluding rare extreme values).

The choice between these two distributions depends on which values between the minimum and maximum values are possible. If *any* values in this range are possible, including even *fractional* values, then the uniform distribution would be preferred over the integer uniform distribution. If only integer values are possible, then the integer uniform distribution would be the preferable one.

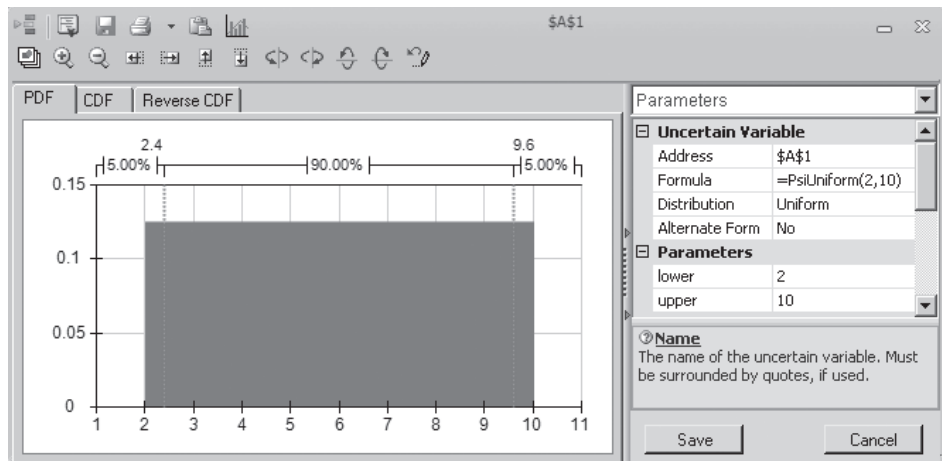
■ **FIGURE 28.20**

The characteristics and dialog boxes for the uniform distributions in ASPE's Distributions menu.

Uniform and Integer Uniform Distribution

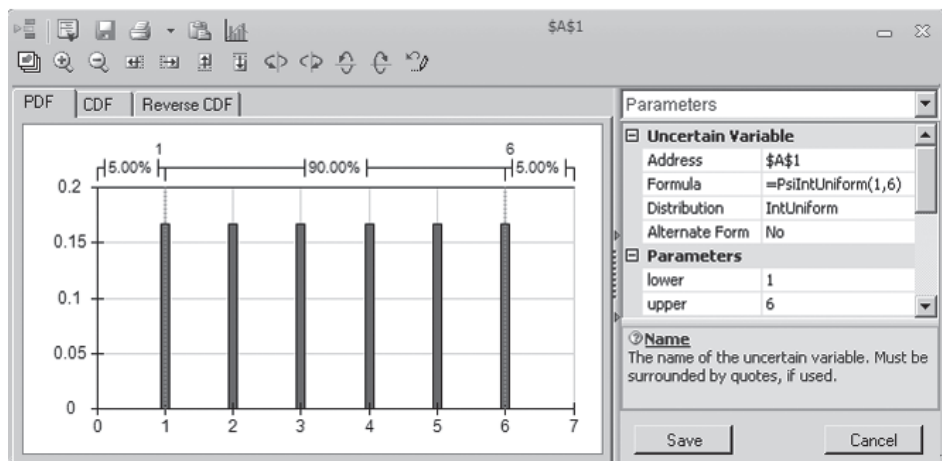
Uniform Distribution:

- Fixed lower and upper limit
- All values equally likely



Integer Uniform Distribution:

- Fixed lower and upper limit
- All integer values equally likely



Either of these distributions is a particularly convenient one because it has only two parameters (lower and upper limit) and both are very intuitive. These distributions receive considerable use for this reason. In our examples of performing simulations on a spreadsheet, the integer uniform distribution was used to represent the demand for a newspaper (in Fig. 20.7 in Sec. 20.6), whereas the uniform distribution was used to generate the bid for a construction project by one competitor (in Fig. 28.2) and the future sale price for real estate property (in Fig. 28.13).

The disadvantage of this distribution is that it usually is only a rough approximation of the true distribution. It is uncommon for either the minimum value or the maximum value to be just as likely as any other value between these bounds while any value barely outside these bounds is impossible.

The Exponential Distribution

If you have studied Chap. 17 on queueing theory, you hopefully will recall that the most commonly used queueing models assume that the time between consecutive arrivals of customers to receive a particular service has an exponential distribution. The reason for this assumption is that, in most such situations, the arrivals of customers are random events and the exponential distribution is the probability distribution of the time between random events. Section 17.4 describes this property of the exponential distribution in some detail.

As first depicted in Fig. 17.3, this distribution has the unusual shape shown in Fig. 28.21. In particular, the peak is at 0 but there is a long tail to the right. This indicates that the most likely times are short ones well below the mean but that very long times also are possible. This is the nature of the time between random events.

Since the only parameter is the mean time until the next random event occurs, this distribution is a relatively easy one to use.

The Poisson Distribution

Although the exponential distribution (like most of the preceding ones) is a continuous distribution, the Poisson distribution is a discrete distribution, as shown in the bottom half of Fig. 28.21. The only possible values are nonnegative integers: 0, 1, 2, However, it is natural to pair this distribution with the exponential distribution for the following reason. If the time between consecutive events has an exponential distribution (i.e., the events are occurring at random), then the number of events that occur within a certain period of time has a Poisson distribution. (This is property 4 of the exponential distribution described in Sec. 17.4.) The Poisson distribution has some other applications as well.

When considering the number of events that occur within a certain period of time, the mean to be entered into the one parameter field in the dialog box should be the average number of events that occur within that period of time.

The Bernoulli and Binomial Distribution

The Bernoulli distribution is a very simple discrete distribution with only two possible values (1 or 0) as shown in the top half of Fig. 28.22. It is used to simulate whether a particular event occurs or not. The only parameter of the distribution is the probability that the event occurs. The Bernoulli distribution gives a value of 1 (representing yes) with this probability; otherwise, it gives a value of 0 (representing no).

As shown in the bottom half of Fig. 28.22, the binomial distribution is an extension of the Bernoulli distribution for when an event might occur a number of times. The binomial distribution gives the probability distribution of the number of *times* a particular event occurs, given the number of independent opportunities (called trials) for the event to occur, where the probability of the event occurring remains the same from trial to trial. For example, if the event of interest is getting heads on the flip of a coin, the binomial distribution (with Prob. = 0.5) gives the distribution of the number of heads in a given number of flips

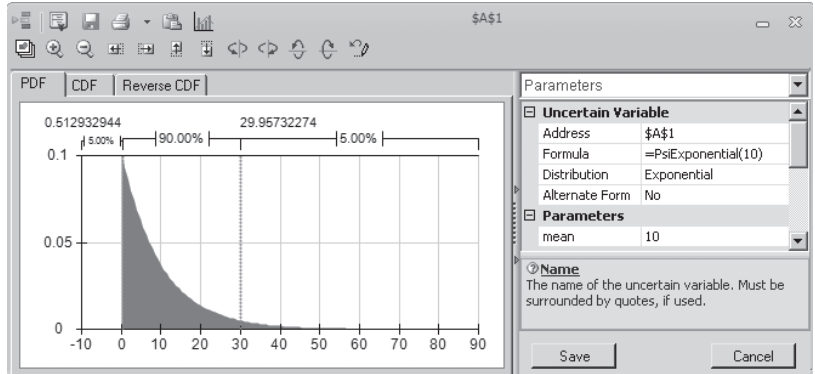
■ FIGURE 28.21

The characteristics and dialog boxes for two distributions that involve random events. These distributions in ASPE's Distributions menu are (1) the exponential distribution and (2) the Poisson distribution.

Distributions for Random Events

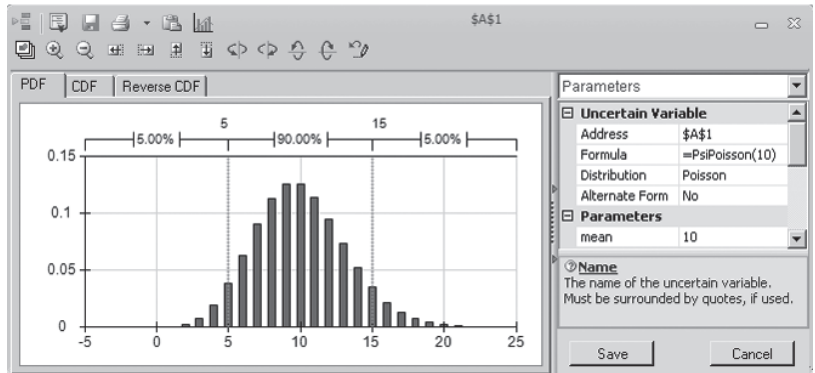
Exponential Distribution:

- Widely used to describe time between random events (e.g., time between arrivals)
- Events are independent
- Mean = average time until the next event occurs



Poisson Distribution:

- Describes the number of times an event occurs during a given period of time or space
- Occurrences are independent
- Any number of events is possible
- Mean = average number of events during period of time (e.g., arrivals per hour), assumed constant over time



of the coin. Each flip constitutes a trial where there is an opportunity for the event (heads) to occur with a fixed probability (0.5). The binomial distribution is equivalent to the Bernoulli distribution when the number of trials is equal to 1.

You have seen another example in the preceding section when the binomial distribution was entered into the uncertain variable cell NumberThatShow (C17) in Fig. 28.16. In this airline overbooking example, the events are customers showing up for the flight and the trials are customers making reservations, where there is a fixed probability that a customer making a reservation actually will arrive to take the flight.

The only parameters for this distribution are the number of trials and the probability of the event occurring on a trial.

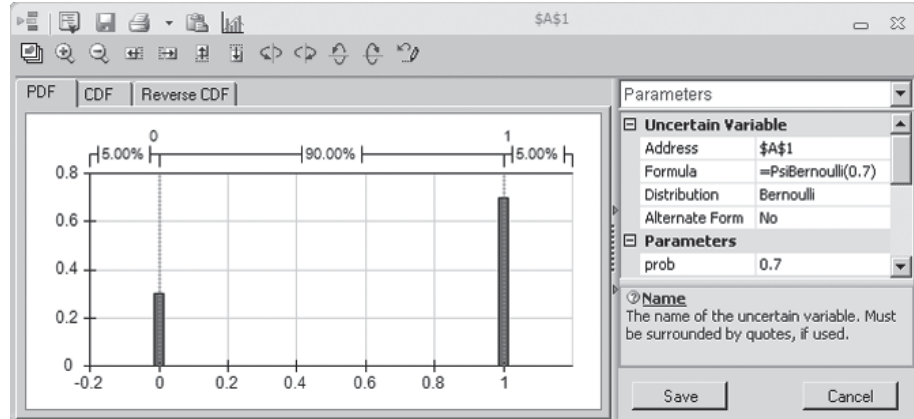
The Geometric and Negative Binomial Distributions

These two distributions displayed in Fig. 28.23 are related to the binomial distribution because they again involve trials where there is a fixed probability on each trial that the event will occur. The *geometric distribution* gives the distribution of the number of trials until the event occurs for the first time. After entering a positive integer into the suc field in its dialog box, the *negative binomial distribution* gives the distribution of the number of trials until the event occurs the number of times specified in the suc field (suc is the number of successful events that must occur). Thus, suc is a parameter for this distribution and the fixed probability of the event occurring on a trial is a parameter for both distributions.

Distributions for Number of Times an Event Occurs

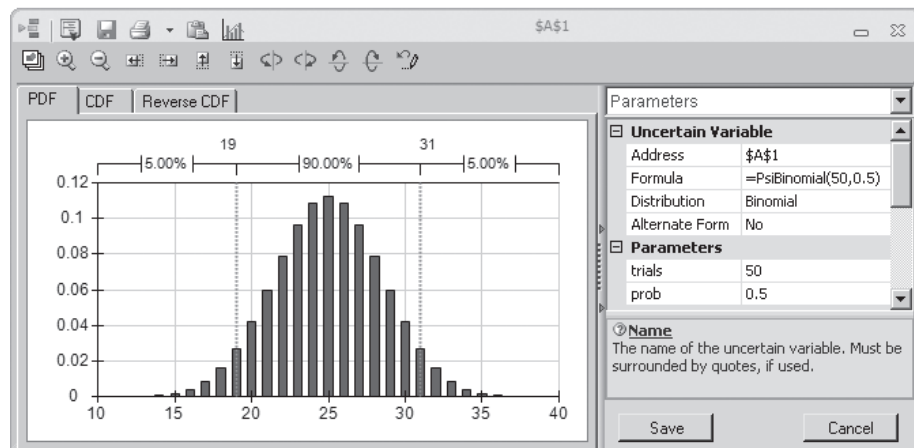
Bernoulli Distribution:

- Describes whether an event occurs or not
- Two possible outcomes: 1 (Yes) or 0 (No)



Binomial Distribution:

- Describes number of times an event occurs in a fixed number of trials (e.g., number of heads in 10 flips of a coin)
- For each trial, only two outcomes possible
- Trials independent
- Probability remains same for each trial



■ **FIGURE 28.22**

The characteristics and dialog boxes for the Bernoulli and binomial distributions in ASPE's Distributions menu.

To illustrate these distributions, suppose you are again interested in the event of getting heads on a flip of a coin (a trial). The geometric distribution (with Prob. = 0.5) gives the distribution of the number of flips until the first head occurs. If you want five heads, the negative binomial distribution (with Prob. = 0.5 and suc = 5) gives the distribution of the number of flips until heads have occurred five times.

Similarly, consider a production process with a 50 percent yield, so each unit produced has an 0.5 probability of being acceptable. The geometric distribution (with Prob. = 0.5) gives the distribution of the number of units that need to be produced to obtain one acceptable unit. If a customer has ordered five units, the negative binomial distribution (with Prob. = 0.5 and suc = 5) gives the distribution of the production run size that is needed to fulfill this order.

Other Distributions

The Distributions menu includes many other distributions as well, such as beta, gamma, Weibull, Pert, Pareto, Erlang, and many more. These distributions are not as widely used in simulations, so they will not be discussed further.

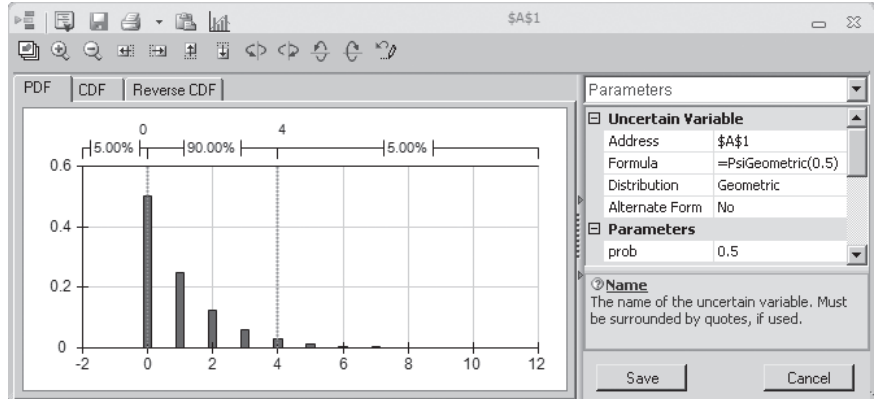
■ FIGURE 28.23

The characteristics and dialog boxes for two distributions that involve the number of trials until events occur. These distributions in ASPE's Distributions menu are (1) the geometric distribution and (2) the negative binomial distribution.

Distributions for Number of Trials Until Event Occurs

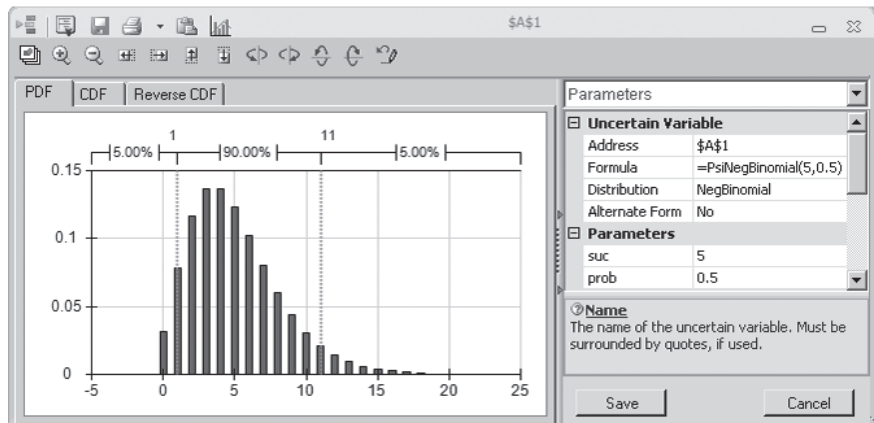
Geometric Distribution:

- Describes number of trials until an event occurs (e.g., number of times to spin roulette wheel until you win)
- Probability same for each trial
- Continue until succeed
- Number of trials unlimited



Negative Binomial Distribution:

- Describes number of trials until an event occurs n times
- Same as geometric when $\text{suc} = n = 1$
- Probability same for each trial
- Continue until n^{th} success
- Number of trials unlimited



There is also a Custom submenu that enables you to design your own distribution when none of the other distributions will do. The next subsection will focus on how this is done.

The Custom Distribution

Of the 46 probability distributions included in the Distributions menu, 39 of them are standard types that might be discussed in a course on probability and statistics. In most cases, one of these standard distributions will be just what is needed for an uncertain variable cell. However, unique circumstances occasionally arise where none of the standard distributions fit the situation. This is where the distributions in the Custom submenu of the Distributions menu enter the picture.

The custom distributions actually are not probability distributions until you make them one. Rather, choosing a member of the Custom submenu triggers a process that enables you to custom-design your own probability distribution to fit almost any unique situation you might encounter.

There are seven choices in the Custom submenu: Cumul (short for cumulative), Discrete, DisUniform (short for discrete uniform), General, Histogram, Sip, and Slurp. The

custom cumulative, *custom general*, and *custom histogram* distributions are all similar in that they are all used to create a continuous distribution with a fixed minimum and maximum value. With the *custom cumulative* distribution, you enter several values in between the minimum and maximum, along with the corresponding cumulative probability at those values. With the *custom general* distribution, you also enter several values in between the minimum and maximum, but instead of cumulative probabilities, you enter relative weights that represent how likely it should be for outcomes near the listed values to occur (relative to outcomes near the other values in the list). Finally, with the *custom histogram* distribution, the range between the minimum and maximum is divided into a number of equal-sized segments and weights are provided for each segment to indicate how likely it should be (relative to the other segments) for a random outcome to fall within that segment.

The *custom discrete* and the *custom discrete uniform* distributions are also similar. With both, you enter a set of discrete values and these values are assumed to be the only possible outcomes. With the *custom discrete* distribution, each value (or outcome) is assigned its own probability, whereas the *custom discrete uniform* distribution assumes that all the discrete values have the same probability.

Finally, the *custom sip* and *custom slurp* distribution are used when you have a set of historical data and you want the uncertain variable to sample directly from the historical data. This might be appropriate if you expect the future to behave similarly to the past.

We will show two examples that use distributions from the Custom submenu. The first utilizes the *custom discrete* distribution whereas the second utilizes the *custom general* distribution.

In the first example, a company is developing a new product but it is unclear which of three production processes will be needed to produce the product. The unit production cost will be \$10, \$12, or \$14, depending on which process is needed. The probabilities for these individual discrete values of the cost are the following:

- 20 percent chance of \$10
- 50 percent chance of \$12
- 30 percent chance of \$14

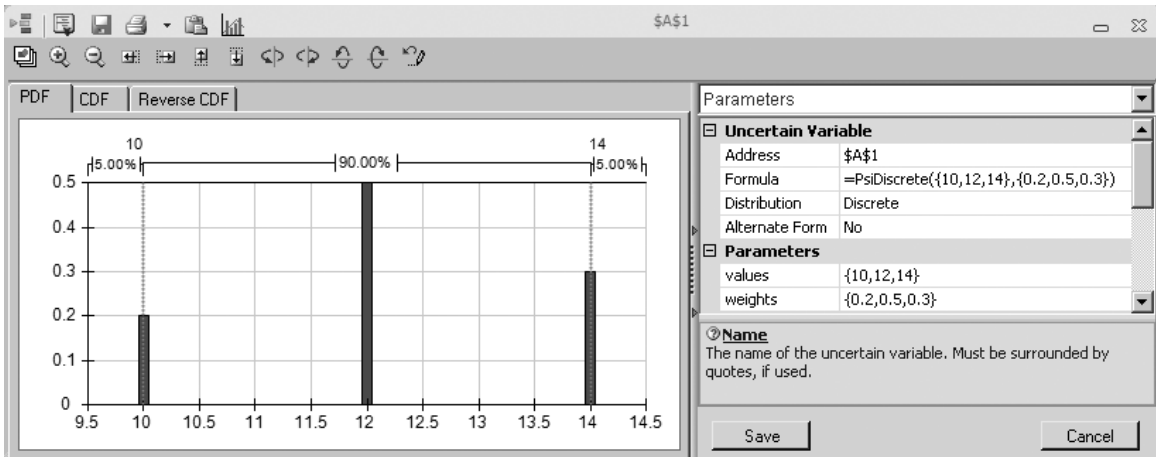
To enter this distribution, first choose Discrete from the Custom submenu under the Distributions menu on the ASPE ribbon. Each discrete value and weight (expressed as a decimal number representing the probability) is then entered in the values and weights boxes as a list within curly brackets, as shown in Fig. 28.24.

The second example also involves a company that is developing a new product. However, the complication in this case is that our company's management has learned that another firm is developing a competitive product. It is unclear which company will be able to bring its product to market first and thereby capture most of the sales. In this light, here are the predicted thousands of sales for our company's new product:

- 0–20 (with 10 the most likely) if the competitive product comes to market first.
- 20–30 (all equally likely) if both products reach the market at the same time.
- 30–50 (with 40 the most likely) if our company's product comes to market first.

The two products are believed to have an equal chance of reaching the market first. Each of the other two cases is considered about three times as likely as both products reaching the market at the same time.

To enter this distribution, first choose General from the Custom submenu of the Distributions menu to bring up the dialog box shown in Fig. 28.25. The first two parameters, min = 0 and max = 50, are used to specify the smallest and largest possible value for sales (in thousands). In the values box, any number of values between the minimum and



■ **FIGURE 28.24**

This dialog box illustrates how ASPE's *custom discrete* distribution can enable you to custom-design your own distribution by entering a set of discrete values and their weights.

maximum can be entered as a list inside of curly brackets. For each value in this list, a corresponding weight needs to be entered into the weights box. The weight is a relative value used to specify the likelihood of each value relative to the other values in the list. Since sales around 10 thousand or 40 thousand are each about three times as likely as sales between 20 thousand and 30 thousand, the weights for the values $\{10, 20, 30, 40\}$ are entered as $\{3, 1, 1, 3\}$. The net result is what is known as a *bimodal distribution*, with two distinct peaks, as shown in the chart on the left side of the dialog box.

Identifying the Continuous Distribution That Best Fits Historical Data

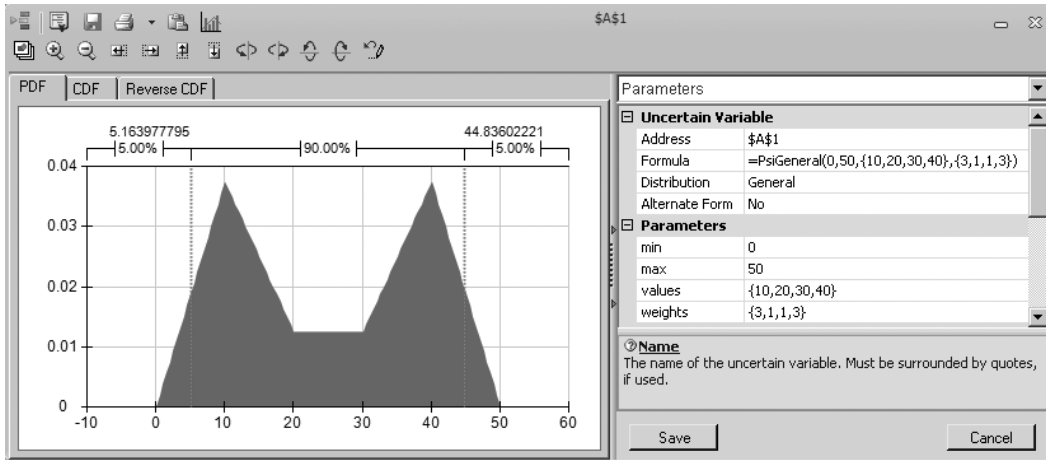
We now have at least mentioned most of the probability distributions in the Distributions menu and have described the characteristics of many of them. This brings us to the question of how to identify which distribution is best for a particular uncertain variable cell. When historical data are available, ASPE provides a powerful feature for doing this by using the Fit button on the ASPE ribbon. We will illustrate this feature next by returning to the example involving Freddie the newsboy that was presented in Sec. 20.6.

Recall that one of the most popular newspapers that Freddie the newsboy sells from his newsstand is the daily *Financial Journal*. Freddie purchases copies from his distributor early each morning. Since excess copies left over at the end of the day represent a loss for Freddie, he is trying to decide what his order quantity should be in the future. This led to the spreadsheet model in Fig. 20.7 that was presented in Sec. 20.6. This model includes the uncertain variable cell Demand (C12). To get started, a discrete uniform distribution between 40 and 70 has been entered into this uncertain variable cell.

To better guide his decision on what the order quantity should be, Freddie has been keeping a record of the *demand* (the number of customers requesting a copy) for this newspaper each day. Figure 28.26 shows a portion of the data he has gathered over the last 60 days in cells F4:F63, along with part of the original spreadsheet model from Fig. 20.7. These data indicate a lot of variation in sales from day to day—ranging from about 40 copies to 70 copies. However, it is difficult to tell from these numbers which distribution in the Distributions menu best fits these data.

■ **FIGURE 28.25**

This dialog box illustrates how ASPE's *custom general* distribution can enable you to custom-design your own continuous distribution. The minimum and maximum value are specified as 0 and 50. Values near 10 and 40 are roughly 3 times as likely to occur as values near 20 or 30.



ASPE provides the following procedure for fitting the best distribution to data:

1. Gather the data needed to identify the best distribution to enter into an uncertain variable cell.
2. Enter the data into the spreadsheet containing your simulation model.
3. Select the cells containing the data.
4. Click the Fit button on the ASPE ribbon, which brings up the Fit Options dialog box.
5. Make sure the Range box in this dialog box is correct for the range of the historical data in your worksheet.
6. Specify which type of distributions are being considered for fitting (continuous or discrete).
7. Indicate whether to allow shifted distributions and whether to run a sample independence test.
8. Also use this dialog box to select which ranking method should be used to evaluate how well a distribution fits the data.
9. Click Fit, which brings up the Fit Results chart that identifies the distribution that best fits the data.
10. If desired, check the box to select distributions to view that are lower on the list on the left side of the dialog box. This identifies the other types of distributions (including their parameter values) that are next in line for fitting the data well.
11. After choosing the distribution (from steps 9 and 10) that you want to use, close the dialog box by using the close box in the upper-right-hand corner and then click Yes to accept the distribution.
12. Click the cell where you want the uncertain variable cell to be. This then enters the chosen distribution into the uncertain variable cell.

Since Fig. 28.26 already includes the needed data in cells F4:F63, applying this procedure to Freddie's problem begins by selecting the data. Then clicking the Fit button brings up the Fit Options dialog box displayed in Fig. 28.27. The range F4:F63 of the data in Fig. 28.26 is already entered into the Range box of this dialog box. When deciding which type of distributions should be considered for fitting, the default option of continuous

■ FIGURE 28.26

Cells F4:F63 contain the historical demand data that have been collected for the example involving Freddie the newsboy that was introduced in Sec. 20.6. Columns B and C come from the simulation model for this example in Fig. 20.7.

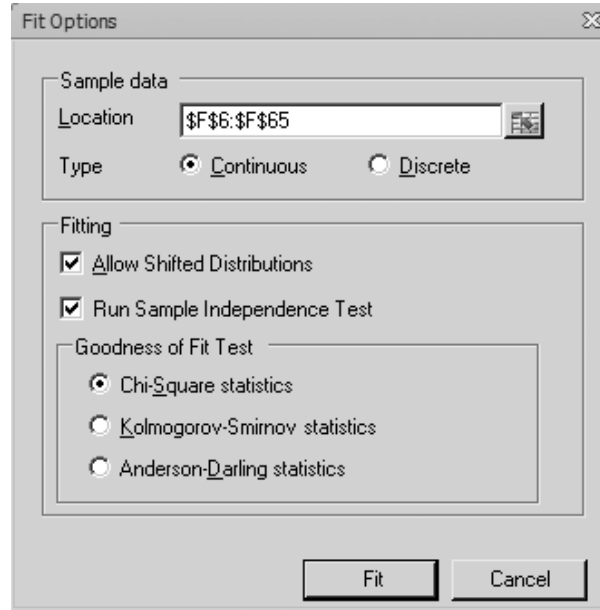
	A	B	C	D	E	F
1		Freddie the Newsboy				
2						
3			Data			Historical
4		Unit Sale Price	\$2.50			Demand
5		Unit Purchase Cost	\$1.50		Day	Data
6		Unit Salvage Value	\$0.50		1	62
7					2	45
8			Decision Variable		3	59
9		Order Quantity	60		4	65
10					5	50
11			Simulation		6	64
12		Demand	44		7	56
13					8	51
14		Sales Revenue	\$110.00		9	55
15		Purchasing Cost	\$90.00		10	61
16		Salvage Value	\$8.00		11	40
17					12	47
18		Profit	\$28.00		13	63
19					14	68
20		Mean Profit	\$46.45		15	67
21					16	67
60					55	41
61					56	42
62					57	64
63					58	45
64					59	59
65					60	70

distributions has been selected here. Sales will always be integer, so discrete might seem the more logical choice. However, when all the integer values over a wide range are possible (all 31 integer values between 40 and 70 in this case), the form of the distribution begins to resemble a continuous distribution. Furthermore, there are many more continuous distributions (31) available in ASPE than discrete distributions (8). Thus, there may be a better chance of finding a continuous distribution that is a good fit. This continuous distribution can then be made to give only integer values by rounding each number in the uncertain variable cell to the nearest integer (as was done in the airline overbooking example of Sec. 28.5 with the ticket demand in cell C11 of Fig. 28.16). The chi-square test also has been selected for the ranking method. Clicking Fit then brings up the Fit Results chart displayed in Fig. 28.28.

The left side of the Fit Results chart in Fig. 28.28 identifies the best-fitting distributions, ranked according to the Chi-Square test. This is a widely used test in statistics where smaller values indicate a better fit. It appears that the uniform distribution would be a good fit. In combination with the fact that demand actually must be integer, this confirms that the choice made in Freddie's original spreadsheet model in Fig. 20.7 to enter the integer

■ **FIGURE 28.27**

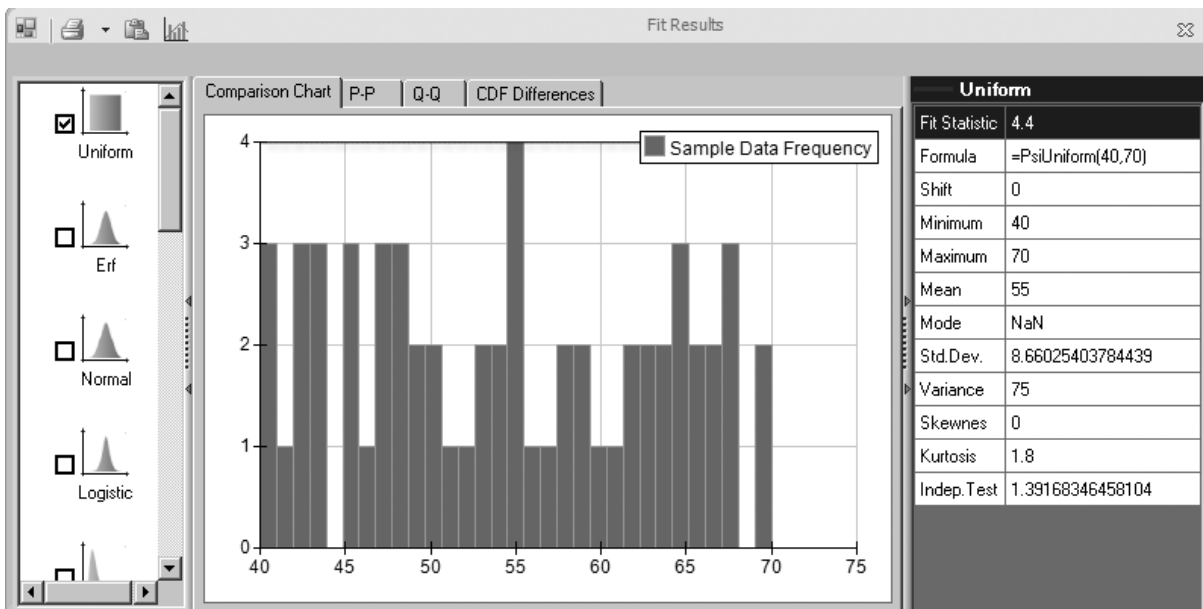
This Fit Options dialog box specifies (1) the range of the data in Fig. 28.26 for Freddie's problem, (2) only continuous distributions will be considered, (3) shifted distributions will be allowed, (4) a sample independence test will be run, and (5) which ranking method will be used (the chi-square test) to evaluate how well each of the distributions fit the data.



uniform distribution into the uncertain variable cell Demand (C12) was reasonable. In fact, if we had chosen Discrete instead of Continuous as the type of distribution to fit in Fig. 28.27, ASPE would have found the integer uniform distribution to be the best fit. Choosing either Continuous or Discrete (or both) would have been reasonable in this case and would have led to the same type of distribution (uniform).

■ **FIGURE 28.28**

This Fit Results dialog box identifies the continuous distributions that provide the best fit, ranked top-to-bottom from best to worst on the left side. For the distribution that provides the best fit (Uniform), the distribution is plotted (the horizontal line at the top of the chart) so that it can be compared with the frequency distribution of the historical demand data. The value of the Fit Statistic (chi-square) is 4.4.



28.7 DECISION MAKING WITH PARAMETER ANALYSIS REPORTS AND TREND CHARTS

Many simulation models include at least one decision variable. For example, the model formulated for both the bidding example in Sec. 28.1 and the overbooking example in Sec. 28.5 included a single decision variable, as listed below:

Bidding example: OurBid (C25) in Fig. 28.2

Overbooking example: ReservationsToAccept (C13) in Fig. 28.16

In both of these cases, you have seen how well simulation with ASPE can *evaluate* a particular value of the decision variable by providing a wealth of output for the results cell(s). However, in contrast to many OR techniques, this approach has not identified an *optimal solution* for the decision variable(s). Fortunately, ASPE provides a way to systematically perform multiple simulations by using parameter cells. This makes it easy to identify at least an approximation of an optimal solution for problems with only one or two decision variables. In this section, we describe this approach and illustrate it by applying it in turn to the two decision variables listed above. (Recall that Sec. 20.6 included still another approach, using the Solver in ASPE to search for an optimal solution for simulation models.)

An intuitive approach for searching for an optimal solution is to use trial and error. Try different values of the decision variable(s), run a simulation for each, and see which one provides the best estimate of the chosen measure of performance. The interactive simulation mode in ASPE makes this especially easy, since the results in the statistic cells are available immediately after changing the value of a decision variable. Using parameter cells allows you to do the same thing in a more systematic way. After defining a parameter cell, all the desired simulations are run and the results soon are displayed nicely in the **parameter analysis report**. If desired, you also can view an enlightening **trend chart**, which can provide additional details about the results.

If you have previously used parameter cells with the Solver in ASPE to generate parameter analysis reports for performing sensitivity analysis systematically (as was done in Chap. 7), the parameter analysis reports in simulation models work in much the same way. Two is the maximum number of decision variables that can be varied simultaneously in a parameter analysis report.

Let us begin by returning to the bidding example mentioned above and use a parameter cell to run multiple simulations.

A Parameter Analysis Report for the Reliable Construction Co. Bidding Problem

We turn now to generating a parameter analysis report for the Reliable Construction Co. bidding problem presented in Sec. 28.1. Since the procedure for how to generate a parameter analysis report already has been presented in Sec. 20.6, our focus here is on summarizing the results.

Recall that the management of the company is concerned with determining what bid it should submit for a project that involves constructing a new plant for a major manufacturer. Therefore, the decision variable in the spreadsheet model in Fig. 28.2 is OurBid (C25). The parameter cell dialog box in Fig. 28.29 is used to further describe this decision variable. Management feels that the bid should be in the range between \$4.8 million and \$5.8 million, so these are the numbers (in units of millions of dollars) that are entered into the entry boxes for Bounds in this dialog box.

■ **FIGURE 28.29**

This parameter cell dialog box specifies the characteristics of the decision variable OurBid (C25) in Fig.28.2 for the Reliable Construction Co. contract bidding problem.

Function Arguments

PsiSimParam

Values_or_lower	4.8	=	4.8
Upper	5.8	=	5.8
Base_case		=	

PsiSimParam provides a list of different values that a variable should have in different simulations, where the value is the same for all trials in one simulation.

Upper

Formula result = 4.8

Help on this function

OK Cancel

Management wants to choose the bid that would maximize its expected profit. Consequently, the results cell in the spreadsheet model is Profit (C29). After choosing Parameter Analysis from the Reports>Simulation menu on the ASPE ribbon, the corresponding dialog box in Fig. 28.30 is used to specify that the mean of the Profit should be shown as the

■ **FIGURE 28.30**

This Parameter Analysis dialog box allows you to specify which parameter cells to vary and which results to show after each simulation run. Here the OurBid (C25) parameter cell will be varied over six different values and the value of the mean will be displayed for each of the six simulation runs.

Multiple Simulations Report

'Bidding!\$C\$29'

Statistics

- Mean
- Mode
- Variance
- StdDev
- Skewness
- Kurtosis

Parameters:

Bidding

- OurBid

Up Down

Vary Parameters Independently

Major Axis Points: 6 Minor Axis Points: 11

OK Cancel

FIGURE 28.31

The parameter analysis report for the Reliable Construction Co. contract bidding problem described in Sec. 28.1.

	A	B
1	OurBid	Mean
2	4.8	0.188
3	5.0	0.356
4	5.2	0.472
5	5.4	0.482
6	5.6	0.257
7	5.8	0.024

parameter cell OurBid is varied over six major axis points. The six values automatically are distributed evenly over the range specified in Fig. 28.29, so simulations will be run for bids of 4.8, 5.0, 5.2, 5.4, 5.6, and 5.8 (in millions of dollars).

Figure 28.31 shows the resulting parameter analysis report. A bid of \$5.4 million gives the largest mean value of the profits obtained on the 1,000 trials of the simulation run. This mean value of \$482,000 in cell B5 should be a close estimate of the expected profit from using this bid. The prototype example in Chap. 22 begins with the company having just won the contract by submitting this bid.

Problem 28.8 asks you to refine this analysis by generating a parameter analysis report that considers all bids between \$5.2 million and \$5.6 million in multiples of \$0.05 million.

A Parameter Analysis Report and Trend Chart for the Transcontinental Airlines Overbooking Problem

As described in Sec. 28.5, Transcontinental Airlines has a popular daily flight from San Francisco to Chicago with 150 seats available. The number of requests for reservations usually exceeds the number of seats by a considerable amount. However, even though the fare is nonrefundable, an average of only 80 percent of the customers who make reservations actually show up to take the flight, so it seems appropriate to accept more reservations than can be flown. At the same time, significant costs are incurred if customers with reservations are not allowed to take the flight. Therefore, the company's OR group is analyzing what number of reservations should be accepted to maximize the expected profit from the flight.

In the spreadsheet model in Fig. 28.16, the decision variable is ReservationsToAccept (C13) and the results cell is Profit (F23). The OR group wants to consider integer values of the decision variable over the range between 150 and 200, so the parameter cell dialog box is used in the usual way to specify these bounds on the variable. The decision is made to test 11 values of ReservationsToAccept (C13), so simulations will be run for values in intervals of five between 150 and 200.

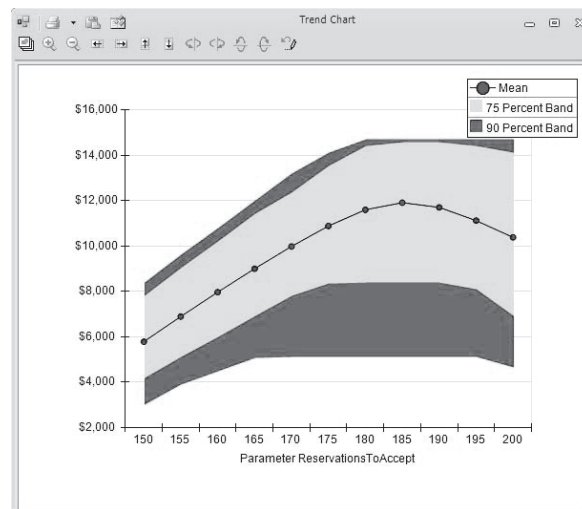
The results are shown in Fig. 28.32. The parameter analysis report on the left side of the figure reveals that the mean of the profit values obtained in the respective simulation runs climbs rapidly as ReservationsToAccept (C13) increases until the mean reaches a peak of \$11,912 at 185 reservations, after which it starts to drop. Only the means at 180 and 190 reservations are close to this peak, so it seems clear that the most profitable number of reservations lies somewhere between 180 and 190. (Now that the range of numbers that need to be considered has been narrowed down this far, Problem 28.10 asks you to continue that analysis by generating a parameter analysis report that considers all integer values over this range.)

The trend chart on the right side of Fig. 28.32 provides additional insight. The bands in this chart trend upward until the number of reservations to accept reaches approximately

■ **FIGURE 28.32**

The parameter analysis report and trend chart for the Transcontinental Airlines overbooking problem described in Sec. 28.5.

	A	B
1	ReservationsToAccept	Mean
2	150	\$5,789
3	155	\$6,896
4	160	\$7,968
5	165	\$9,001
6	170	\$9,982
7	175	\$10,880
8	180	\$11,592
9	185	\$11,912
10	190	\$11,712
11	195	\$11,124
12	200	\$10,395



185; then they start trending slowly downward. This indicates that the *entire* frequency distribution from the respective simulation runs keeps shifting upward until the run for 185 reservations and then starts shifting downward. Also note that the width of the entire set of seven bands increases until about the simulation run for 180 reservations and then remains about the same thereafter. This indicates that the amount of variability in the profit values also increases until the simulation run for 180 reservations and then remains about the same thereafter.

■ 28.8 SUMMARY

Increasingly, spreadsheet software is being used to perform simulations. As illustrated in Secs. 20.1 and 20.4, the standard Excel package sometimes is sufficient to do this. In addition, some Excel add-ins now are available that greatly extend these capabilities. ASPE is an especially powerful add-in of this kind.

When using ASPE, each input cell that has a random value is referred to as an *uncertain variable cell*. The procedure for defining an uncertain variable cell includes selecting one of 46 types of probability distributions from the Distributions menu to enter into the cell. When historical data are available, ASPE also has a procedure for identifying which continuous distribution fits the data best.

An output cell that is used to forecast a measure of performance is called a *results cell*. Each trial of a simulation run generates a value in each results cell. When the simulation run is completed, ASPE provides the results in a variety of useful forms, including a frequency distribution, a statistics table, a percentiles table, and a cumulative chart.

When a simulation model has one or two decision variables, ASPE provides a *parameter analysis report* that systematically applies simulation to identify at least an approximation of an optimal solution. A *trend chart* also provides additional insights to aid in decision making.

In addition, ASPE includes a powerful optimization module called Solver. This module efficiently uses a series of simulation runs to search for an optimal solution for a simulation model with any number of decision variables.

The availability of such powerful software now enables managers to add simulation to their personal tool kit of OR techniques for analyzing some key managerial problems. A variety of examples in this chapter illustrate some of the many possibilities for important applications of simulation.

■ SELECTED REFERENCES

For general references on simulation, see the Selected References given for Chap. 20. For further information regarding Frontline Systems, Analytic Solver Platform, and ASPE, go to www.solver.com.

■ LEARNING AIDS FOR THIS CHAPTER ON THIS WEBSITE

See the learning aids for Chap. 20. Additional learning aids are Excel files that provide the spreadsheet models for the examples in this chapter, as well as Sales Data 1 and Sales Data 2 for two end-of-chapter problems.

■ PROBLEMS

ASPE should be used for all of the following problems.

28.1. Consider the Reliable Construction Co. project scheduling example presented in Sec. 28.2. Recall that simulation was used to estimate the probability of meeting the deadline and that Fig. 28.8 revealed that the deadline was met on 57.7 percent of the trials from one simulation run. As discussed while interpreting this result, the percentage of trials on which the project is completed by the deadline will vary from simulation run to simulation run. This problem will demonstrate this fact and investigate the impact of the number of trials per simulation on this randomness. The spreadsheet model is available on this website. Make sure

that the Monte Carlo sampling method is chosen in Simulation Options.

- (a) Set the trials per simulation to 100 in Simulation Options and run the simulation of the project five times. Note the mean completion time and the percentage of trials on which the project is completed within the deadline of 47 weeks for each simulation run.
- (b) Repeat part *a* except set the trials per simulation to 1,000 in Simulation Options.
- (c) Compare the results from part *a* and part *b* and comment on any differences.

Activity	Predecessors
A. Secure funding	–
B. Design building	A
C. Site preparation	A
D. Foundation	B, C
E. Framing	D
F. Electrical	D
G. Plumbing	E
H. Walls and roof	F, G
I. Finish construction	H
J. Landscaping	H

Construction Time Estimates (months)

Activity	Optimistic Scenario	Most Likely Scenario	Pessimistic Scenario
C. Site preparation	1.5	2	2.5
D. Foundation	1.5	2	3
E. Framing	3	4	6
F. Electrical	2	3	5
G. Plumbing	3	4	5
H. Walls and roof	4	5	7
I. Do the finish work	5	6	7

28.2. Consider the historical data contained in the Excel File “Sales Data 1” on this website. Use ASPE to fit continuous distributions to these data.

- Which distribution provides the closest fit to the data? What are the parameters of the distribution?
- Which distribution provides the second-closest fit to the data? What are the parameters of the distribution?

28.3. Consider the historical data contained in the Excel File “Sales Data 2” on this website. Use ASPE to fit continuous distributions to these data.

- Which distribution provides the closest fit to the data? What are the parameters of the distribution?
- Which distribution provides the second-closest fit to the data? What are the parameters of the distribution?

28.4. Ivy University is planning to construct a new building for its engineering school. This project will require completing all of the activities in the above table. For most of these activities, a set of predecessor activities must be completed before the activity begins. For example, the foundation cannot be laid until the building is designed and the site prepared.

Obtaining funding likely will take approximately six months (with a standard deviation of one month). Assume that this time has a normal distribution. The architect has estimated that the time required to design the building could be anywhere between 6 and 10 months. Assume that this time has a uniform distribution. The general contractor has provided three estimates for each of the construction tasks—an optimistic scenario (minimum time required if the weather is good and all goes well), a most likely scenario, and a pessimistic scenario (maximum time required if there are weather and other problems). These estimates are provided in the table that follows. Assume that each of these construction times has a triangular distribution. Finally, the landscaper has guaranteed that his work will be completed in five months.

Use ASPE to perform 1,000 trials of a simulation for this project. Use the results to answer the following questions.

- What is the mean project completion time?
- What is the probability that the project will be completed in 36 months or less?
- Generate a sensitivity chart. Based on this chart, which activities have the largest impact on the project completion time?

28.5. The employees of General Manufacturing Corp. receive health insurance through a group plan issued by Wellnet. During the past year, 40 percent of the employees did not file any health insurance claims, 30 percent filed only a small claim, and 20 percent filed a large claim. The small claims were spread uniformly between 0 and \$2,000, whereas the large claims were spread uniformly between \$2,000 and \$20,000.

Based on this experience, Wellnet now is negotiating the corporation’s premium payment per employee for the upcoming year. To obtain a close estimate of the average cost of insurance coverage for the corporation’s employees, use ASPE with a spreadsheet to perform 1,000 trials of a simulation of an employee’s health insurance experience. Generate a frequency chart and a statistics table.

28.6. Refer to the financial risk analysis example presented in Sec. 28.4, including its results shown in Fig. 28.15. Think-Big management is quite concerned about the risk profile for the proposal. Two statistics are causing particular concern. One is that there is nearly a 20 percent chance of losing money (a negative NPV). Second, there is more than a 6 percent chance of losing more than half (\$10 million) as much as the mean gain (\$18 million). Therefore, management is wondering whether it would be more prudent to go ahead with just one of the two projects. Thus, in addition to option 1 (the proposal), option 2 is to take a 16.50 percent share of the hotel project only (so no participation in the shopping center project), and option 3 is to take a 13.11 percent share of the shopping center only (so no participation in the hotel project). Management wants to choose one of the three options. Risk profiles now are needed to evaluate the latter two.

- Estimate the mean NPV and the probability that the NPV will be greater than 0 for option 2 after performing a simulation with 1,000 trials for this option.
- Repeat part *a* for option 3.

(c) Suppose *you* were the CEO of the Think-Big Development Co. Use the results in Fig. 28.15 for option 1 along with the corresponding results obtained for the other two options as the basis for a managerial decision on which of the three options to choose. Justify your answer.

28.7. Susan is a ticket scalper. She buys tickets for Los Angeles Lakers games before the beginning of the season for \$100 each. Since the games all sell out, Susan is able to sell the tickets for \$150 on game day. Tickets that Susan is unable to sell on game day have no value. Based on past experience, Susan has predicted the probability distribution for how many tickets she will be able to sell, as shown in the following table.

Tickets	Probability
10	0.05
11	0.10
12	0.10
13	0.15
14	0.20
15	0.15
16	0.10
17	0.10
18	0.05

- (a) Suppose that Susan buys 14 tickets for each game. Use ASPE to perform 1,000 trials of a simulation on a spreadsheet. What will be Susan's mean profit from selling the tickets? What is the probability that Susan will make at least \$0 profit? (*Hint:* Use the Custom Discrete distribution to simulate the demand for tickets.)
- (b) Generate a parameter analysis report to consider all nine possible quantities of tickets to purchase between 10 and 18. Which purchase quantity maximizes Susan's mean profit?
- (c) Generate a trend chart for the nine purchase quantities considered in part *b*.
- (d) Use ASPE's Solver to search for the purchase quantity that maximizes Susan's mean profit.

28.8. Consider the Reliable Construction Co. bidding problem discussed in Sec. 28.1. The spreadsheet model is available on this website. The parameter analysis report generated in Sec. 28.7 (see Fig. 28.31) for this problem suggests that \$5.4 million is the best bid, but this table only considered bids that were a multiple of \$0.2 million.

- (a) Refine the search by generating a parameter analysis report for this bidding problem that considers all bids between \$5.2 million and \$5.6 million in multiples of \$0.05 million.
- (b) Use ASPE's Solver to search for the bid that maximizes Reliable Construction Co.'s mean profit. Assume that the bid may be any value between \$4.8 million and \$5.8 million.

28.9. Consider the Everglade cash flow problem analyzed in Sec. 28.3. The spreadsheet model is available on this website.

- (a) Generate a parameter analysis report to consider five possible long-term loan amounts between \$0 million and \$20 million and forecast Everglade's mean ending balance. Which long-term loan amount maximizes Everglade's mean ending balance?
- (b) Generate a trend chart for the five long-term loan amounts considered in part *a*.
- (c) Use ASPE's Solver to search for the long-term loan amount that maximizes Evergreen's mean ending balance.

28.10. Consider the airline overbooking problem discussed in Sec. 28.5. The spreadsheet model is available on this website. The parameter analysis report generated in Sec. 28.7 (see Fig. 28.32) for this problem suggests that 185 is the best number of reservations to accept in order to maximize profit, but the only numbers considered were a multiple of five.

- (a) Refine the search by generating a parameter analysis report for this overbooking problem that considers all integer values for the number of reservations to accept between 180 and 190.
- (b) Generate a trend chart for the 11 forecasts considered in part *a*.
- (c) Use ASPE's Solver to search for the number of reservations to accept that maximizes the airline's mean profit. Assume that the number of reservations to accept may be any integer value between 150 and 200.

■ ACKNOWLEDGMENT

A somewhat longer version of this chapter (with slight differences) also appears as Chap. 13 in the 5th edition of *Introduction to Management Science: A Modeling and Case Studies Approach with Spreadsheets* by Frederick S. Hillier

and Mark S. Hillier, McGraw-Hill/Irwin, 2014. We gratefully acknowledge the major role that Mark S. Hillier played in developing this chapter.