

# 28

## CHAPTER

---

# Examples of Performing Simulations on Spreadsheets with Crystal Ball

Section 20.6 introduced the subject of how to perform simulations on a spreadsheet while using the powerful Excel add-in, Crystal Ball (Professional Edition). This chapter elaborates considerably further on the same subject.

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 Crystal Ball. Section 28.6 then describes Crystal Ball’s Distribution Gallery and the characteristics of the 21 probability distributions it provides. A special procedure that Crystal Ball provides for identifying the continuous distribution that best fits historical data is also outlined. Section 28.7 illustrates the application of Crystal Ball’s Decision Table tool to two of the previous examples.

## 28.1 BIDDING FOR A CONSTRUCTION PROJECT

Managers frequently must make decisions whose outcome 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 decision of setting the price for a new product that is being brought to the market. How well this decision works out will depend heavily on the pricing decisions being made nearly simultaneously by other firms marketing similar new products. Similarly, the success of a decision on when to market a product under development will be determined largely by whether this product reaches the market before other firms release competitive products.

When a decision must be made before learning of the corresponding decisions being made by competitors, the analysis will need to take into account the uncertainty regarding what the competitors’ decisions will be. Simulation provides a natural way of doing this by using assumption cells to represent the 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 Crystal Ball to guide its choice of \$5.4 million as the bid for the project. You will not need to review the presentation in Chap. 22 to follow the current example.

The first step was to estimate what the company's total cost would be if it were to undertake this project. This was determined to be \$4.55 million. (This amount excludes the penalty for missing the deadline for completion of the project, as well as the bonus for completion well before the deadline, since management considers either event to be relatively unlikely.) There is also 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 realized 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 is also a particularly unpredictable bidder because of its inability to accurately estimate the true costs of a project. 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 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. Therefore, 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, are the following.

- 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 28.1 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 *assumption 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, entering any number into the cell,

28.1 BIDDING FOR A CONSTRUCTION PROJECT

	A	B	C	D	E
1		<b>Reliable Construction Co. Contract Bidding</b>			
2					
3		<b>Data</b>			
4		Our Project Cost (\$million)	4.550		
5		Our Bid Cost (\$million)	0.050		
6					
7		<b>Competitor Bids</b>	Competitor 1	Competitor 2	Competitor 3
8		Bid (\$million)	5.839	5.688	5.688
9					
10		Distribution	Triangular	Triangular	Uniform
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		<b>Minimum Competitor</b>			
23		<b>Bid (\$million)</b>	5.688		
24					
25		<b>Our Bid (\$million)</b>	5.4		
26					
27		<b>Win Bid?</b>	1	(1=yes, 0=no)	
28					
29		<b>Profit (\$million)</b>	0.800		

	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	Range Name	Cells
22	<b>Minimum Competitor</b>		CompetitorBids	C8:E8
23	<b>Bid (\$million)</b>	=MIN(C8:E8)	MinimumCompetitorBid	C23
24			OurBid	C25
25	<b>Our Bid (\$million)</b>	5.4	OurBidCost	C5
26			OurProjectCost	C4
27	<b>Win Bid?</b>	=IF(OurBid<MinimumCompetitorBid,1,0)	Profit	C29
28			WinBid?	C27
29	<b>Profit (\$million)</b>	=WinBid?*(OurBid-OurProjectCost)-OurBidCost		

■ FIGURE 28.1

A spreadsheet model for applying simulation to the Reliable Construction Co. contract bidding problem. The assumption cells are CompetitorBids (C8:E8), the forecast cell is Profit (C29), and the decision variable is OurBid (C25).

and then clicking on the appropriate distribution in the Distribution Gallery, which brings up the dialogue box for that distribution. Figure 28.2 shows the Triangular Distribution dialogue box that has been used to set the parameter values (Minimum, Likeliest, and Maximum) for competitor 1. Competitor 2 would be handled in a similar way. 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 dialogue box, such as the one shown in Fig. 20.9 in Sec. 20.6, is used instead to set the parameter values for 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.

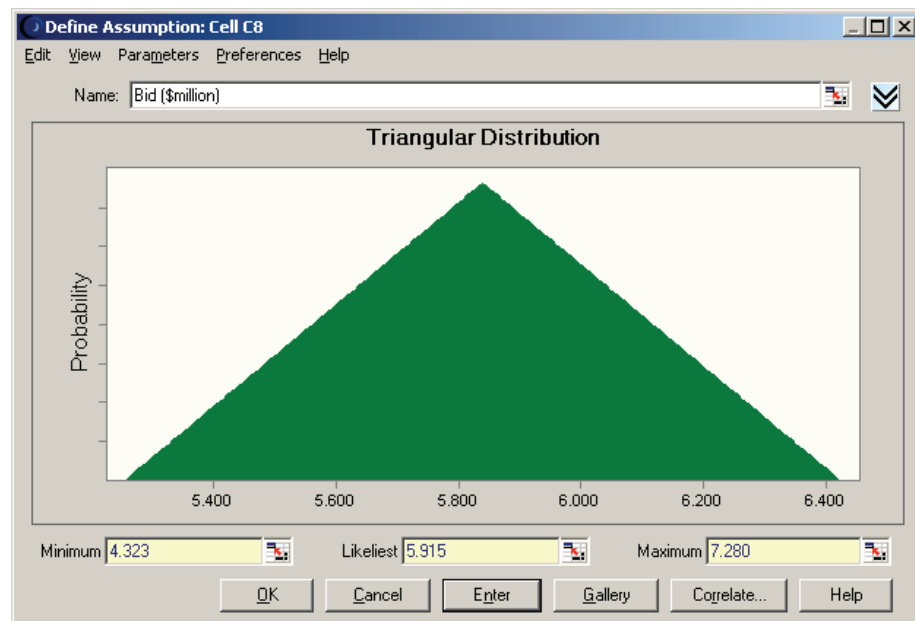
Since management wants to maximize the expected profit from the bidding process and then (if the bid wins) from doing the project, the forecast cell in this model is Profit (C29). The profit achieved on a given trial depends on whether the company wins the bid. If not, the profit actually represents a loss of \$50,000 (the cost of preparing the bid). However, if the bid wins, the profit is the amount by which the bid exceeds the sum of the project cost and the bid-preparation cost. The equation entered into Profit (C29) performs this calculation for whichever case applies.

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

**Assumption cells:** CompetitorBids (C8:E8)  
**Decision variable:** OurBid (C25)  
**Forecast cell:** Profit (C29)

■ **FIGURE 28.2**

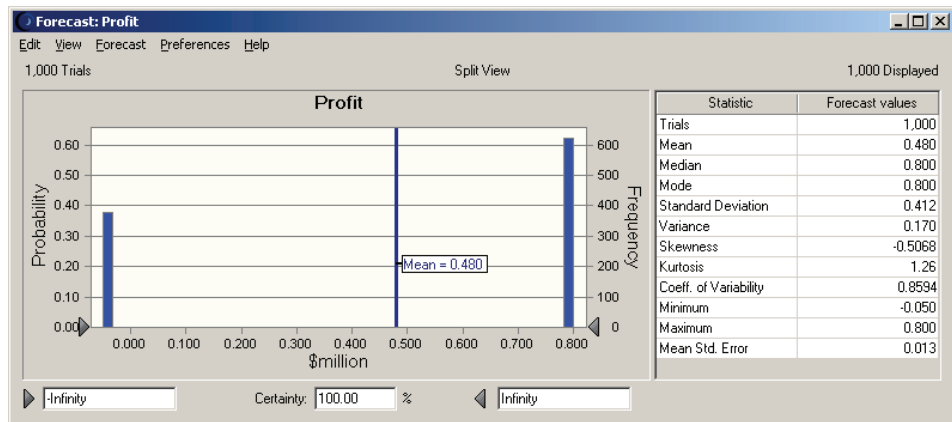
The Triangular Distribution dialogue box. It is being used here to enter a triangular distribution with the parameters 4.323 (=C18), 5.915 (=C19), and 7.280 (=C20) into the assumption cell C8 in the spreadsheet model in Fig. 28.1.

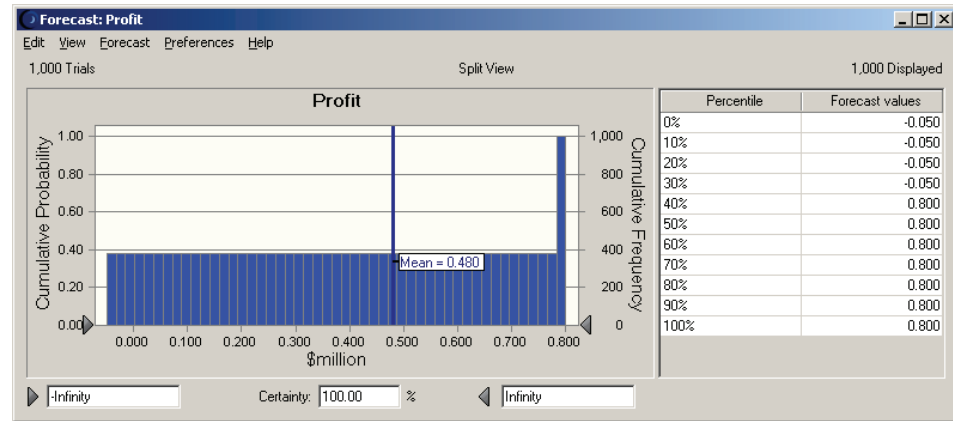


**The Simulation Results**

To evaluate a possible bid of \$5.4 million that is entered into OurBid (C25), this simulation model was run for 1,000 trials. Figure 28.3 shows the results in the form of a frequency chart and a statistics table, while Fig. 28.4 displays the corresponding percentiles table and cumulative chart. 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 trials. This resulted in a mean profit of  $0.480$  (\$480,000) from all 1,000 trials, as well

**FIGURE 28.3**  
The frequency chart and statistics table that summarize the results of running the simulation model in Fig. 28.1 for the Reliable Construction Co. contract bidding problem.





■ **FIGURE 28.4**  
Further results for the Reliable Construction Co. contract bidding problem.

as the other statistics recorded in the statistics table. In Fig. 28.4, note how the possibility of only two profit values results in only these values appearing in the Percentiles table and also results in having a flat cumulative chart until the upper value is reached.

By themselves, these results do not show that \$5.4 million is the best bid to submit. We still need to perform additional simulation runs to estimate whether a larger expected profit could be obtained by using a different bid value. Section 28.7 will describe how doing this with a decision table 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 manager estimate the probability of meeting the deadline with any particular plan? Section 22.4 described one method using PERT/CPM. We now will illustrate how simulation provides an alternative method that usually yields a better estimate of meeting the project deadline.

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 assumption cells where the random inputs are *times*. Another interesting feature of this example is its use of a special kind of Crystal Ball chart called the *sensitivity chart*. This chart will provide a key insight into how the project plan should be revised to help ensure that the deadline is met.

### The Problem Being Addressed

As in the preceding section, this example 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 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. The Reliability Construction Co. 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, which is 47 weeks away. 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 Fig. 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. Those six sequences are listed below.

Path 1: START → A → B → C → D → G → H → M → FINISH

Path 2: START → A → B → C → E → H → M → FINISH

Path 3: START → A → B → C → E → F → J → K → N → FINISH

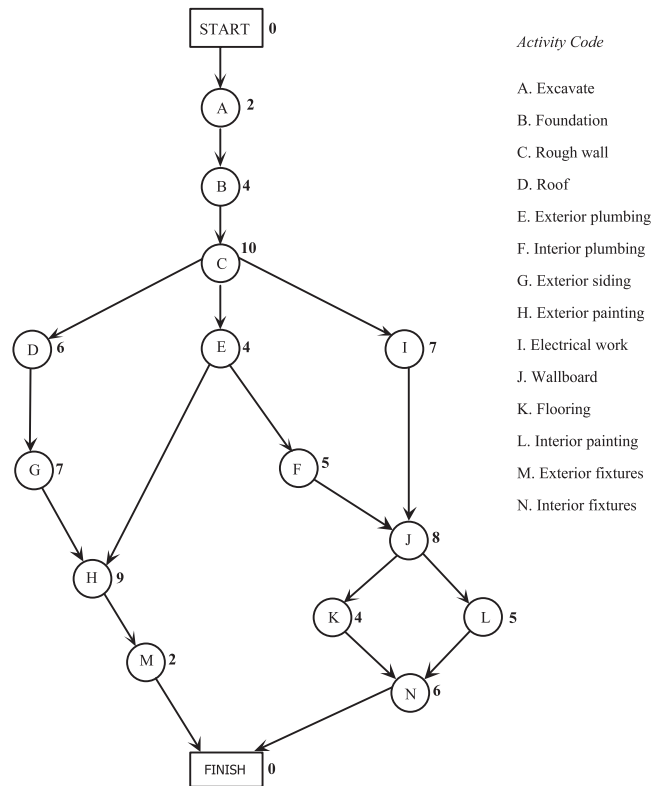
Path 4: START → A → B → C → E → F → J → L → N → FINISH

Path 5: START → A → B → C → I → J → K → N → FINISH

Path 6: START → A → B → C → I → J → L → N → FINISH

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 on. 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 differ substantially



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

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 project duration.

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.

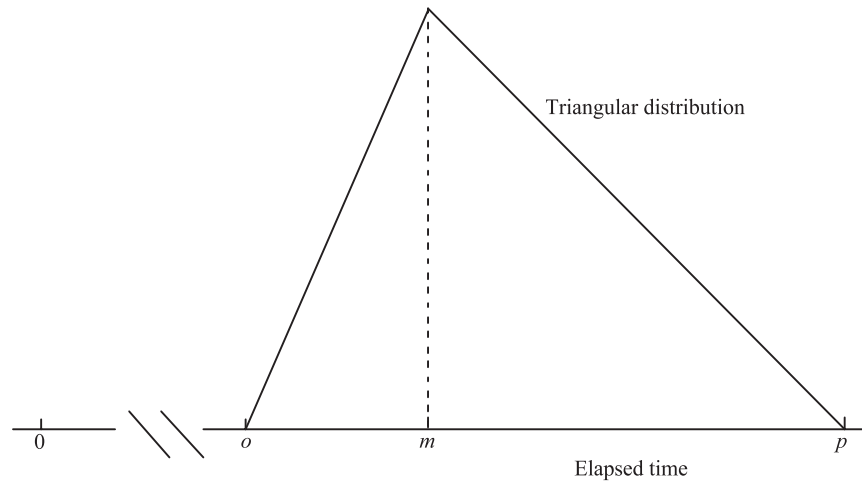
One key advantage of simulation is that it does not need to make most of the simplifying approximations that may be required by analytical methods. Another is that there is great flexibility about which probability distribution to use. It is not necessary to choose an analytically convenient one.

When dealing with the duration of an activity, simulations commonly use a *triangular distribution* as the distribution of this duration. A triangular distribution for the duration of an activity has the shape shown in Fig. 28.6, where *o*, *m*, and *p* are the labels for the optimistic estimate, the most likely estimate, and the pessimistic estimate, respectively. For each *assumption cell* containing this distribution, a Triangular Distribution dialogue box (such as the one shown in Fig. 28.2) is used to enter the values of *o*, *m*, and *p* by entering their respective cell references into the Minimum, Likeliest, and Maximum boxes.



■ **FIGURE 28.6**

The shape of a triangular distribution for the duration of an activity, where the minimum lies at the optimistic estimate  $o$ , the most likely value lies at the most likely estimate  $m$ , and the maximum lies at the pessimistic estimate  $p$ .



### A Spreadsheet Model for Applying Simulation

Figure 28.7 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 forecast cell ProjectCompletion (I21).

Since the activity times generally are variable, the cells H6:H19 all need to be *assumption cells* with the one exception of cell H16. Since  $o = m = p = 4$  for activity K, its activity time actually is a constant 4, so this constant is entered into cell H16. (Crystal Ball would give an error message if cell H16 were specified to be an assumption cell with a triangular distribution where Minimum = Likeliest = Maximum.) Figure 28.8 shows the Triangular Distribution dialogue box after it has been used to specify the parameters for the first assumption cell H6. Rather than repeating this process for all the other assumption cells, it is quicker to simply copy and paste the parameters for the other assumption cells. This is begun by selecting cell H6 and clicking on the Copy Data button on the Crystal Ball tab or toolbar. Then select the cells in which to paste the data (H7:H15 and H17:H19), and choose Paste Data by clicking this button on the Crystal Ball tab or toolbar. The cell references for the parameters in Fig. 28.8 are entered as =D6, =E6, and =F6 (relative references without \$ signs). Therefore, the row numbers will update appropriately to refer to the data cells in the correct row during this copy-and-paste process. For example, the cell references for the parameters of the triangular distribution in cell H7 will update to =D7, =E7, and =F7.

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

**Assumption cells:** Cells H6:H15 and H17:H19

**Forecast cell:** ProjectCompletion (I21)

### The Simulation Results

We now are ready to perform simulation with the spreadsheet model in Fig. 28.7. Using the Run Preferences dialogue box to specify 1,000 trials, Fig. 28.9 shows the results in the form of a frequency chart, a statistics table, and a percentiles table. These results show

	A	B	C	D	E	F	G	H	I
1	<b>Simulation of Reliable Construction Co. Project</b>								
2									
3								Activity	
4		Immediate		Time Estimates			Start	Time	Finish
5	Activity	Predecessor	o	m	p	Time	(triangular)	Time	
6	A	—	1	2	3	0	2	2	
7	B	A	2	3.5	8	2	4.5	6.5	
8	C	B	6	9	18	6.5	11	17.5	
9	D	C	4	5.5	10	17.5	6.5	24	
10	E	C	1	4.5	5	17.5	3.5	21	
11	F	E	4	4	10	21	6	27	
12	G	D	5	6.5	11	24	7.5	31.5	
13	H	E, G	5	8	17	31.5	10	41.5	
14	I	C	3	7.5	9	17.5	6.5	24	
15	J	F, I	3	9	9	27	7	34	
16	K	J	4	4	4	34	4	38	
17	L	J	1	5.5	7	34	4.5	38.5	
18	M	H	1	2	3	41.5	2	43.5	
19	N	K, L	5	5.5	9	38.5	6.5	45	
20									
21							Project Completion	45	

	G	H	I
3		Activity	
4	Start	Time	Finish
5	Time	(triangular)	Time
6	0	2	=AStart+ATime
7	=AFinish	4.5	=BStart+BTime
8	=BFinish	11	=CStart+CTime
9	=CFinish	6.5	=DStart+DTime
10	=CFinish	3.5	=EStart+ETime
11	=EFinish	6	=FStart+FTime
12	=DFinish	7.5	=GStart+GTime
13	=MAX(EFinish,GFinish)	10	=HStart+HTime
14	=CFinish	6.5	=IStart+ITime
15	=MAX(FFinish,IFinish)	7	=JStart+JTime
16	=JFinish	4	=KStart+KTime
17	=JFinish	4.5	=LStart+LTime
18	=HFinish	2	=MStart+MTime
19	=MAX(KFinish,LFinish)	6.5	=NStart+NTime
20			
21		Project Completion	=MAX(MFinish,NFinish)

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
MStart	G18
MTime	H18
NFinish	I19
NStart	G19
NTime	H19
ProjectCompletion	I21

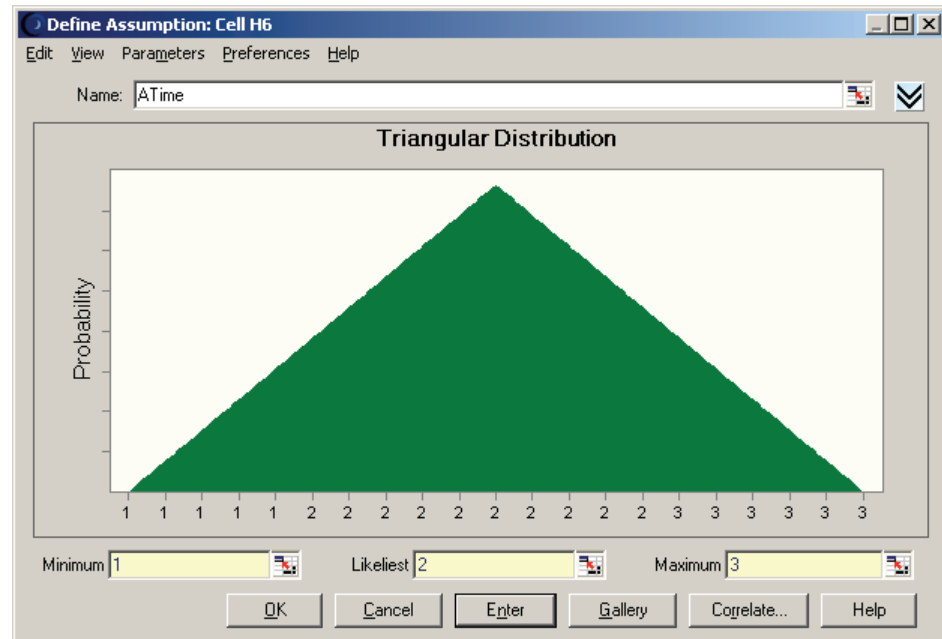
■ **FIGURE 28.7**

A spreadsheet model for applying simulation to the Reliable Construction Co. project scheduling problem. The assumption cells are cells H6:H15 and H17:H19. The forecast cell is ProjectCompletion (I21).

a very wide range of possible project durations. Out of the 1,000 trials, both the statistics table and percentiles table indicate that one trial had a duration as short as 35.98 weeks while another was as long as 58.81 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 46.20 weeks, which is much too close to the deadline of 47 weeks to leave much margin for slippage in the project schedule. (The small mean standard error of 0.12 week reported at the bottom of the

■ **FIGURE 28.8**

A triangular distribution with parameters 1 (=D6), 2 (=E6), and 3 (=F6) is being entered into the first assumption cell H6 in the spreadsheet model in Fig. 28.7.



statistics table shows that the sample average of 46.20 weeks from the 1,000 trials probably is extremely close to the true mean of the underlying probability distribution of project duration.) The 70 percent percentile of 48.05 weeks in the percentiles table reveals that 30 percent of the trials missed the deadline by a least a week.

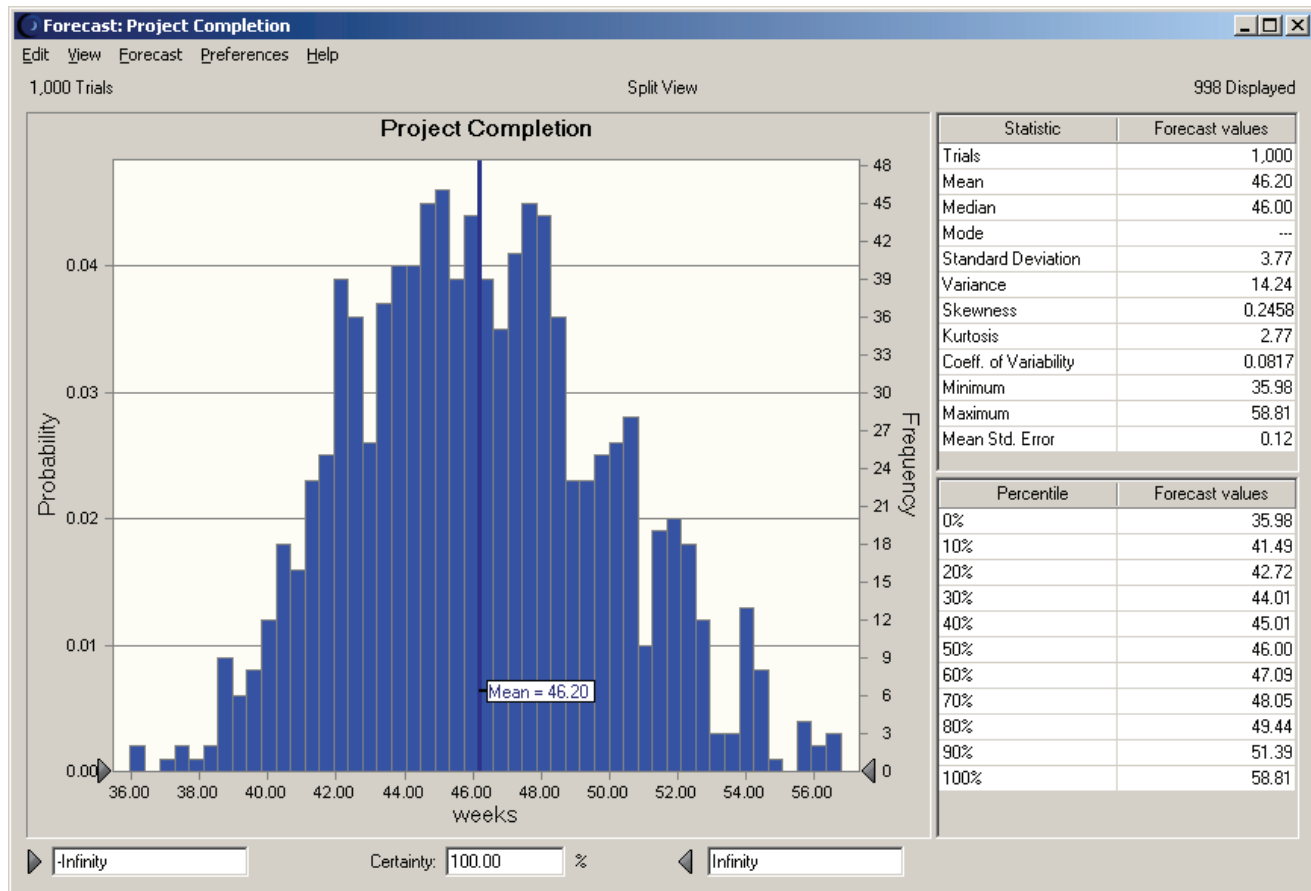
A statistic of special interest to Reliable's management is the probability of meeting the deadline of 47 weeks under the current project plan. (Remember that the contract includes a severe penalty of \$300,000 for missing this deadline.) The 60 percent percentile of 47.09 weeks in the percentiles table suggests that slightly under 60 percent of the trials made the deadline, but nothing else in Fig. 28.9 pins down this percentage any closer. Figure 28.10 shows that all you need to do to identify the exact percentage is to type the deadline of 47.00 in the box in the lower right-hand corner of the frequency chart. The Certainty box then reveals that about 58.88 percent of the trials met the deadline.<sup>1</sup>

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 0.588 provided by the Certainty box in Fig. 28.10 is a close estimate of the true probability of meeting the deadline under the assumptions of the spreadsheet model in Fig. 28.7. 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

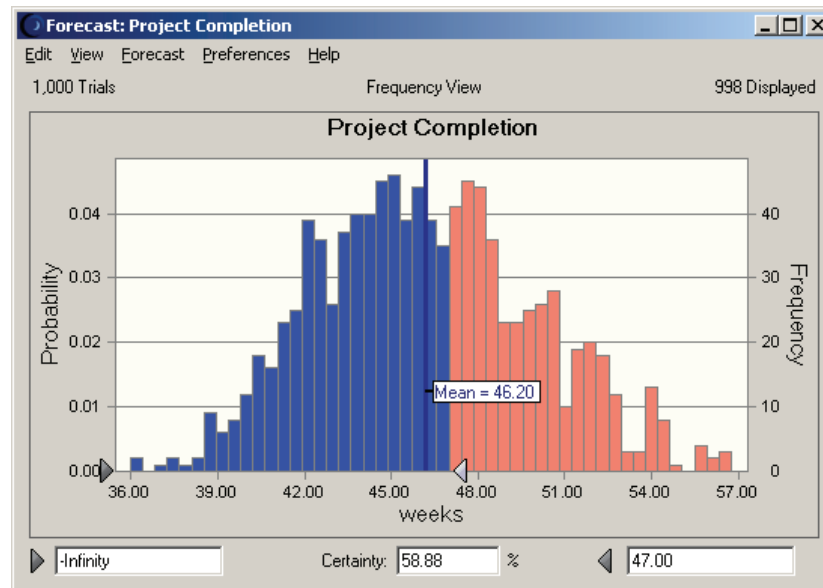
Given such a low probability (0.615) of meeting the project deadline, Reliable's project manager (David Perty) will want to revise the project plan to improve the probability substantially. Crystal Ball has another tool, called the *sensitivity chart* that provides guidance in

1. Actually 588 of the 1,000 trials met the deadline. The certainty is shown as 58.88 rather than 58.8 because Crystal Ball further refines the certainty estimate based on where the 47-week deadline falls relative to the 588th longest trial (which just barely met the deadline) and the 589th longest (which just barely did not).



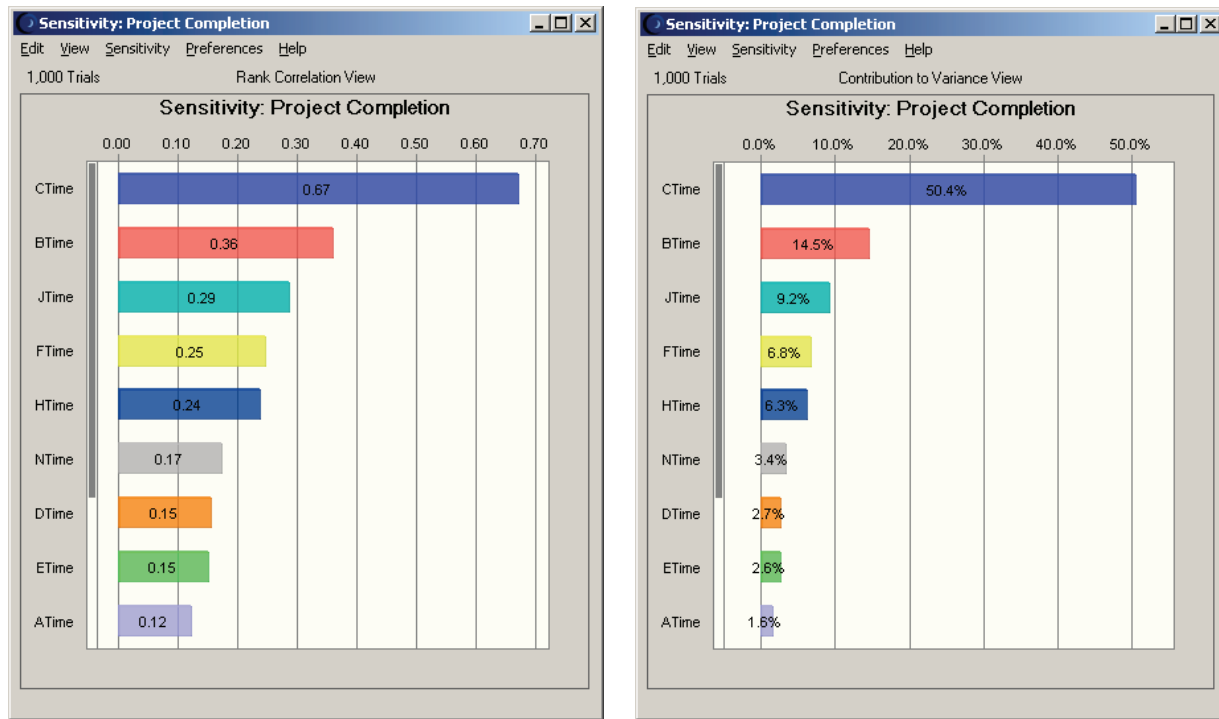
**FIGURE 28.9**  
 The frequency chart, statistics table, and percentiles table that summarize the results of running the simulation model in Fig. 28.7 for the Reliable Construction Co. project scheduling problem.

■ **FIGURE 28.10**  
After entering the project deadline of 47 weeks into the box in the lower right-hand corner, the Certainty box reveals that 58.88 percent of the trials resulted in the project being completed by the deadline.



identifying which revisions in the project plan would be most beneficial. To open a sensitivity chart after running a simulation, click on the Open Sensitivity Chart button on the Crystal Ball tab or toolbar. Both a contribution to variance chart and a rank correlation chart are then available under the View menu, as shown in Fig. 28.11. Using range names, the left side of both charts identify various assumption cells (activity times) in column H of the spreadsheet model in Fig. 28.7. The other key cell in the spreadsheet model that is considered in the sensitivity charts (as indicated at the top of both charts) is the forecast cell ProjectCompletion (I21) that gives the duration of the project. The contribution to variance chart indicates what percentage of the variability that exists in the forecast cell is due to the variability of each assumption cell. The rows in the chart are listed in descending order. For example, roughly half the variability in project completion time is due to variability in CTime (the time to complete Activity C: Rough Wall).

The rank correlation chart gives the correlation coefficient (based on rank values) between each assumption cell and the forecast cell.



■ **FIGURE 28.11**

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

A correlation coefficient between two variables measures the strength of the relationship between these variables. Thus, each correlation coefficient in Fig. 28.12 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 28.11 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 28.7 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. 28.7 indicate that CTime is highly variable, with an unusually large spread of nine 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 suggests that the best way to reduce the project completion time 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. Crystal Ball'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 it is impossible to predict the future with certainty, simulation is needed to take future uncertainties into account when making decisions. For example, businesses typically have a great deal of uncertainty about what their future cash flows will be. An attempt is often 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 future cash flows. Simulation does this by using assumption cells to represent 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 at the end of 2002 to manage its cash flow problems. Because of both a temporary decline in business and high construction costs, the company at that time was facing some negative cash flows in the next few years as well as in a couple more distant years. As first provided in Table 21.1, Table 28.1 shows projected net cash flows for the next 10 years (2010 to 2019). The company has some new retirement communities that will be opening during the 10-year period, so it is anticipated (or at least hoped) that a large positive cash flow will occur in 2019. Therefore, the problem confronting Everglade management is how to best arrange Everglade's financing to tide the company over until its investments in the 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 immediately (the beginning of 2010) and a series of short-term (1-year) loans as needed

■ **TABLE 28.1** Projected net cash flows for the Everglade Golden Years Company (as of the end of 2009) over the next 10 years

Year	Projected Net Cash Flow (millions of dollars)
2010	-8
2011	-2
2012	-4
2013	3
2014	6
2015	3
2016	-4
2017	7
2018	-2
2019	10

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 2020 when all of the loans have been paid off. Figure 21.5 in Chap. 21 shows the complete spreadsheet model after using the Excel Solver to obtain the optimal solution. (Figure 21.5 is repeated here as Fig. 28.12.) The changing cells, LTLoan (D11) and STLoan (E11:E20), give the size of the long-term loan and of the short-term loans at the beginning of the various years. The target cell EndBalance (J21) indicates that the resulting cash balance at the end of the 10 years (the beginning of 2020) would be \$2.92 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. 28.12 is an excellent first step in developing a final plan. However, the drawback of this spreadsheet model 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 (2010) 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 thus needed to assess the effect of these uncertainties.

### A Spreadsheet Model for Applying Simulation

Figure 28.13 shows the modification of the spreadsheet model in Fig. 28.12 that is needed to apply simulation. One key difference is that the constants in CashFlow (C11:C20) in Fig. 28.12 have turned into random inputs in CashFlow (F12:F21) in Fig. 28.13. Thus, the latter cells CashFlow (F12:F21) are *assumption cells*. (The numbers appearing in these cells have been entered arbitrarily as a first step in defining these assumption cells.) 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 (minimum, most likely, and maximum) for each of the years, as presented in cells C12:E21.

The number 6.65 entered into LTLoan (G12) is the size of the long-term loan (in millions of dollars) that was obtained in Fig. 28.12. 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. 28.12. 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



28.3 CASH FLOW MANAGEMENT

1	<b>Everglade Cash Flow Management Problem</b>										
2											
3		LT Rate	7%								
4		ST Rate	10%								
5											
6		Start Balance	1	(all cash figures in millions of dollars)							
7		Minimum Cash	0.5								
8											
9			Cash	LT	ST	LT	ST	LT	ST	Ending	Minimum
10	Year	Flow	Loan	Loan	Interest	Interest	Payback	Payback	Payback	Balance	Balance
11	2010	-8	6.649	0.851						0.500	>= 0.5
12	2011	-2		3.401	-0.465	-0.085		-0.851		0.500	>= 0.5
13	2012	-4		8.207	-0.465	-0.340		-3.401		0.500	>= 0.5
14	2013	3		6.493	-0.465	-0.821		-8.207		0.500	>= 0.5
15	2014	6		1.607	-0.465	-0.649		-6.493		0.500	>= 0.5
16	2015	3		0.000	-0.465	-0.161		-1.607		1.266	>= 0.5
17	2016	-4		3.699	-0.465	0.000		0.000		0.500	>= 0.5
18	2017	7		0.000	-0.465	-0.370		-3.699		2.965	>= 0.5
19	2018	-2		0.000	-0.465	0.000		0.000		0.500	>= 0.5
20	2019	10		0.000	-0.465	0.000		0.000		10.035	>= 0.5
21	2020				-0.465	0.000	-6.649	0.000		2.920	>= 0.5

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

**Solver Options**

Assume Linear Model

Assume Non-Negative

**Solver Parameters**

Set Target Cell: EndBalance

Equal To:  Max  Min  Value Of

By Changing Cells: LTLoan,STLoan

Subject to the Constraints: EndingBalance >= MinimumBalance

	F	G	H	I	J	K	L
9	LT	ST	LT	ST	Ending		Minimum
10	Interest	Interest	Payback	Payback	Balance		Balance
11					=StartBalance+SUM(C11:I11)	>=	=MinimumCash
12	=-LTRate*LTLoan	=-STRate*E11		=-E11	=J11+SUM(C12:I12)	>=	=MinimumCash
13	=-LTRate*LTLoan	=-STRate*E12		=-E12	=J12+SUM(C13:I13)	>=	=MinimumCash
14	=-LTRate*LTLoan	=-STRate*E13		=-E13	=J13+SUM(C14:I14)	>=	=MinimumCash
15	=-LTRate*LTLoan	=-STRate*E14		=-E14	=J14+SUM(C15:I15)	>=	=MinimumCash
16	=-LTRate*LTLoan	=-STRate*E15		=-E15	=J15+SUM(C16:I16)	>=	=MinimumCash
17	=-LTRate*LTLoan	=-STRate*E16		=-E16	=J16+SUM(C17:I17)	>=	=MinimumCash
18	=-LTRate*LTLoan	=-STRate*E17		=-E17	=J17+SUM(C18:I18)	>=	=MinimumCash
19	=-LTRate*LTLoan	=-STRate*E18		=-E18	=J18+SUM(C19:I19)	>=	=MinimumCash
20	=-LTRate*LTLoan	=-STRate*E19		=-E19	=J19+SUM(C20:I20)	>=	=MinimumCash
21	=-LTRate*LTLoan	=-STRate*E20	=-LTLoan	=-E20	=J20+SUM(C21:I21)	>=	=MinimumCash

■ FIGURE 28.12

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	M	N	O	P	
1	<b>Everglade Cash Flow Management Problem When Applying Simulation</b>																
2																	
3		LT Rate	7%														
4		ST Rate	10%														
5																	
6		Start Balance	1														
7		Minimum Cash	0.5														
8																	
9		Cash Flow (Triangular Distribution)					Simulated						Balance				
10		Year	Minimum	Most Likely	Maximum	Cash Flow	LT Loan	LT Interest	ST Interest	LT Payback	ST Payback	Before ST Loan	ST Loan	Ending Balance		Minimum Balance	
11		2010	-9	-8	-7	-8.00	6.65					-0.35	0.85	0.50		>= 0.50	
12		2011	-4	-2	1	-1.67		-0.47	-0.09			-2.57	3.07	0.50		>= 0.50	
13		2012	-7	-4	0	-3.67		-0.47	-0.31			-7.01	7.51	0.50		>= 0.50	
14		2013	0	3	7	3.33		-0.47	-0.75			-4.89	5.39	0.50		>= 0.50	
15		2014	3	6	9	6.00		-0.47	-0.54			0.11	0.39	0.50		>= 0.50	
16		2015	1	3	5	3.00		-0.47	-0.04			2.60	0.00	2.60		>= 0.50	
17		2016	-6	-4	-2	-4.00		-0.47	0			-1.86	2.36	0.50		>= 0.50	
18		2017	4	7	12	7.67		-0.47	-0.24			5.10	0.00	5.10		>= 0.50	
19		2018	-5	-2	4	-1.00		-0.47	0			3.64	0.00	3.64		>= 0.50	
20		2019	5	10	18	11.00		-0.47	0			14.17	0.00	14.17		>= 0.50	
21		2020						-0.47	0	-6.65	0	7.05		7.05		>= 0.50	

	H	I	J	K	L	M	N	O	P
10	LT	ST	LT	ST	Before	ST	Ending		Minimum
11	Interest	Interest	Payback	Payback	ST Loan	Loan	Balance		Balance
12					=StartBalance+SUM(F12:K12)	=MAX(MinimumBalance-BalanceBeforeSTLoan,0)	=BalanceBeforeSTLoan+STLoan		>= -MinimumCash
13	=LTRate*LTLoan	=STRate*M12		=M12	=N12+SUM(F13:K13)	=MAX(MinimumBalance-BalanceBeforeSTLoan,0)	=BalanceBeforeSTLoan+STLoan		>= -MinimumCash
14	=LTRate*LTLoan	=STRate*M13		=M13	=N13+SUM(F14:K14)	=MAX(MinimumBalance-BalanceBeforeSTLoan,0)	=BalanceBeforeSTLoan+STLoan		>= -MinimumCash
15	=LTRate*LTLoan	=STRate*M14		=M14	=N14+SUM(F15:K15)	=MAX(MinimumBalance-BalanceBeforeSTLoan,0)	=BalanceBeforeSTLoan+STLoan		>= -MinimumCash
16	=LTRate*LTLoan	=STRate*M15		=M15	=N15+SUM(F16:K16)	=MAX(MinimumBalance-BalanceBeforeSTLoan,0)	=BalanceBeforeSTLoan+STLoan		>= -MinimumCash
17	=LTRate*LTLoan	=STRate*M16		=M16	=N16+SUM(F17:K17)	=MAX(MinimumBalance-BalanceBeforeSTLoan,0)	=BalanceBeforeSTLoan+STLoan		>= -MinimumCash
18	=LTRate*LTLoan	=STRate*M17		=M17	=N17+SUM(F18:K18)	=MAX(MinimumBalance-BalanceBeforeSTLoan,0)	=BalanceBeforeSTLoan+STLoan		>= -MinimumCash
19	=LTRate*LTLoan	=STRate*M18		=M18	=N18+SUM(F19:K19)	=MAX(MinimumBalance-BalanceBeforeSTLoan,0)	=BalanceBeforeSTLoan+STLoan		>= -MinimumCash
20	=LTRate*LTLoan	=STRate*M19		=M19	=N19+SUM(F20:K20)	=MAX(MinimumBalance-BalanceBeforeSTLoan,0)	=BalanceBeforeSTLoan+STLoan		>= -MinimumCash
21	=LTRate*LTLoan	=STRate*M20		=M20	=N20+SUM(F21:K21)	=MAX(MinimumBalance-BalanceBeforeSTLoan,0)	=BalanceBeforeSTLoan+STLoan		>= -MinimumCash
22	=LTRate*LTLoan	=STRate*M21	=-LTLoan	=M21	=N21+SUM(F22:K22)		=BalanceBeforeSTLoan+M22		>= -MinimumCash

Range Name	Cells
BalanceBeforeSTLoan	L12:L22
CashFlow	F12:F21
EndBalance	N22
EndingBalance	N12:N22
LTLoan	G12
LTRate	C3
MinimumBalance	P12:P22
MinimumCash	C7
StartBalance	C6
STLoan	M12:M22
STRate	C4

■ FIGURE 28.13

A spreadsheet model for applying simulation to the Everglade Golden Years Company cash flow management problem. The assumption cells are CashFlow (F12:F21), and the forecast cell is EndBalance (N22).

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. 28.13.

The target cell EndBalance (J21) in Fig. 28.12 becomes the forecast cell EndBalance (N22) in Fig. 28.13. On any trial of the simulation, if the simulated cash flows in CashFlow (F12:F21) in Fig. 28.13 are more favorable than the projected cash flows given in Table 28.1 (as is the case for the current numbers in Fig. 28.13), then EndBalance (N22) in Fig. 28.13 would be larger than EndBalance (J21) in Fig. 28.12. However, if the simulated cash flows are less favorable than the projections, then EndBalance (N22) in Fig. 28.13 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 2020 (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.

**Assumption cells:** CashFlow (F12:F21)

**Forecast cell:** EndBalance (N22)

### The Simulation Results

Figure 28.14 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 left-hand box in the frequency chart. The Certainty box in the lower middle box then indicates that nearly 89 percent of the trials resulted in a positive cash balance at the end. Furthermore, both the frequency chart and the cumulative chart show that many of these positive cash balances are reasonably large, with some exceeding \$10 million. The overall mean is \$6.72 million.

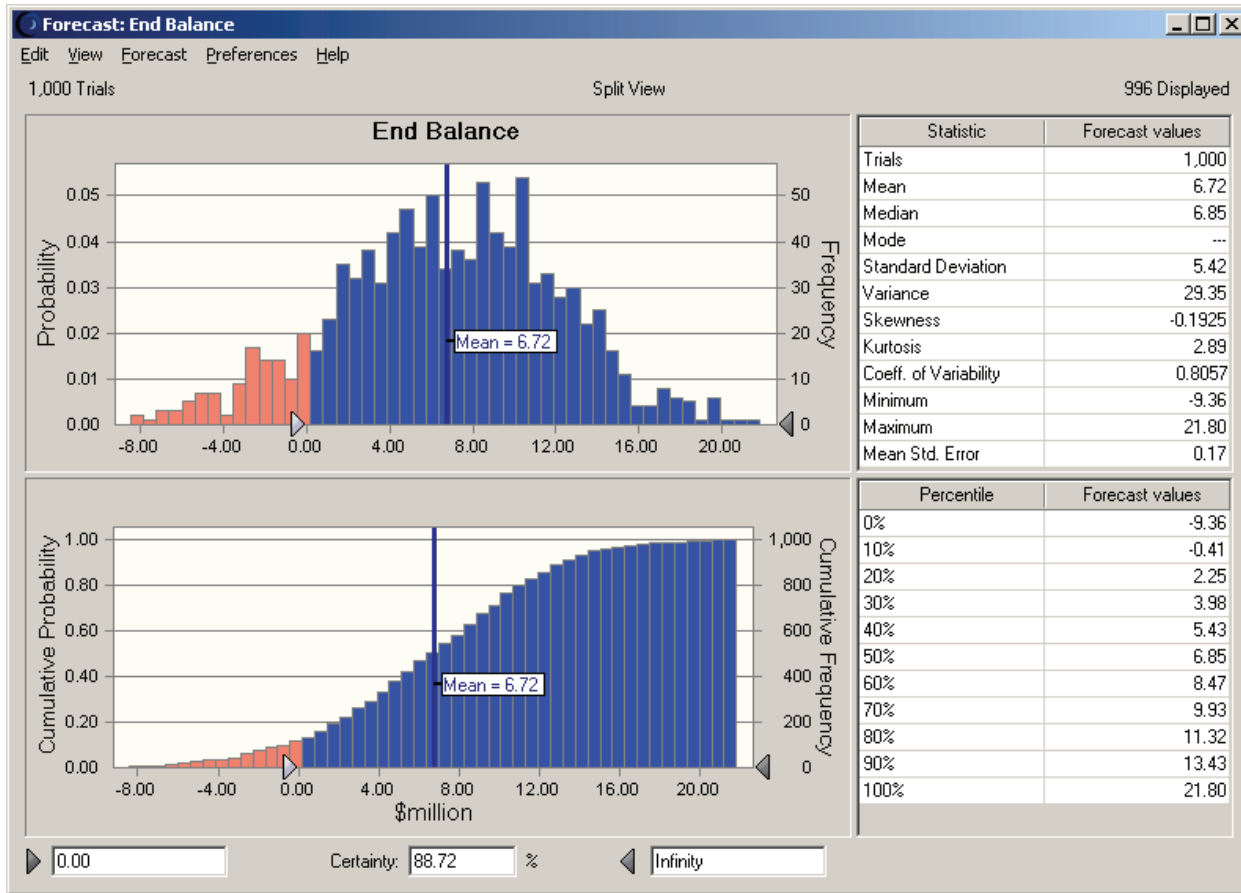
On the other hand, it is worrisome that more than 11 percent of the trials resulted in a negative cash balance at the end. Although huge losses were rare, most of these negative cash balances were quite significant, ranging from \$1 million to \$3 million.

### 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 11 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 late years if the cash flows are not as good as currently projected. This possibility is investigated in Problem 28.7.

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



■ **FIGURE 28.14**

The frequency chart and cumulative chart that summarize the results of running the simulation model in Fig. 28.13 for the Everglade Golden Years Company cash flow management problem. The Certainty box below the frequency chart reveals that nearly 89 percent of the trials resulted in a positive cash balance at the end of the 10 years.

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. 28.13. In particular, a 10-year loan of \$6.65 million will be taken immediately (the beginning of 2010). In addition, a 1-year loan will be taken at the beginning of each year from 2010 to 2019 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 a 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 \$18.1 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 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 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 the total NPV for each project.

### A Spreadsheet Model for Applying Simulation

A spreadsheet model has been formulated for this problem in Fig. 28.15. 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 assumption cells. (The numbers in these cells in Fig. 28.15 happen to be mean values that were entered merely to start the process of defining these assumption cells.) 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 28.16 shows the Normal Distribution dialogue box that is used to enter the parameters (mean and standard deviation) for the normal distribution into the first assumption cell D7 by referencing cells F7 and G7. The parameters for the other normal distributions are then copied and pasted into the corresponding assumption cells. The Uniform Distribution dialogue box (like the one displayed in Fig. 20.9 in Chap. 20) is used in a similar way to enter the parameters (minimum and maximum) for this kind of distribution into the assumption 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. 28.15) 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 forecast cell is NetPresentValue (D37).

■ **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)

28.4 FINANCIAL RISK ANALYSIS

	A	B	C	D	E	F	G	H
1	<b>Simulation of Think-Big Development Co. Problem</b>							
2								
3				<b>Project Simulated</b>				
4				<b>Cash Flow</b>				
5	<b>Hotel Project:</b>			<b>(\$millions)</b>				
6	Construction Costs:	Year 0		-80				
7		Year 1		-80	Normal	-80	5	(mean, st. dev.)
8		Year 2		-80	Normal	-80	10	(mean, st. dev.)
9		Year 3		-70	Normal	-70	15	(mean, st. dev.)
10	Revenue per Share	Year 4		30	Normal	30	20	(mean, st. dev.)
11		Year 5		40	Normal	40	20	(mean, st. dev.)
12		Year 6		50	Normal	50	20	(mean, st. dev.)
13	Selling Price per Share	Year 7		522	Uniform	200	844	(min, max)
14								
15	<b>Shopping Center Project</b>							
16	Construction Costs:	Year 0		-90				
17		Year 1		-50	Normal	-50	5	(mean, st. dev.)
18		Year 2		-20	Normal	-20	5	(mean, st. dev.)
19		Year 3		-60	Normal	-60	10	(mean, st. dev.)
20	Revenue per Share	Year 4		15	Normal	15	15	(mean, st. dev.)
21		Year 5		25	Normal	25	15	(mean, st. dev.)
22		Year 6		40	Normal	40	15	(mean, st. dev.)
23	Selling Price per Share	Year 7		387.5	Uniform	160	615	(min, max)
24								
25				<b>Think Big's</b>				
26				<b>Simulated Cash Flow</b>				
27				<b>(\$millions)</b>				<b>Share</b>
28		Year 0		-24.999			Hotel	16.50%
29		Year 1		-19.755			Shopping Center	13.11%
30		Year 2		-15.822				
31		Year 3		-19.416			Cost of Capital	10%
32		Year 4		6.917				
33		Year 5		9.878				
34		Year 6		13.494				
35		Year 7		136.931				
36								
37		Net Present Value (\$millions)		18.120				

	C	D
25		<b>Think Big's</b>
26		<b>Simulated Cash Flow</b>
27		<b>(\$millions)</b>
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)

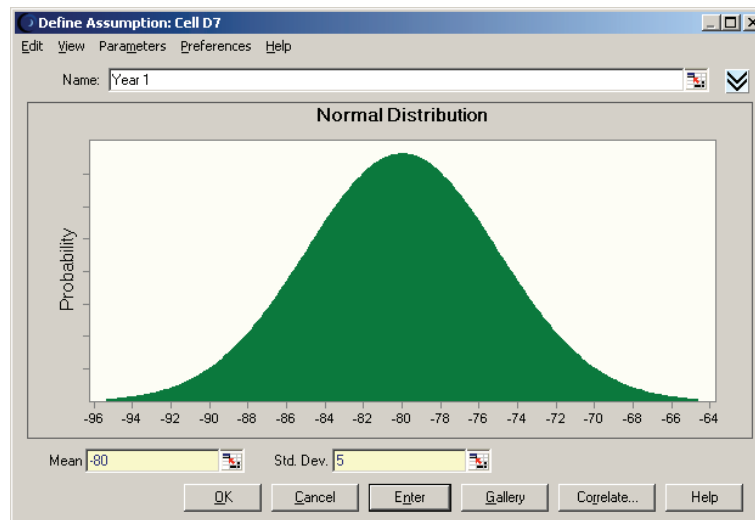
Range Name	Cells
CashFlowYear0	D28
CashFlowYear1To7	D29:D35
CostOfCapital	H31
HotelShare	H28
NetPresentValue	D37
ShoppingCenterShare	H29

■ FIGURE 28.15

A spreadsheet model for applying simulation to the Think-Big Development Co. financial risk analysis problem. The assumption cells are cells D7:D13 and D17:D23, the forecast cell is NetPresentValue (D37), and the decision variables are HotelShare (H28) and ShoppingCenterShare (H29).



■ **FIGURE 28.16**  
A normal distribution with parameters  $-80$  ( $=F7$ ) and  $5$  ( $=G7$ ) is being entered into the first assumption cell D7 in the spreadsheet model in Fig. 28.15.



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

**Assumption cells:** Cells D7:D13 and D17:D23  
**Decision variables:** HotelShare (H28) and ShoppingCenterShare (H29)  
**Forecast cell:** NetPresentValue (D37)

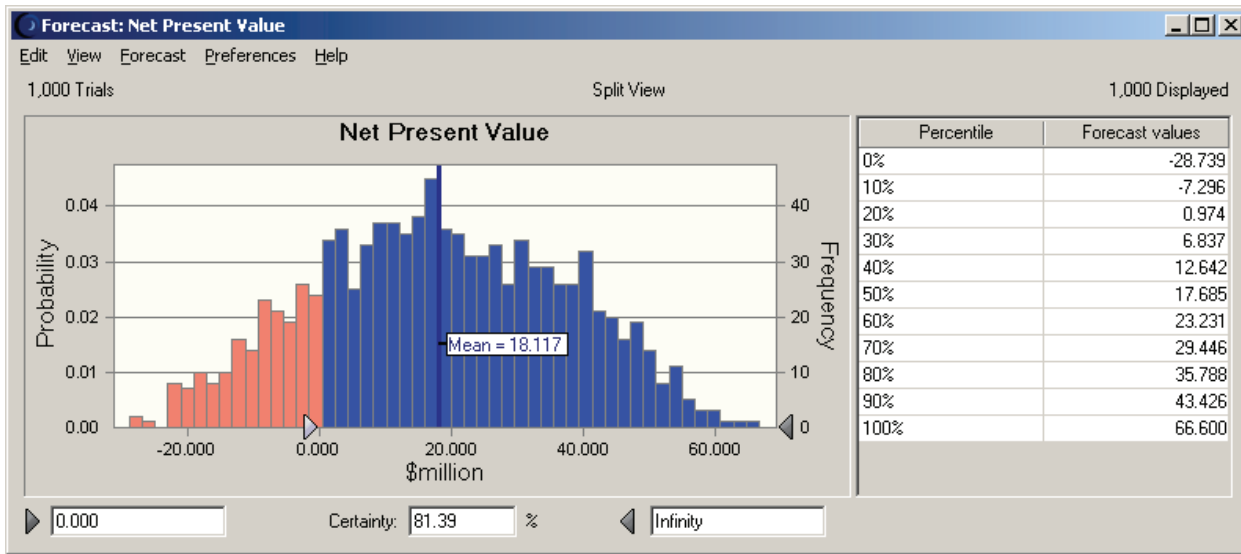
### The Simulation Results

Using the Run Preferences dialogue box to specify 1,000 trials, Fig. 28.17 shows the results of performing a simulation with the spreadsheet model in Fig. 28.15. The frequency chart in Fig. 28.17 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 \$18.117 million, which is very attractive. However, the 1,000 trials generated an extremely wide range of NPV values, all the way from about  $-\$29$  million to over  $\$66$  million. Thus, there is a significant chance of incurring a huge loss. By entering 0 into the box in the lower left-hand corner of the frequency chart, the Certainty box indicates that 81.39 percent of the trials resulted in a profit (a positive value of NPV). This also gives the bad news that there is nearly 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.

The percentiles table in Fig. 28.17 also provides management with some specific numbers for better assessing the risk. The 10 percent percentile of  $-7.296$  indicates a 10 percent chance of incurring a loss greater than about \$7.3 million. On the other hand, the 90 percent percentile of 43.426 indicates a 10 percent chance of achieving a huge profit (NPV) exceeding \$43.4 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 28.17**

The frequency chart and percentiles table that summarize the results of running the simulation model in Fig. 28.15 for the Think-Big Development Co. financial risk analysis problem. The Certainty box under the frequency chart reveals that 81.39 percent of the trials resulted in a positive value of the 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. As described in an article in the January–February 1992 issue of *Interfaces*, 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 given day.

For most of these flights, the number of requests for reservations considerably exceeds the number of seats that are 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 it roughly follows 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 based on 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 of its flights, since approximately 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 can be 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 has been a significant number of empty seats. While gathering its data, the OR group has discovered the reason why. An average of only 80 percent of the customers who make reservations for this particular 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 necessary 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. 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 (150). Thus, 190 appears to be a good first guess for an appropriate trade-off between having too many empty seats and having to bump too 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 each 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 28.18 shows a spreadsheet model for this problem. Because there are three measures of interest here, the spreadsheet model needs three forecast cells. These forecast cells are Profit (F23), NumberOfFilledSeats (C20), and NumberDeniedBoarding (C21). 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 assumption 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 dialogue box (shown earlier in Fig. 28.16) 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 assumption 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 cell F17 since this is the *average* percentage that have shown up for the flight in recent months.

However, the *actual* percentage that 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 assumption cell? Section 28.6 will describe the characteristics of various distributions. The one that has the characteristics to fit this assumption 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,

	A	B	C	D	E	F
1	<b>Transcontinental Airlines Overbooking</b>					
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	195	Normal	195	30
11		Demand (rounded)	195			
12						
13		Reservations to Accept	190			
14						
15					Tickets	Probability
16					Purchased	to Show up
17		Number that Show	152	Binomial	190	80%
18						
19						
20		Number of Filled Seats	150		Ticket Revenue	\$45,000
21		Number Denied Boarding	2		Bumping Cost	\$900
22					Fixed Cost	\$30,000
23					Profit	\$14,100

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

	B	C
20	Number of Filled Seats	=MIN(AvailableSeats,NumberThatShow)
21	Number Denied Boarding	=MAX(0,NumberThatShow-AvailableSeats)

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

Range Name	Cell
AvailableSeats	C4
AverageFare	C6
BumpingCost	F21
CostOfBumping	C7
Demand	C11
FixedCost	C5
NumberDeniedBoarding	C21
NumberOfFilledSeats	C20
NumberThatShow	C17
Profit	F23
ReservationsToAccept	C13
SimulatedTicketDemand	C10
TicketRevenue	F20
TicketsPurchased	E17

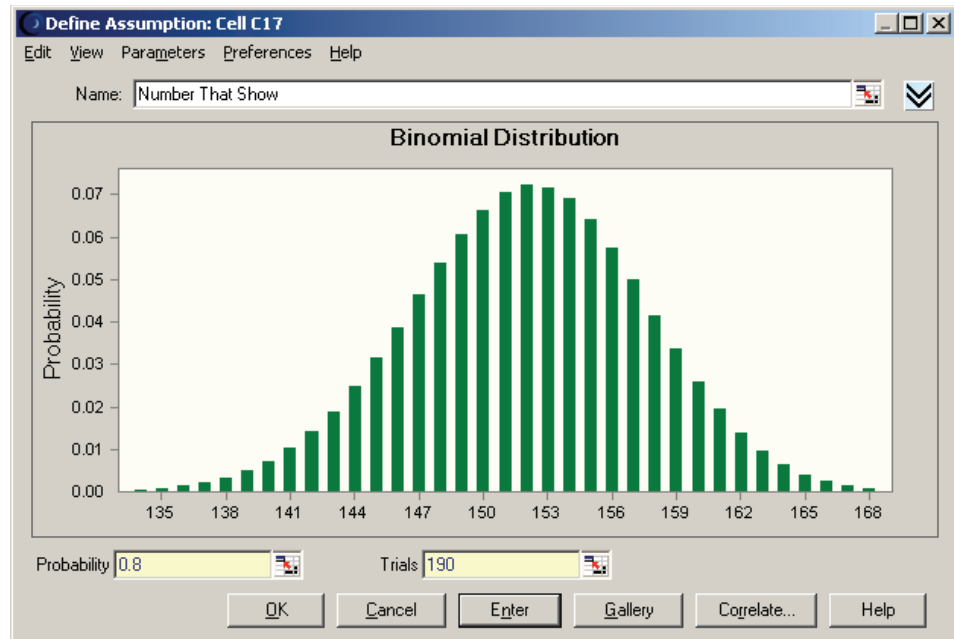
■ **FIGURE 28.18**

A spreadsheet model for applying simulation to the Transcontinental Airlines overbooking problem. The assumption cells are SimulatedTicketDemand (C10) and NumberThatShow (C17). The forecast cells are Profit (F23), NumberOfFilledSeats (C20), and NumberDeniedBoarding (C21). The decision variable is ReservationsToAccept (C13).

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 *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 28.19 displays the Binomial Distribution dialogue box that enters this distribution into NumberThatShow (C17) by referencing the parameters in cells F17 and E17.

■ **FIGURE 28.19**  
A binomial distribution with parameters 0.8 (=F17) and 190 (=E17) is being entered into the assumption cell NumberThatShow (C17).



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 random ticket demand. Crystal Ball therefore must determine the value for TicketsPurchased (E17) *before* it can randomly generate NumberThatShow (C17). Fortunately, Crystal Ball automatically takes care of the order in which to generate the various assumption cells so that this is not a problem.

The equations entered into all the output cells and forecast cells are given at the bottom of Fig. 28.18.

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

**Assumption cells:** SimulatedTicketDemand (C10) and  
NumberThatShow (C17)

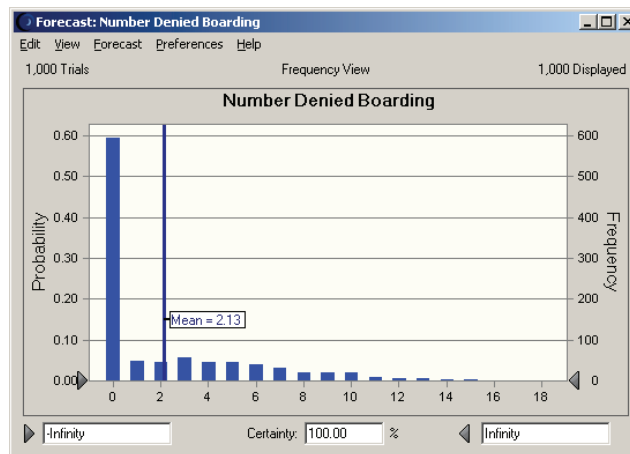
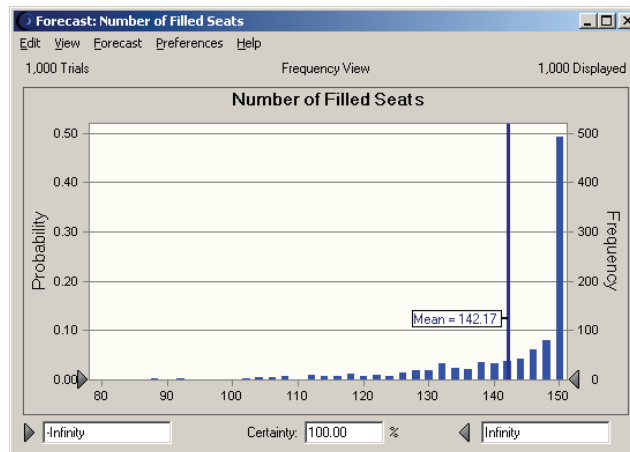
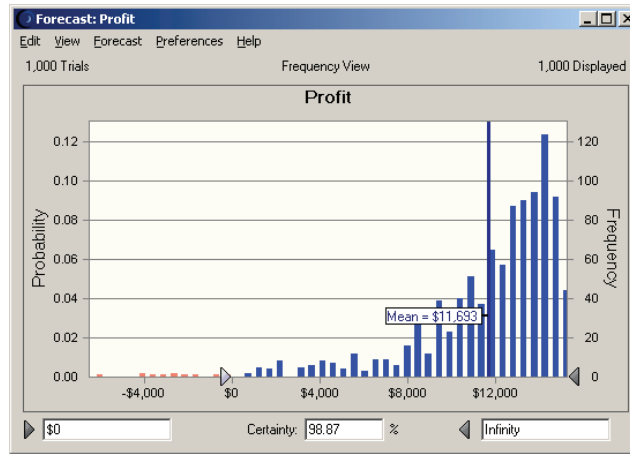
**Decision variable:** ReservationsToAccept (C13)

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

### The Simulation Results

Figure 28.20 shows the frequency chart obtained for each of the three forecast cells after applying simulation for 1,000 trials to the spreadsheet model in Fig. 28.18 with ReservationsToAccept (C13) set at 190.

The profit results estimate that the mean profit per flight would be \$11,693. 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-left box, the Certainty box reports that 98.87 percent of the trials resulted in a profit for that day's flight.



**FIGURE 28.20** The frequency charts that summarize the results for the respective forecast cells—Profit (F23), NumberOfFilledSeats (C20), and NumberDeniedBoarding (C21)—from running the simulation model in Fig. 28.18 for the Transcontinental Airlines overbooking problem. The Certainty box below the top frequency chart reveals that 98.87 percent of the trials resulted in a positive profit.

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 142.17 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 over 40 percent of the time. On nearly all of these trials, the number ranged between one and ten. Considering that no customers were denied boarding for 60 percent of the trials, the mean number is only 2.13.

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 ReservationsToAccept (C13) to pin down the optimal value of this decision variable. We will demonstrate how to do this efficiently with the help of a decision table in Sec. 28.7.

## 28.6 CHOOSING THE RIGHT DISTRIBUTION

As mentioned in Sec. 20.6, Crystal Ball's Distribution Gallery provides a wealth of options for choosing a probability distribution to use. Any of 21 distributions can be selected as the one to be entered into any assumption cell. In the preceding sections, we have illustrated the use of five of these distributions (the discrete uniform, triangular, normal, and binomial distributions). However, not much was actually 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 21 distributions and how these characteristics help to identify the best choice. We next describe a special feature of Crystal Ball that is available for creating a custom distribution when none of the other 20 choices in the Distribution Gallery are acceptable. We then return to the example analyzed in Sec. 20.6 to illustrate the use of another special feature of Crystal Ball. When historical data are available, this feature has the capability to identify which of the available continuous distributions provides the best fit to the 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.

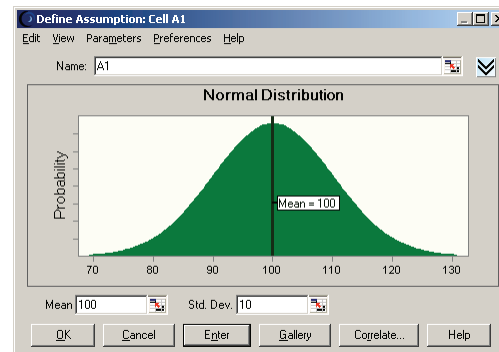
### Characteristics of the Available Distributions

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. 28.18.) Crystal Ball's Distribution Gallery includes both continuous and discrete distributions. We will begin by looking at the continuous distributions.

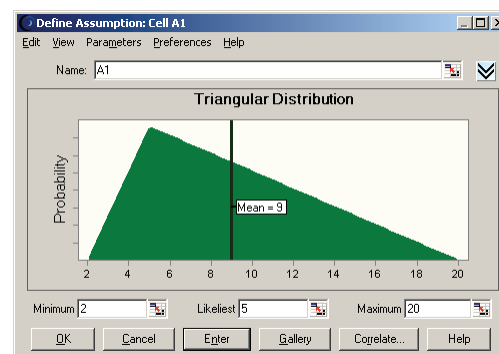
The right-hand side of Fig. 28.21 shows the dialogue box for three commonly used continuous distributions from the Distribution Gallery. The dark figure in each dialogue box displays a typical *probability density function* for that distribution. The height of the

**Normal Distribution:**

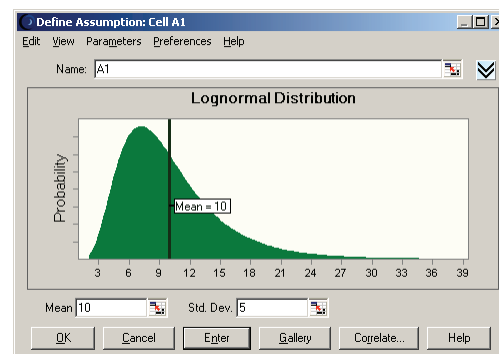
- Some value most likely (the mean)
- Values close to mean more likely
- Symmetric (as likely above as below mean)
- Extreme values are possible, but rare

**Triangular Distribution:**

- Some value most likely
- Values close to most-likely 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

**FIGURE 28.21**

The characteristics and dialogue boxes for three popular central-tendency distributions in Crystal Ball's Distribution Gallery: (1) the normal distribution, (2) the triangular distribution, and (3) the lognormal distribution.

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 to as *central-tendency distributions*. The characteristics of each of these distributions are listed on the left-hand side of Fig. 28.21.



**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 table for this distribution.) One reason that it arises so frequently is that, according to the *central limit theorem*, the sum of many independent and identically distributed random variables has 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 assumption cell in Fig. 28.18 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 ( $3 \times 30 = 90$ ), so a negative value essentially could never occur. (When normal distributions were entered into assumption cells in Fig. 28.15 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. 28.21 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 (the minimum value, most likely value, and maximum value) are intuitive ones, so they are relatively easy to estimate.

These advantages have made the triangular distribution a popular choice to use for simulations. They are the reason why this distribution was used in previous examples to represent competitors' bids for a construction contract (in Fig. 28.1), activity times (in Fig. 28.7), and cash flows (in Fig. 28.13).

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 a fixed minimum value and maximum value. 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. 28.21 combines some of the advantages of both the normal and triangular distributions. It has a curve with a gradually changing slope. It has a very long tail to the right, even more so than for the normal distribution, so it 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 so the normal distribution should not be used.

This distribution always is “positively skewed,” meaning that the long tail is always to the right. This forces the most likely value to be toward the left side (with the mean being 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 Discrete Uniform Distributions.** Although the preceding three distributions are all central-tendency distributions, the uniform distributions shown in Fig. 28.22 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 another. 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 discrete uniform distribution. If only integer values are possible, then the discrete uniform distribution would be the preferable one.

Either of these distributions is a particularly convenient one because it has only two parameters (the minimum value and the maximum value) and both are very easy to interpret, so estimating the parameters is relatively straightforward. These distributions receive considerable use for this reason. In our examples of performing simulations on a spreadsheet, the discrete uniform distribution has been 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. 28.1) and the future sale price for real estate property (in Fig. 28.15).

The disadvantage of either distribution is that it usually is only a rough approximation of the true distribution. It is uncommon for either the minimum value or maximum value to be just as likely as any other value between these bounds while any value barely outside these bounds is impossible. Nevertheless, one of these distributions frequently is chosen because of its ease of use when none of the other standard distributions provides a clearly better approximation of the true distribution.

**The Weibull and Beta Distributions.** We will describe the Weibull and beta distributions together because, as suggested by their similar shapes in Fig. 28.23, these two distributions have similar characteristics. In contrast to the uniform distribution, both are central-tendency distributions. In contrast to the normal, lognormal, and uniform distributions, both have more than two parameters. The three (for the Weibull) or four (for the beta) parameters provide great flexibility in adjusting the shape of the curve to fit the situation. This enables making the distributions positively skewed, symmetric, or negatively skewed as desired. This flexibility is the key advantage of these distributions. Another advantage is that they have a minimum value, so negative values can be avoided. The location parameter sets the minimum value for the Weibull distribution. The Weibull distribution allows rare extreme values to the right, whereas two of the parameters for the beta distribution set the minimum value and the maximum value.

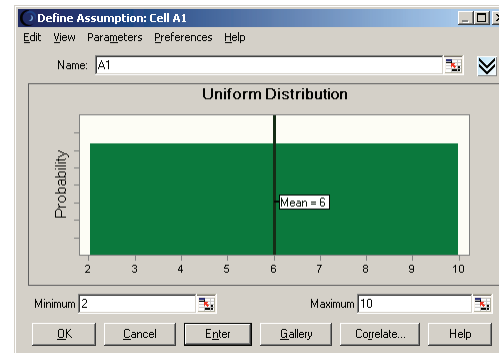
Certain parameters, such as the location parameter mentioned above for the Weibull distribution or the minimum parameter and maximum parameter for the beta distribution, have an intuitive interpretation. The scale parameter for the Weibull distribution simply sets the width of the distribution (excluding rare extreme values). However, neither the shape parameter for the Weibull distribution nor the alpha and beta parameters for the beta distribution are particularly intuitive. Therefore, these distributions are mainly used only when historical data are available for calculating estimates of these parameters. (One

■ **FIGURE 28.22**

The characteristics and dialogue boxes for the uniform and discrete uniform distributions in Crystal Ball's Distribution Gallery.

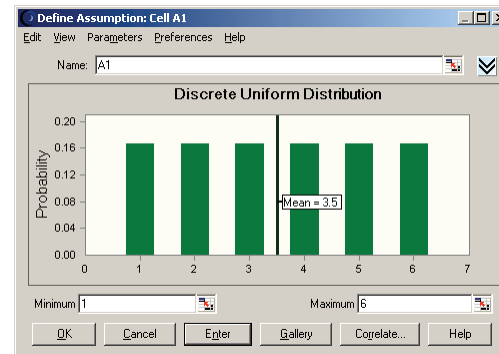
**Uniform Distribution:**

- Fixed minimum and maximum value
- All values equally likely



**Discrete Uniform Distribution:**

- Fixed minimum and maximum value
- All integer values equally likely

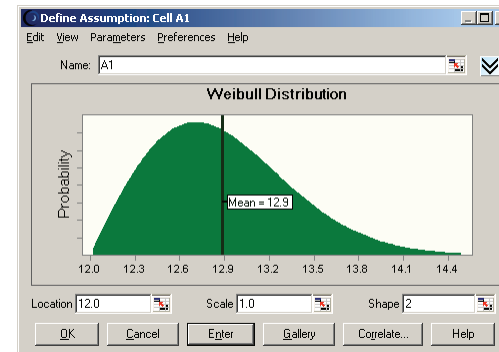


■ **FIGURE 28.23**

The characteristics and dialogue boxes for three- and four-parameter distributions in Crystal Ball's Distribution Gallery: (1) the Weibull distribution and (2) the beta distribution.

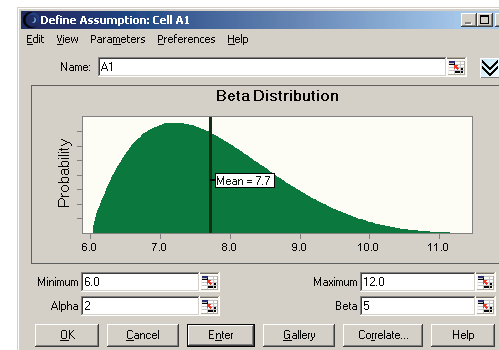
**Weibull Distribution:**

- Random value above some number (Location)
- Shape > 0 (usually  $\leq 10$ )
- Shape < 3 becomes more positively skewed (below mean more likely) until it resembles exponential distribution (equivalent at Shape = 1)
- Symmetrical at Shape = 3.25, becomes more negatively skewed above that
- Scale defines width



**Beta Distribution:**

- Random value between 0 and some positive number (scale)
- Shape specified using two positive values (alpha, beta)
- Alpha < Beta: positively skewed (below mean more likely)
- Beta < Alpha: negatively skewed



exception is the innovative use of the beta distribution by the PERT three-estimate approach described in Sec. 22.4, which uses the minimum, most likely, and maximum values to calculate approximate values of the mean and standard deviation.) With historical data available, the choice between these two distributions (and other options as well) can then be based on which one provides the best fit with the data after obtaining statistical estimates of its parameters. We will describe a little later how Crystal Ball can make the choice in this way.

**The Exponential Distribution.** If you have studied Chap. 17 on queueing theory, you will recall that the most commonly used queueing models (those based on the birth-and-death process) assume that interarrival times and service times both have exponential distributions. This assumption for interarrival times is a valid one for the common situation where customer arrivals are random events. Similarly, this assumption for service times implies that service completions by a busy server are random events. Section 17.4 describes in detail the various properties of the exponential distribution and how these properties fit the characteristics of many queueing systems.

As first depicted in Fig. 17.3, this distribution has the unusual shape shown in the top half of Fig. 28.24. 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.

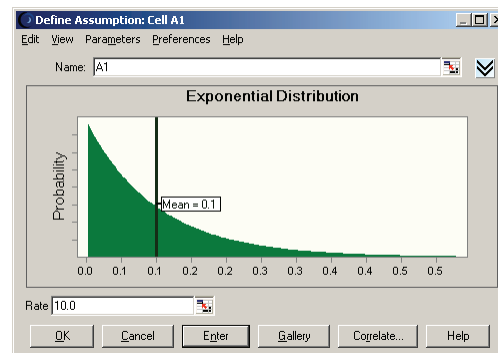
Since the only parameter is the rate at which the random events occur on the average, this distribution is a relatively easy one to use.

#### ■ FIGURE 28.24

The characteristics and dialogue boxes for two distributions that involve random events. These distributions in Crystal Ball's Distribution Gallery are (1) the exponential distribution and (2) the Poisson distribution.

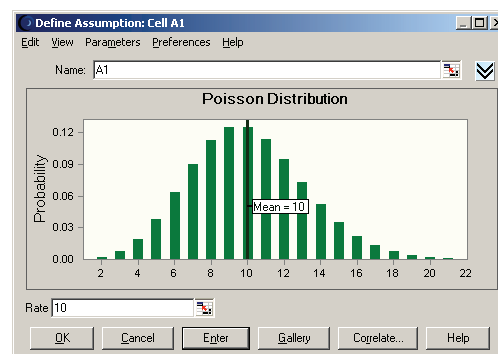
#### Exponential Distribution:

- Widely used to describe time between random events (e.g. time between arrivals)
- Events are independent
- Rate = avg. number of events per unit time (e.g., arrivals per hour)



#### 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
- Rate = avg. # of events during period of time (e.g., arrivals per hour), assumed constant over time



**The Poisson Distribution.** Although the exponential distribution (like all the preceding ones) is a continuous distribution, the Poisson distribution is a discrete distribution, as shown in the bottom half of Fig. 28.24. The only possible values are nonnegative integers: 0, 1, 2,  $\dots$ . 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 “Rate” to be entered into the one parameter field in the dialogue box should be the expected number of events that occur within that period of time.

**The Yes-No and Binomial Distributions.** The yes-or-no distribution is a very simple discrete distribution with only two possible values (0 or 1) as shown in the top half of Fig. 28.25. It is used to simulate whether a particular event occurs or not. The only parameter of the 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. 28.25, the binomial distribution is an extension of the yes-no 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 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).

You have seen another example in the preceding section when the binomial distribution was entered into the assumption cell NumberThatShow (C17) in Fig. 28.18. 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. 28.26 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 Shape field in its dialogue box, the *negative binomial distribution* gives the distribution of the number of trials until the event occurs the number of times specified in the Shape field. Thus, shape is a parameter for this distribution, and the fixed probability of the event occurring on a trial is a parameter for both distributions.

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 Shape = 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 a probability of 0.5 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 Shape = 5) gives the distribution of the production run size that would be needed to fulfill this order.

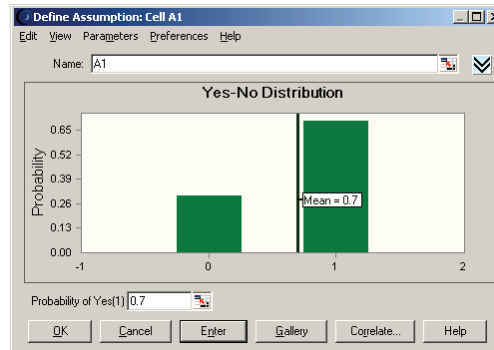
■ **FIGURE 28.25**

The characteristics and dialogue box for the yes-no and binomial distributions in Crystal Ball's Distribution Gallery.

## Distributions for Number of Times an Event Occurs

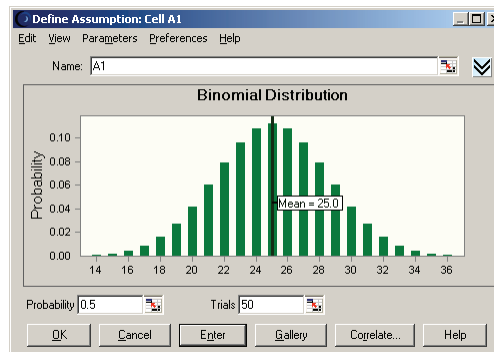
### Yes-No Distribution:

- 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



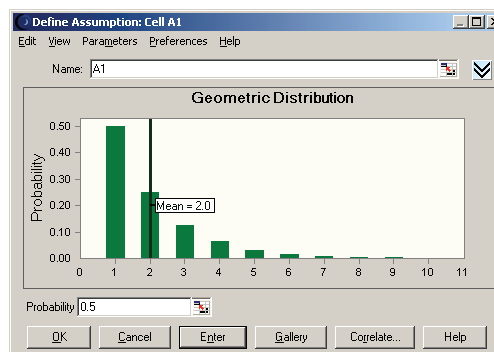
■ **FIGURE 28.26**

The characteristics and dialogue boxes for two distributions that involve the number of trials until events occur. These distributions in Crystal Ball's Distribution Gallery are (1) the geometric distribution and (2) the negative binomial distribution.

## Distributions for Number of Trials Until Events Occur

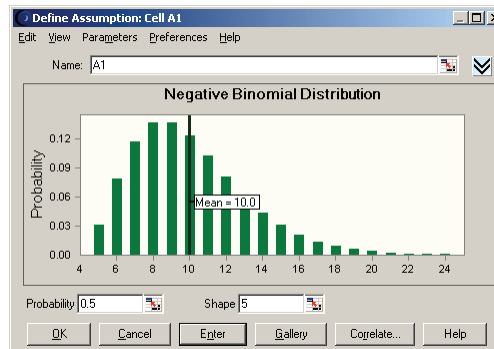
### Geometric Distribution:

- 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 Shape =  $n = 1$
- Probability same for each trial
- Continue until  $n^{\text{th}}$  success
- Number of trials unlimited





**Other Distributions.** The Distribution Gallery includes eight additional distributions: the custom, hypergeometric, gamma, logistic, Pareto, student's t, max extreme, and min extreme distributions.<sup>1</sup>

The custom distribution is an especially useful one because it enables you to design your own distribution when none of the other 20 distributions in the Distribution Gallery are acceptable. The next subsection will focus on the custom distribution.

The remaining seven distributions are not widely used in simulations, so they will not be discussed further.

**The Custom Distribution.** Of the 21 probability distributions included in the Distribution Gallery, 20 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 assumption cell. However, unique circumstances occasionally arise where none of the standard distributions fit the situation. This is where the 21th member of the Distribution Gallery, called the "custom distribution," enters the picture.

The custom distribution actually is not a probability distribution until you make it one. Rather, choosing this member of the Distribution Gallery triggers a process that enables you to custom-design your own probability distribution to fit almost any unique situation you might encounter. In designing your distribution, you are given five choices under the Parameters menu: unweighted values, weighted values, continuous ranges, discrete ranges, and sloping ranges. The choice determines how many parameters are available. For unweighted values, there is just one parameter (*Value*). A list of values are entered and each of these values are assumed to be equally likely. For weighted values, there are two parameters (*Value* and *Probability*) allowing each value to have a different probability. Continuous ranges have three parameters (*Minimum*, *Maximum*, and *Probability*) to specify the range (or ranges) of the distribution and the total probability of each range. Discrete ranges add a fourth parameter (*Step*) to specify the distance between the possible discrete values in the range. Finally, sloping ranges have two parameters (*Height of Minimum* and *Height of Maximum*) to allow the distribution to slope from the minimum to the maximum. Furthermore, these values or ranges can be combined in any possible way, such as combining a set of discrete values with a continuous range. We will show two examples.

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 weighted values under the Parameters menu in the Custom Distribution dialogue box. Each discrete value and probability (expressed as a decimal number) is then entered in the Value and Probability columns as shown in Fig. 28.27.

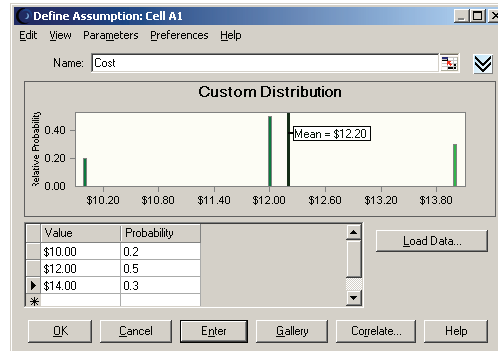
The other distribution types (unweighted values, continuous ranges, discrete ranges, and sloping ranges) are entered in a similar fashion by choosing the appropriate option under the Parameter menu. We next illustrate how to enter a combination of a weighted value and a continuous range.

<sup>1</sup>The gamma distribution has been introduced previously when the Erlang distribution (a special type of gamma distribution) was described in Sec. 17.7.

## Custom Distribution (Weighted Values)

### Weighted Values (Custom Distribution):

- Enter set of values with varying probabilities
- For each discrete value, enter *Value* and *Probability*



■ **FIGURE 28.27**

Following the instructions on the left, this dialogue box illustrates how Crystal Ball's custom distribution can enable you to custom-design your own distribution to enter a set of discrete values and their probabilities.



The example in this case is a variation of the probability distribution of the revenue from drilling for oil in the Goferbroke Company example introduced in Sec. 15.1. In that example, there was a probability of 0.75 of not finding oil (so a revenue of 0) and a probability of 0.25 of finding oil, in which case the revenue would be \$800,000. However, there actually is some uncertainty about how much revenue would be received if oil is found, so the following distribution would be more realistic.

Revenue of 0 has probability 0.75.

A range of revenue values: \$600,000 to \$1,000,000.

The values over this range are equally likely.

Probability of a value in this range: 0.25.

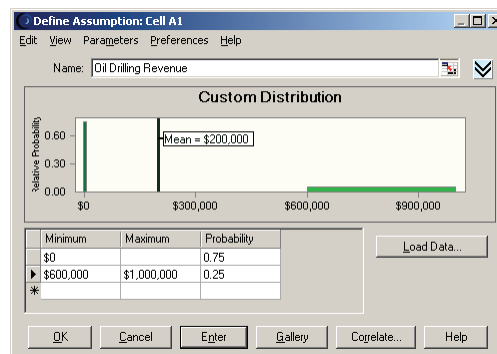
This distribution combines a single discrete value at 0 with a continuous portion between \$600,000 and \$1,000,000. To enter such a combination distribution, choose the member of the combination under the Parameters menu with the most parameters needed. The weighted value requires two parameters (*Value* and *Probability*) while the continuous range requires three (*Minimum*, *Maximum*, and *Probability*), so choose Continuous Range under the Parameters menu. This leads to the Custom Distribution dialogue box with three columns for data (*Minimum*, *Maximum*, and *Probability*), as shown in Fig. 28.28.

For a discrete value, only two parameters are needed (*Value* = \$0 and *Probability* = 0.75), so the middle column is left blank in the first row of the table near the bottom of Fig. 28.28. For the continuous range, we enter the *Minimum* (\$600,000), *Maximum* (\$1,000,000), and *Probability* (0.25) in the second row of this table in Fig. 28.28.

## Combinations with the Custom Distribution

### Combination (Custom Distribution):

- For discrete values, enter *Value* and *Probability*.
- For continuous range, enter *Minimum*, *Maximum*, and *Probability*



■ FIGURE 28.28

Following the instructions on the left, this dialogue box for the custom distribution demonstrates combining two of its options: (1) enter a discrete value at \$0 with a probability of 0.75 and (2) enter a uniform distribution between \$600,000 and \$1,000,000 and then assign a total probability of 0.25 to this portion of the overall distribution.

### Identifying the Continuous Distribution That Best Fits Historical Data

We now have at least mentioned all 21 types of probability distributions in the Distribution Gallery 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 assumption cell. When historical data are available, Crystal Ball provides a powerful feature for doing this with the continuous distributions by using the Fit button on the Distribution Gallery dialogue box. We illustrate this feature below 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 shown in Sec. 20.6. This model includes an assumption cell SimulatedDemand (C12). To get started, a discrete uniform distribution between 40 and 70 has been entered into this assumption 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 28.29 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 show a lot of variation in sales from day to day over a range from about 40 to 70 copies, so the underlying distribution can be closely approximated by a continuous distribution. However, it is difficult to tell from these numbers which continuous distribution in the Distribution Gallery best fits these data.

Crystal Ball provides the following procedure for addressing this issue.

#### Procedure for Fitting the Best Continuous Distribution to Data

1. Gather the data needed to identify the best distribution to enter into an assumption cell.
2. Enter the data into the spreadsheet containing your simulation model.
3. Select the cell that you want to define as an assumption cell that contains the distribution that best fits the data.
4. Choose Define Assumption from the Crystal Ball tab or toolbar, which brings up the Distribution Gallery dialogue box.
5. Click the Fit button on this dialogue box, which brings up the Fit Distribution dialogue box.
6. Use the Range box in this dialogue box to enter the range of the historical data in your worksheet.



10. If desired, the Next button can be clicked repeatedly for identifying 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 you want to use, click the Accept button while that distribution is showing. This will enter the appropriate parameters into the dialogue box for this distribution. Clicking OK then enters this distribution into the assumption cell.

Since Fig. 28.29 already includes the needed data in cells F4:F63, applying this procedure to Freddie's problem begins by selecting SimulatedDemand (C12) as the cell we want to define as an assumption cell that contains the distribution that best fits the data. Then applying steps 4 and 5 brings up the first Fit Distribution dialogue box displayed in Fig. 28.30. The range F4:F63 of the data in

■ **FIGURE 28.30**

The Fit Distribution dialogue box specifies the range of the data in Fig. 28.29 for Freddie's problem, which continuous distributions will be considered (all) and which ranking method will be used (the chi-square test) to evaluate how well each of the distributions fit the data.

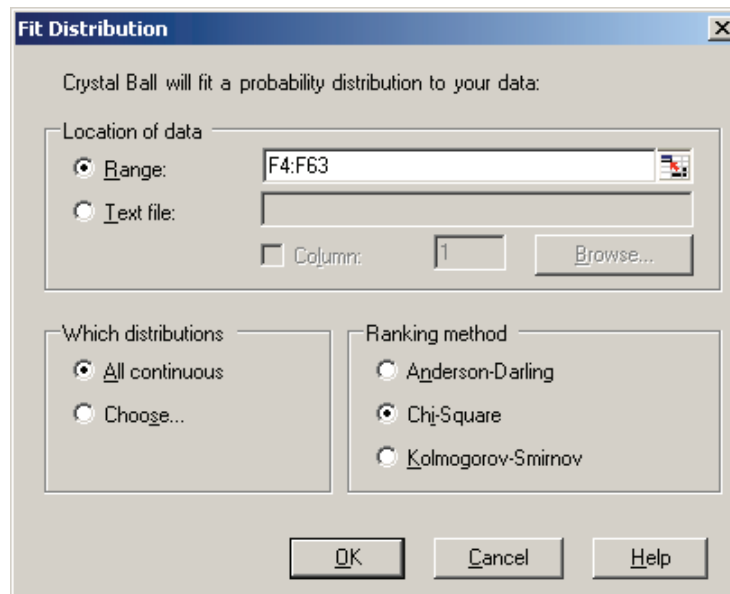
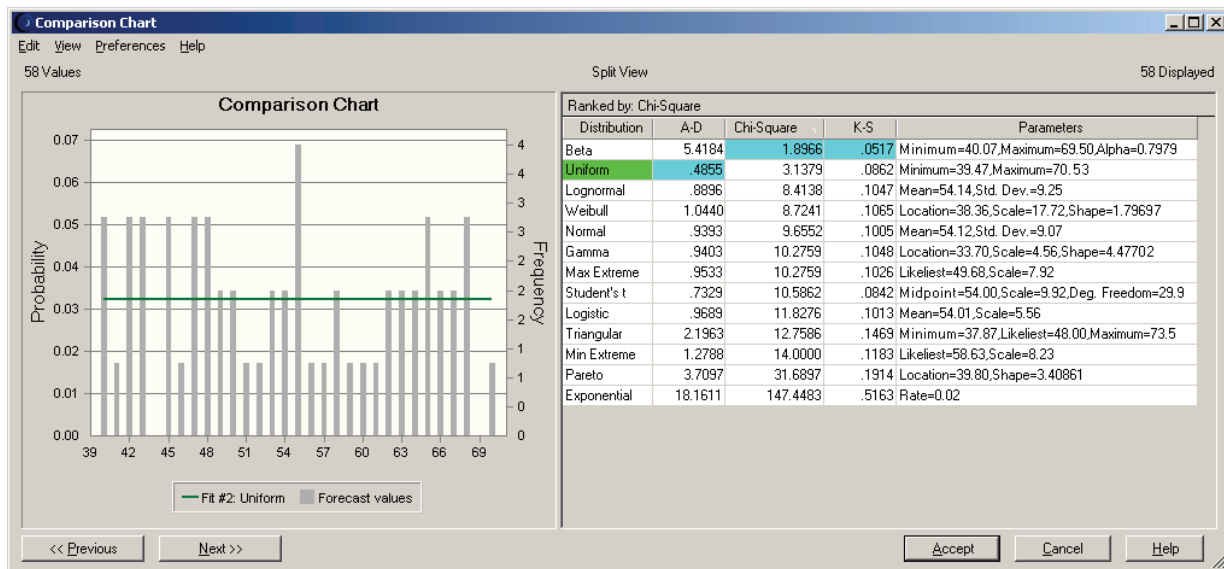


Fig. 28.29 is entered into the Range box of this dialogue box. When deciding which continuous distributions should be considered for fitting, the default option of considering all the continuous distributions in the Distribution Gallery has been selected here. The default choice of the Chi-Square Test also has been selected in Test also has been selected for the ranking method. Clicking OK then brings up the comparison chart in Fig. 28.31. The vertical lines show the frequency chart for the data in Fig. 28.29. The right side of the comparison chart identifies the best-fitting distribution according to the Anderson-Darling, Chi-Square, and Kolmogorov-Smirnov tests. Lower numbers in these columns indicate a better fit. The Anderson-Darling test indicates that the uniform distribution is the best fit, while the Chi-Square and Kolmogorov-Smirnov test indicate that the beta distribution is a slightly better fit.

This procedure compares the continuous distributions. Of these distributions, the continuous uniform distribution is either the best fit or nearly the best fit to the table. In combination with the fact that demand must be integer, this confirms that the choice made in Freddie's original spreadsheet model in Fig. 20.7 to enter the discrete uniform distribution into the assumption cell Demand (C12) was reasonable.

■ FIGURE 28.31

On the right, this comparison chart identifies the continuous distribution that provides the best fit with the historical demand data in Fig. 28.29. This distribution then is plotted (the horizontal line) so that it can be compared with the frequency distribution of the historical demand data.



## 28.7 DECISION MAKING WITH DECISION TABLES

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. 28.1.

Overbooking example: ReservationsToAccept (C13) in Fig. 28.18.

In both of these cases, you have seen how well simulation with Crystal Ball can *evaluate* a particular value of the decision variable by providing a variety of enlightening statistical output for the forecast cell(s). However, in contrast to many OR techniques, this approach has not identified an *optimal solution* for the decision variable.

Near the end of Sec. 20.6, we described a special feature of Crystal Ball called the *Decision Table tool* that enables using a series of simulation runs to identify at least an approximation of an optimal solution for problems with only one or two decision variables. We also illustrated the use of this tool by applying it to the relatively straightforward example involving Freddie the newsboy that was analyzed throughout Sec. 20.6.

This tool will enable us to further analyze the two examples cited above.

### Application of the Decision Table Tool to the Reliable Construction Co. Bidding Problem

We turn now to the application of the Decision Table tool to the Reliable Construction Co. bidding problem presented in Sec. 28.1. Since the procedure for how to generate a Decision Table 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 the project of constructing a new plant for a major manufacturer. Therefore, the decision variable in the spreadsheet model in Fig. 28.1 is OurBid (C25). The Define Decision Variable dialogue box in Fig. 28.32 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 Variable Bounds in this dialogue box. Because the bid

■ **FIGURE 28.32**

This dialogue box specifies the characteristics of the decision variable OurBid (C25) in Fig. 28.1 for the Reliable Construction Co. contract bidding problem.

The image shows a dialog box titled "Define Decision Variable: Cell C25". It has a "Name" field containing "Our Bid". Below this is a "Bounds" section with "Lower" set to 4.8 and "Upper" set to 5.8. To the right is a "Type" section with two radio buttons: "Continuous" (which is selected) and "Discrete". Below the "Discrete" radio button is a "Step" field which is empty. At the bottom of the dialog box are three buttons: "OK", "Cancel", and "Help".

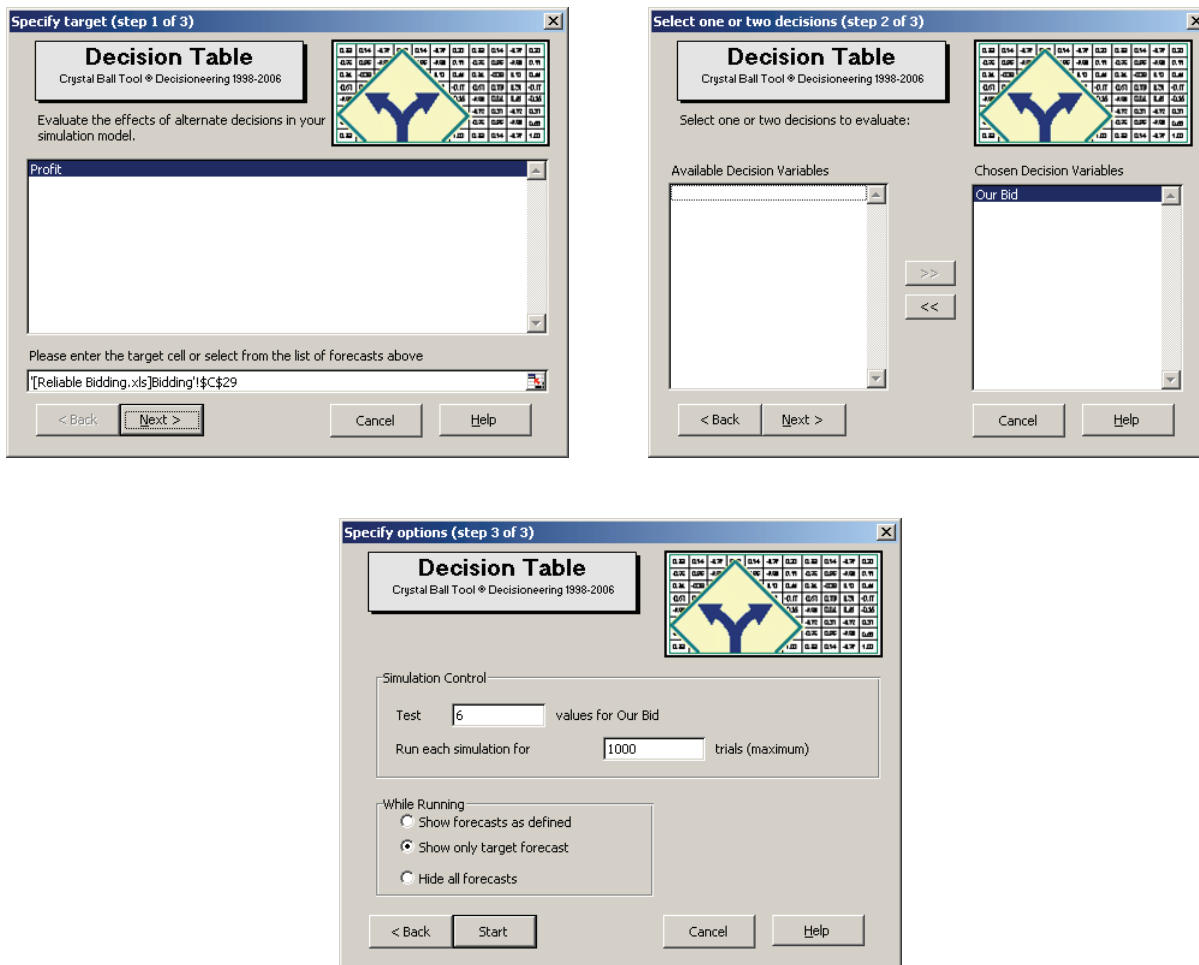
can be *any* amount (down to the cent) within this range, the variable has been classified as a continuous variable on the right side of the dialogue box.

Management wants to choose the bid that would maximize its expected profit. Consequently, the forecast cell in the spreadsheet model is Profit (C29). In the Decision Table dialogue box for Step 1 shown in Fig. 28.33, this forecast cell has been selected to be the target cell. In the Step 2 dialogue box, the decision variable OurBid (C25) already has been brought over to the right-hand side as the one to be considered. The decision has been made in the Step 3 dialogue box to test six values for this decision variable. The six values are then automatically distributed evenly over the range specified in Fig. 28.32, so that simulations will be run for bids of 4.8, 5.0, 5.2, 5.4, 5.6, and 5.8 (in millions of dollars). The Step 3 dialogue box also has been used to specify that each simulation will run for 1,000 trials.

Figure 28.34 shows the resulting Decision Table. 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

■ FIGURE 28.33

The three Decision Table dialogue boxes for the Reliable Construction Co. contract bidding problem.



	A	B	C	D	E	F	G
	Trend Chart	OurBid (4,800)	OurBid (5,000)	OurBid (5,200)	OurBid (5,400)	OurBid (5,600)	OurBid (5,800)
	Overlay Chart						
1	Forecast Charts						
2		0.188	0.357	0.476	0.501	0.252	0.010
3		1	2	3	4	5	6

■ **FIGURE 28.34**  
The Decision Table for the Reliable Construction Co. contract bidding problem described in Sec. 28.1.

of \$501,000 in cell E4 should be a close estimate of the expected profit from using this bid. The prototype example considered in Chap. 22 begins with the company having just won the contract by submitting this bid.

Problem 28.6 asks you to refine this analysis by generating a Decision Table that considers all bids between \$5.2 million and \$5.6 million in multiples of \$0.05 million. (It also asks you to use the OptQuest technique presented in Supplement 3 to Chap. 20 to search for the precise optimal bid.)

**Application of the Decision Table Tool to 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.

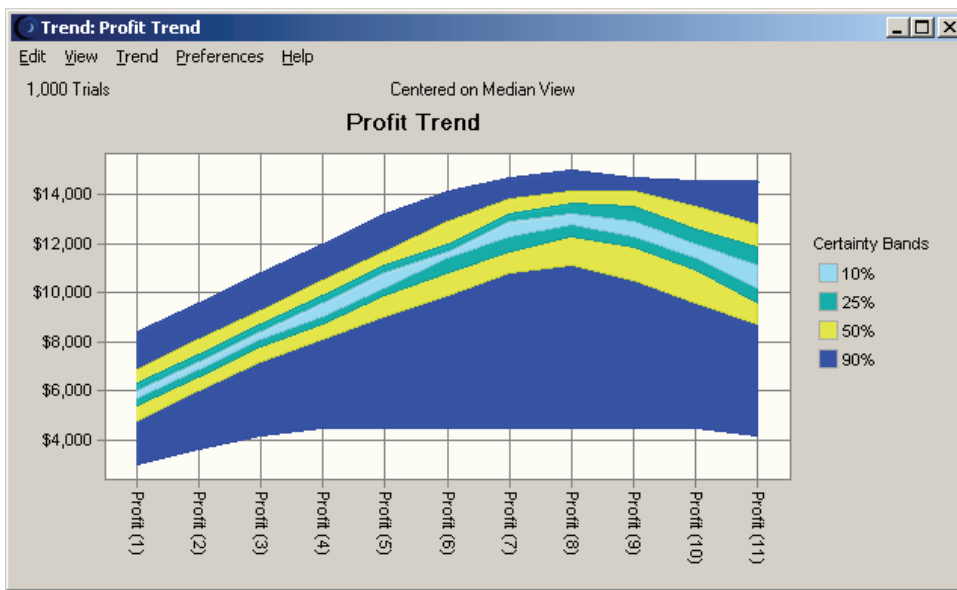
In the spreadsheet model in Fig. 28.18, the decision variable is ReservationsToAccept (C13), and the forecast cell is Profit (F23). The OR group wants to consider integer values of the decision variable over the range between 150 and 200, so the Define Decision Variable dialogue box is used in the usual way to specify these bounds on the variable and to define the variable as being discrete. The three Decision Table dialogue boxes also are used in the usual way. In the Step 3 dialogue box, 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 number of trials for each simulation run also is set at 1,000 in this dialogue box.

The results are shown in Fig. 28.35. The Decision Table in the upper part 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 \$11,926 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, Prob. 28.8 asks you to continue the analysis by generating a Decision Table that considers all integer values over this range, as well as by applying OptQuest to search for the optimal number of reservations to accept.)



28.7 DECISION MAKING WITH DECISION TABLES

	A	B	C	D	E	F	G	H	I	J	K	L
	Trend Chart	Reservations To Accept (150)	Reservations To Accept (155)	Reservations To Accept (160)	Reservations To Accept (165)	Reservations To Accept (170)	Reservations To Accept (175)	Reservations To Accept (180)	Reservations To Accept (185)	Reservations To Accept (190)	Reservations To Accept (195)	Reservations To Accept (200)
1	Overlay Chart											
2	Forecast Charts	\$5,792	\$6,898	\$7,973	\$9,006	\$9,984	\$10,889	\$11,613	\$11,926	\$11,693	\$11,063	\$10,303
3		1	2	3	4	5	6	7	8	9	10	11



■ **FIGURE 28.35** The Decision Table and trend chart for the Transcontinental Airlines overbooking problem described in Sec. 28.5.

The trend chart in the lower part of Fig. 28.35 provides additional insight for making a decision on the number of reservations to accept. All seven bands in this chart trend upward until the number of reservations to accept reaches approximately 185 and then starts trending slowly downward. This indicates that the *entire* frequency distribution from the respective simulation runs keeps shifting upward until the eighth run (the one for 185 reservations) and then starts shifting downward. Also note that the width of the entire set of seven bands increases until about the eighth simulation run and then remains about the same thereafter. This indicates that the amount of variability in the profit values also increases until about the eighth simulation run and then remains about the same thereafter.

## 28.8 CONCLUSIONS

Spreadsheet software is increasingly 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. Crystal Ball is an especially powerful add-in of this kind. Section 20.6 introduced the main features of Crystal Ball, and this chapter has elaborated considerably further on these features.

When using Crystal Ball, each input cell that has a random value is referred to as an *assumption cell*. The procedure for defining an assumption cell includes selecting one of 21 types of probability distributions from a Distribution Gallery to enter into the cell. When historical data are available, Crystal Ball also has a procedure for identifying which continuous distribution best fits the data.

An output cell that is used to forecast a measure of performance is called a *forecast cell*. Each trial of a simulation run generates a value in each forecast cell. When the simulation run is completed, Crystal Ball presents 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, Crystal Ball provides a *Decision Table tool* 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, the Professional Edition of Crystal Ball includes a powerful optimization module called OptQuest that is described and illustrated in Supplement 3 to Chap. 20. Based on years of research in optimization and artificial intelligence, 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.

## SELECTED REFERENCES

For general references on simulation, see the Selected References given for Chap. 20. For details regarding Crystal Ball and OptQuest, go to [www.crystalball.com](http://www.crystalball.com).

## LEARNING AIDS FOR THIS CHAPTER ON THIS WEBSITE

See the learning aids given 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

**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. 28.10 revealed that the deadline was met on 58.88 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 on this randomness. The spreadsheet model is available on this website. Make sure that the “Use Same Sequence of Random Numbers” option is *not* checked and that the Monte-Carlo Sampling Method is selected in the Sampling tab of Run Preferences.

(a) Set the number of trials to 100 in Run Preferences, 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 number of trials to 1,000 in Run Preferences.
- (c) Compare the results from part (a) and part (b), and comment on any differences.

**28-2.** Consider the historical data contained in the Excel File “Sales Data 1” on this website. Use Crystal Ball to fit all continuous distributions to these data.

- (a) Which distribution provides the closest fit to the data? What are the parameters of the distribution?
- (b) 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 Crystal Ball to fit all continuous distributions to these data.

- (a) Which distribution provides the closest fit to the data? What are the parameters of the distribution?
- (b) 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 business school. This project will require completing all of the activities in the following 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.

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

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 below. 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.

**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 finish work	5	6	7

Use Crystal Ball to generate 1,000 trials of a simulation for this project. Use the results to answer the following questions.

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

**28-5.** Refer to the financial risk analysis example presented in Sec. 28.4, including its results shown in Fig. 28.17. 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 a 10 percent chance of losing more than a full third (\$7 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.

- (a) Generate a frequency chart and a percentiles table for option two after performing a simulation with 1,000 trials for this option.
- (b) Repeat part (a) for option 3.
- (c) Suppose *you* were the CEO of the Think-Big Development Co. Use the results in Fig. 28.17 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.

**28-6.** Consider the Reliable Construction Co. bidding problem discussed in Sec. 28.1. The spreadsheet model is available on this website. The Decision Table generated in Sec. 28.7 (see Fig. 28.34) 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 Decision Table for this bidding problem that considers all bids between \$5.2 million and \$5.6 million in multiples of \$0.05 million.

- (b) Use OptQuest to find 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.

**28-7.** Consider the Everglade cash flow problem analyzed in Sec. 28.3. The spreadsheet model is available on this website.

- (a) Generate a Decision Table 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 Evergreen's mean ending balance?
- (b) Generate a trend chart for the five long-term loan amounts considered in part (a).
- (c) Use OptQuest to find the long-term loan amount that maximizes Everglade's mean ending balance.

**28-8.** Consider the airline overbooking problem discussed in Sec. 28.5. The spreadsheet model is available on this website. The Decision Table generated in Sec. 28.7 (see Fig. 28.35) 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 Decision Table 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 OptQuest to find 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.