

Using Excel to Perform Monte Carlo Simulations

By Thomas E. McKee, CMA, CPA, and Linda J.B. McKee, CPA

A critical aspect of any capital budgeting decision is evaluating the risk surrounding key variables in the decision. Whether it involves assumptions about revenue growth, expense growth, the costs of capital, or other input variables, any forecasting model invites an element of uncertainty. Performing sensitivity analysis helps reduce that uncertainty by examining the impact that different values for the assumptions might have on a project's net present value (NPV) or internal rate of return (IRR).

There are several different approaches to performing sensitivity analysis. Single variable sensitivity analysis involves varying the base input variables one at a time to

see the effect on various model outputs. Scenario analysis involves constructing worst case, best case, and most likely case scenarios. And Monte Carlo simulation uses a probability distribution to represent the uncertainty or risk associated with key variables. This is usually the preferred method to use when there is enough information or data available to identify potential value ranges for the input variables.

Commercial software is available that will enable you to perform Monte Carlo simulation, but it isn't always feasible to purchase software. For one-time projects or in situations with limited budgets, it's possible to use Excel to create a Monte Carlo simulation using three of the most common probability distributions.

Probability Distributions

The first step in creating a Monte Carlo simulation is to decide which probability distribution to apply. There are a number of ways to determine the appropriate choice, including:

- ◆ Judgment based on experience and individual knowledge,
- ◆ Computed distribution based on historical data,
- ◆ A survey of key managers about their beliefs concerning a specific variable, or
- ◆ The Delphi technique, where knowledgeable individuals (experts) are repeatedly polled on their beliefs until a consensus belief emerges.

Out of the dozens of possible probability distributions, the three used most often in capital budgeting simulations are rectangular (or continuous uniform) distribution, normal distribution, and triangle distribution. Figure 1 shows what each of these distributions looks like and describes their key variables. Each is possible to create within an Excel worksheet.

Rectangular Distribution

In a rectangular distribution, all the values within a range—from the lowest to highest—are equally probable. In Excel, this can be created using the `RANDBETWEEN` function, which generates a random number between two specified limits. For example, if you wanted to generate a value for revenue, occurring with equal probability anywhere between \$14,000 and \$16,000, the function would be `=RANDBETWEEN(14000,16000)`.

Note that `RANDBETWEEN` is a volatile function. Every time your Excel sheet recalculates, you will get a new number from the formula.

Normal Distribution

In a normal distribution, the values assume a bell-shaped curve. As indicated in Figure 1, you need to identify a mean and standard deviation. For a normal curve, two-thirds of the values (approximately 67%) fall within a range plus or minus one standard deviation around the mean.

To create a normal distribution in Excel, we can use the `RAND` function to create a standard deviate. `RAND` generates an evenly distributed random number greater than or equal to 0 and less than 1.

The average value of `=RAND()` is 0.5, but, for our purposes, we need a standard deviate centered on 0 instead. The solution is to generate 12 random numbers using `RAND`, add them together, and then subtract 6 from the

sum. This will generate a standard normal deviate centered on 0 with a range of -6 to 6. Within an Excel worksheet, this can be accomplished within a simple formula that adds up 12 `RAND()` functions and then subtracts 6.

A more elegant solution involves using a macro. The macro would follow the same process—generating a random number 12 times, summing the results, and then subtracting 6. It is relatively easy to write Visual Basic macros for Excel. The following code would return a normal deviate to cell A1:

```
Sub StandardNormalDeviate()  
SND = 0  
For Index = 1 To 12  
SND = SND + Rnd()  
Next Index  
SND = SND - 6  
Cells(1, 1) = SND  
End Sub
```

The primary reason for generating the standard deviate using a macro is that you can write it, test it carefully, and then use the `CALL` function any time you need a standard normal deviate. This might reduce the risk of a copying or typing error when developing a simulation with a large number of active cells.

To use a normal distribution in Excel, the formula would be `= Mean + (Standard Deviation * Standard Normal Deviate)`. For example, if we wanted to generate revenue with a mean of \$15,000 and standard deviation of \$1,000 (in other words, 67% of the values fall between \$14,000 and \$16,000), the formula would be `=15000+(1000*A1)`, where A1 is the cell reference for the standard normal deviate.

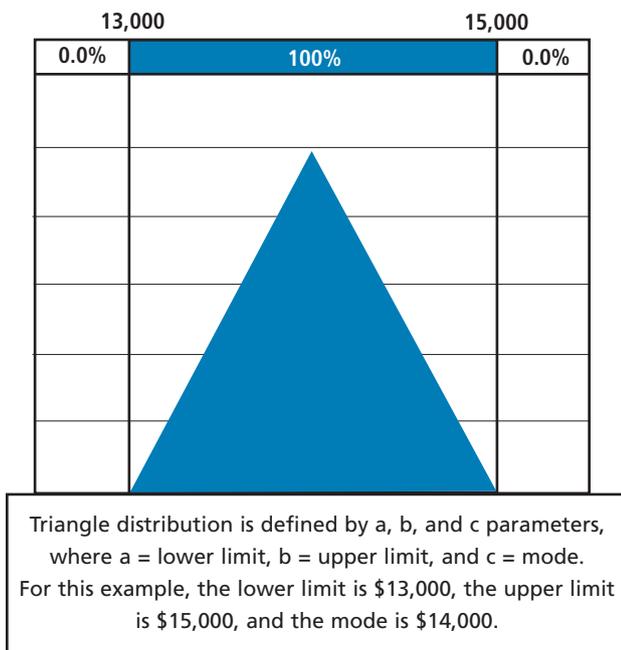
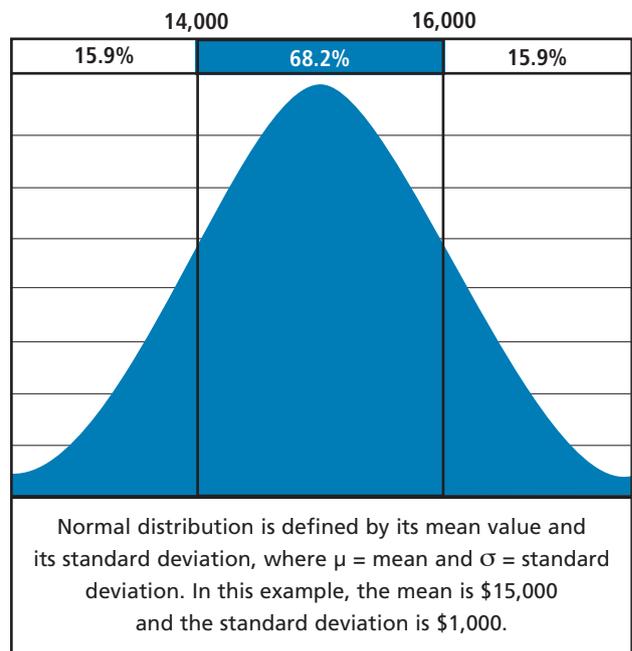
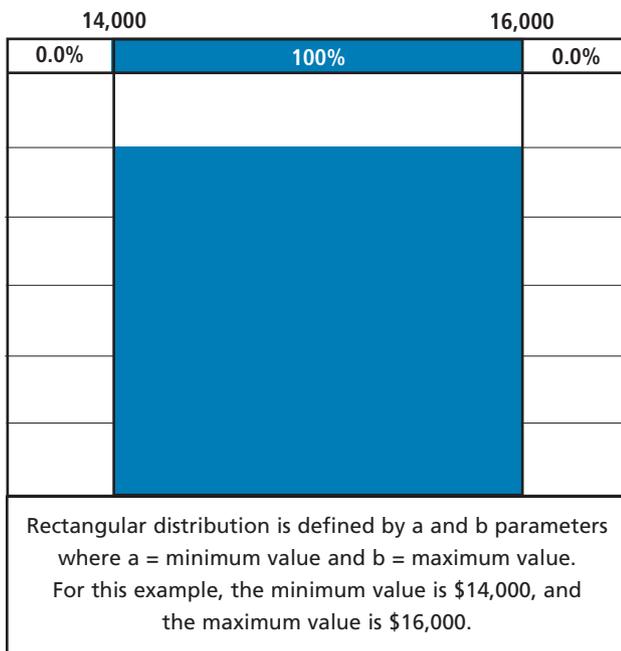
Triangle Distribution

In a triangle distribution, the values increase in probability from the minimum, peaking at the middle (mode), and then decrease in probability to maximum. The variables you need to generate a symmetrical triangle distribution in Excel are the minimum, maximum, mode, and standard deviate. For example, a triangle distribution with a minimum value of \$13,000, a maximum of \$15,000, and a mode of \$14,000 would have the following formula: `=13000+(ABS(A1)*(15000-13000)*(14000-13000))^0.5`, where A1 is again the cell reference to the standard normal deviate.

Running the Monte Carlo Simulation

Monte Carlo simulation involves assigning probability distributions to one or more key variables in an analysis.

Figure 1: Rectangular, Normal, and Triangle Distributions



In order to demonstrate this in a short article, we will perform a simple analysis involving two variables, revenues and expenses, which are used to compute the net income. Assume the initial analysis, based on the information available and the most likely assumptions regarding future performance, forecasts that revenues for the project will be \$15,000 and that expenses will be \$14,000. The result is \$1,000 in net income. But this solution doesn't incorporate the uncertainty about the forecasted rev-

enues and expenses.

To perform Monte Carlo analysis, we must apply a probability distribution to each of the variables. We will use a normal distribution for revenue since most managers would typically have sufficiently detailed revenue data to determine a reliable mean and standard deviation. For expenses, a manager would normally be able to estimate the minimum, maximum, and most likely values without having to do a detailed analysis, so we'll use a symmetrical triangle distribution.

A single iteration would only produce one value for net income based on one revenue value generated from the normal distribution and one expense value generated from the symmetrical triangle distribution. This wouldn't be particularly informative since it represents a single interaction of the two distributions, i.e., a sample size of 1. A more informative output would be the net income results for a large number of interactions from those two distributions. For example, we could generate 1,000 different net income solutions. Using the FREQUENCY and bar graph features, it's then possible to compute the output frequencies and view the shape of the distribution for the 1,000 values.

The Simulations

In order to keep the illustration manageable for this article, we'll generate 10 different net income solutions. In Figure 2, column B contains 10 different revenue values

generated from the normal distribution, and column C contains 10 different expense values generated from the triangle distribution. The formula in cell B2 is $=15000+(1000*B19)$. The formula in cell C2 is $=13000+(ABS(C19)*(15000-13000)*(14000-13000))^{0.5}$.

In these formulas, the references to cells B19 and C19 are for the standard deviate. Note that the references to these cells aren't fixed. As you can see in Figure 2, there is a small table of standard normal deviates in cells A18:C28. (For illustration purposes, this table is located below the actual simulations in Figure 2. If you were creating 1,000 simulations, it would be more sensible to put the table on a different worksheet.) Set up this way, you only need to enter the formulas in cells B2 and C2. Then you can click on the bottom-right corner of those cells and drag the mouse down the column to generate the values for the remaining simulations.

The net incomes in column D are simply the result of subtracting expenses from revenues. The formula in cell D2, for example, is $=B2-C2$.

The Histogram

The next step is to use the data from the 10 simulations to create a histogram. This provides a graphical representation of the distribution of data.

The first step is to divide the entire range of data into smaller intervals, or bins. Cells E2:E11 contain the desired

bins for the Net Income results: $-\$1,000$, $-\$500$, $\$0$, $\$500$, $\$1,000$, $\$1,500$, $\$2,000$, $\$2,500$, $\$3,000$, $\$3,500$. These bins can be thought of as points that separate the solutions into relevant categories. You would, of course, adjust the bins to fit your data.

To create the histogram, we need to know how frequently the net income results fall within each particular bin. For example, the first bin is for values less than $-\$1,000$, and we want to determine how many times the net incomes of our simulations are less than $-\$1,000$. Excel can calculate this with the FREQUENCY function.

FREQUENCY has two arguments. The first is the actual data, and the second is the bins. Because we need a frequency for all 10 bins, we need to enter the formula as a matrix command:

1. Select cell F2.
2. Highlight cells F2:F11.
3. Type in the formula: $=\text{FREQUENCY}(D2:D11, E2:E11)$.
4. Press Ctrl+Shift+Enter.

That last step tells Excel that the formula is a matrix command for all cells in the highlighted range. If you select one of the cells containing the formula, it should appear as $\{=\text{FREQUENCY}(D2:D11E2:E11)\}$.

The final step for the histogram is to use the data in columns E and F to create a simple bar graph (his-

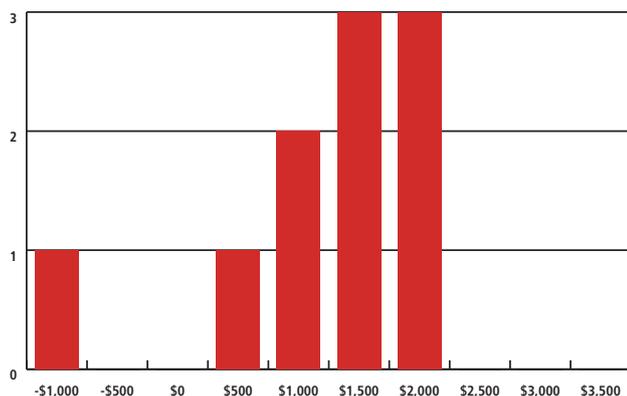
togram), as shown in Figure 3. When the histogram is created with significantly more simulations than the 10 in our example, it will provide a much richer picture of the final result, similar to one of the distributions in Figure 1. If we were to run 100 simulations instead of 10, for example, the result might be more similar to Figure 4. This graph will be very valuable in presenting this information to other parties since most individuals intuitively understand images better than numbers.

In addition to viewing a frequency graph of outcomes, the average value and standard deviation of the solution outcome can be calculated (see rows 13 and 15 of Figure 2). This provides additional information about risk because you can see

Figure 2: Monte Carlo Net Income Analysis

	A	B	C	D	E	F
	Simulation #	Revenues	Expenses	Net Income	Bin Limits	Frequency
1						
2	1	\$14,868	\$14,413	\$455	-\$1,000	1
3	2	\$14,451	\$13,689	\$762	-\$500	0
4	3	\$15,871	\$14,370	\$1,501	\$0	0
5	4	\$14,866	\$13,449	\$1,416	\$500	1
6	5	\$12,854	\$14,496	-\$1,643	\$1,000	2
7	6	\$16,159	\$14,278	\$1,881	\$1,500	3
8	7	\$14,832	\$14,222	\$611	\$2,000	3
9	8	\$16,085	\$14,503	\$1,582	\$2,500	0
10	9	\$14,624	\$13,518	\$1,107	\$3,000	0
11	10	\$15,676	\$14,524	\$1,152	\$3,500	0
12						
13	Average	\$15,029	\$14,146	\$882		
14						
15	Standard Deviation			\$996		
16						
17		Standard Deviates				
18	Simulation #	Revenues	Expenses			
19	1	-0.131627711	0.99845759			
20	2	-0.548628516	0.23761816			
21	3	0.871206129	-0.9389644			
22	4	-0.134334021	-0.1008904			
23	5	-2.146437205	1.11934896			
24	6	1.158721141	0.81685648			
25	7	-0.167871357	0.74603714			
26	8	1.085332375	-1.129815			
27	9	-0.375856554	-0.1339637			
28	10	0.675712381	-1.1606952			

Figure 3: Frequency Distribution for Net Income with 10 Simulations



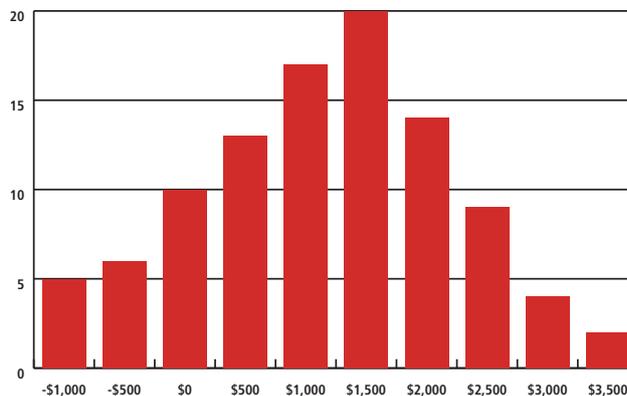
what extreme values might be possible or, by using normal curve probabilities, determine the probability of a value occurring within an interval or above or below a specific value. Knowing the extreme values helps answer questions such as, “What is the worst case that might occur?”

Factoring in Uncertainty

The simple analysis we did with 10 simulations reveals that although the most likely net income estimate is \$1,000, possible outcomes actually range from a loss of \$1,643 to a gain of \$1,881. By considering the uncertainty about revenue and expenses, we can see that a loss is possible. At the same time, a look at the histogram shows a positive outcome for a large proportion of simulations. In fact, many of the simulations had results better than the original gain of \$1,000. This should provide a more detailed picture when making the final decision on the project.

Though we only simulated 10 solutions, a larger number of simulation runs will give you more accurate solutions. There’s no exact formula in choosing the number of simulations. When doing this in Excel, there’s a trade-off between the effort to generate more simulations and the increased accuracy of the outputs. It’s generally best to do as many simulations as necessary to determine if you have a rich, stable solution. You can see from the gaps in Figure 3 that 10 simulations don’t produce a rich solution. For a simple problem with two variables like ours, 100 to 1,000 individual solutions might be enough to get a good approximation, as shown in Figure 4. A complex problem with many uncertain variable inputs might require 10,000 to one million iterations.

Figure 4: Frequency Distribution for Net Income with 100 Simulations



The larger the number of iterations, the harder it is to perform by entering formulas directly into Excel worksheets. In those instances, you might want to consider programming the calculations in Visual Basic, using loops like we did in the macro for the standard normal deviate, to get the larger number of solutions needed.

No decision that involves predicting the future can be guaranteed, but simulations can provide an accurate projection with more information about the uncertainty faced in the future. When you capture the variability in your input data by using probability distributions instead of simple point estimates, your result will give you a probability distribution that provides a better picture of the actual risks faced.

Once you get used to the basic process we described here, you’ll have a simple, readily available model that can provide greater insight into the risk and uncertainty surrounding complex budgeting decisions. **SF**

Thomas E. McKee, CMA, CPA (licensed in N.C., inactive in Tenn.), CIA, Ph.D., is a professor in the Department of Healthcare Leadership & Management at the Medical University of South Carolina in Charleston, S.C., and a professor II in the Department of Accounting, Auditing, and Law at the Norwegian School of Economics, Bergen, Norway. He also is a member of IMA’s Mountain Empire Chapter. You can reach him at temckee@musc.edu.

Linda J.B. McKee, CPA, Ph.D., is an associate professor in the Department of Accounting and Legal Studies at the College of Charleston in Charleston, S.C. You can reach her at bradleyl@cofc.edu.