
Chapter 5: @RISK Modeling Techniques

| | |
|--|------------|
| Introduction | 93 |
| Models from the book <i>Financial Models Using Simulation and Optimization</i> | 94 |
| Loading the Example Models | 94 |
| Modeling Interest Rates and Other Trends | 95 |
| Projecting Known Values into the Future | 97 |
| Increasing Uncertainty Over Time | 97 |
| Modeling Uncertain or "Chance" Events | 99 |
| Will the Flood Occur or the Competitor Enter the Market? | 99 |
| Oil Wells and Insurance Claims | 101 |
| Modeling an Uncertain Number of Events, Each with Uncertain Parameters | 101 |
| Adding Uncertainty Around a Fixed Trend | 103 |
| I Have To Use This Projection But Don't Trust It | 103 |
| Dependency Relationships | 105 |
| Using Variable Arguments and Correlations – These Values Will Be Affected By What Happens Somewhere Else | 105 |
| Sensitivity Simulation | 107 |
| How Do Changes in Model Variables Affect My Simulation Results? | 107 |
| Simulating a New Product: The Hippo Example | 109 |
| Tornado Graphs and Scenarios | 115 |
| Finding Value at Risk (VAR) of a Portfolio | 119 |
| Simulating the NCAA Tournament | 123 |
| Remarks | 125 |

Introduction

The *@RISK Modeling Techniques* chapter will show you how to translate typical "risky" situations into @RISK models. These risky situations have been identified from real-life modeling problems that Excel users often encounter. As you use @RISK to analyze uncertainty in your Excel worksheets, look through the examples and illustrations provided in this chapter. You just may find some helpful tips or techniques which will make your @RISK models better representations of uncertain situations.

Seven @RISK techniques are presented here to illustrate common modeling situations under uncertainty. To help you understand the modeling techniques employed, example Excel worksheets and accompanying @RISK simulations are provided with your @RISK system. The simulations are even "pre-run" so you can just look at the results if you like. As you work through each modeling technique discussed, look at the corresponding worksheet and simulation. It will help you understand the @RISK concepts and techniques involved in modeling each risky situation.

The seven modeling techniques illustrated here are:

- **Modeling Interest Rates and Other Trends** - random trends over time and "random walks".
- **Projecting Today's Known Values Into the Future** - an increasingly uncertain future or "increasing variability".
- **Will the Flood Occur or the Competitor Enter the Market?** – modeling uncertain "chance" events.
- **Oil Wells and Insurance Claims** - modeling an uncertain number of events, each with uncertain parameters.
- **"I have to use this projection but don't trust it"** - adding uncertainty around a fixed trend using "error terms".
- **"These values will be affected by what happens someplace else"** - dependency relationships using variable arguments and correlations.
- **Sensitivity Simulation** - how model changes affect simulation results.

Models from the book *Financial Models Using Simulation and Optimization*

In addition to the seven models discussed here, this chapter also includes three @RISK examples from the book **Financial Models Using Simulation and Optimization** by Wayne Winston. These models illustrate how @RISK is applied to everyday business modeling. The full **Financial Models** book includes 63 examples of how @RISK and other add-ins can be applied to a wide variety of financial problems. For more information on purchasing the full *Financial Models* book, contact Palisade Corporation or visit www.palisade.com.

Loading the Example Models

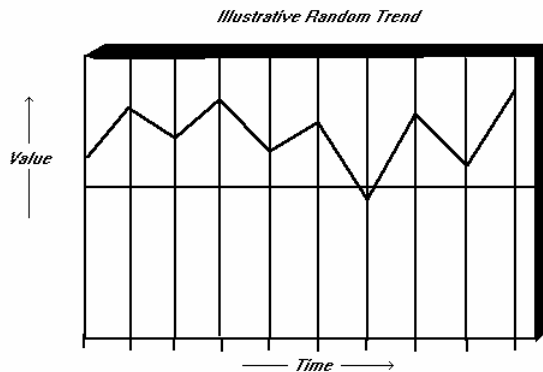
All example spreadsheet models discussed here can be found in the default installation location C:\PROGRAM FILES\PALISADE\RISK45\EXAMPLES.

Modeling Interest Rates and Other Trends

Projecting Trends

Example Model:RATE.XLS

Whether you are getting a mortgage or evaluating the cost of a variable rate loan, projections of future interest rates are highly uncertain. The movement of the rate of interest you are charged is often viewed as random - moving up and down erratically year to year. This movement may be completely random or it may be a random fluctuation around an underlying known trend. In either case, modeling the random portion of any projection is an important technique in Risk Analysis.



Simulation accounts for randomness in a trend over time in a very powerful way - repetitively trying a different possible series of rates with each iteration of a simulation. For example, you could set up a random trend to project the interest rate over ten years. For each iteration, a new, randomly selected value is chosen for each year's interest rate, and results are calculated. By doing this, simulation includes the effects of all possible future interest rates in your results, instead of just a single, most likely projection.

A "random trend" can be easily and directly included in an Excel worksheet with @RISK. And by using the Excel Copy commands, you can place a random trend anywhere in your worksheet.

Simple Random Trends

The simplest random trend is a distribution copied across time. The value randomly selected in one period is independent of the value selected in any other time period:

- 1) Enter a distribution function for the first cell of the trend
- 2) Copy the distribution over the range of cells

In this case, a new value will be sampled every period - a completely random trend with no correlation over time.

A "Random Walk" with Correlation Period to Period

Maybe you don't think future rates are entirely random. Perhaps next year's rate will be influenced by this year's. In Excel terms, there will be some correlation between one cell in the range and the next. Here's a simple way to model this:

- 1) Enter a distribution function for the first cell in the range.
- 2) Enter a distribution function for the second cell in the range that uses the value sampled for the first cell as one of its arguments (such as its mean or most likely value).
- 3) Copy the formula of the second cell across the range. The referenced argument in the formula is a relative reference - the third cell will use the value in the second as its referenced argument and likewise for the fourth, fifth, etc.

For example:

A1: RiskNormal(100,10)

A2: RiskNormal(A1,10)

A3: RiskNormal(A2,10)

A4: RiskNormal(A3,10)

In this manner there is some correlation between one cell and the next cell in the range.

Refining Your Random Trends

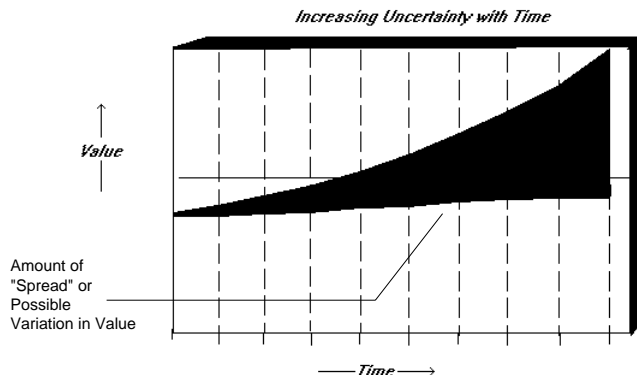
These are just simple examples of modeling random processes over time. As you get more sophisticated, you can include your random terms in formulas which place limits or caps on the amount of change, increase the amount of possible change with time, or other such extensions or variations. And remember, interest rates are just one application for random trends. Look at your Excel worksheets and the uncertain situations you model and you will undoubtedly see others.

Projecting Known Values into the Future

Increasing Uncertainty Over Time

Example Model: VARIABLE.XLS

You know today's values for critical variables in your models, but what about values for these same variables in the future? Time often has a very important impact on estimates - they become less and less certain the further out in time your projections extend. As a consequence, results based on your single "best estimates" become more and more risky the further out in time they are projected.



The widening variation around the trend of best estimates illustrates this problem. @RISK lets you model the effect of time on your estimates by allowing you to easily increase the variability in a random value over time.

The range of possible values for a worksheet cell is given in your distribution functions. As you move out in time - across a range of worksheet cells - the argument of the function which specifies the range of possible values can increase. For example:

A1: RiskLognorm(10,10)

A2: RiskLognorm(10,15)

A3: RiskLognorm(10,20)

A4: RiskLognorm(10,25)

The standard deviation of the LOGNORM distribution function controls the possible variation in value. In this example, as you move out across the range of cells, the standard deviation increases.

Increasing the possible variance in value as your projections extend further and further into the future is a good "rule of thumb" to follow. By doing this, your results will more accurately reflect the greater uncertainty which exists in your knowledge of the distant future.

Modeling Uncertain or "Chance" Events

Will the Flood Occur or the Competitor Enter the Market?

Example Model: DISCRETE.XLS

Uncertainty often shows up in the form of chance events which may have significant impact on your results. Either we will strike oil, or we won't. That competitor will either enter the market or he won't, but if he does! There's a 25% chance of a hailstorm that will wipe out this year's crop.

Including the possibility of these types of events in your models is an important technique in Risk Analysis. If you leave them out, the outcomes caused by these events will not be included in your results and your models will be incomplete. Using the DISCRETE function provided by @RISK and Excel's IF function, modeling these events is easy.

Including the Discrete Event

The DISCRETE function is the means by which you can include probabilities for chance events in your worksheet models.

RiskDiscrete({0,1},{50,50})

This example models the characteristic "coin flip" - the simplest chance event. In this case, an outcome of 0 represents Heads and 1 represents Tails, and each is equally likely to occur. A more complex example illustrates four possible scenarios for annual storm damage from floods:

RiskDiscrete({0,1,2,3},{20,40,30,10})

In this case, outcomes valued 0 to 3 represent four possible levels of flood damage ranging from none (0) through low (1), medium (2) and high (3). The probability of occurrence of no damage is 20%, low 40%, medium 30% and high 10%.

Including the Effects of Chance Events in Your Worksheets

The DISCRETE function returns a value for each iteration that indicates which event has occurred. Your worksheet model has to recognize which event has occurred and calculate different results appropriate to the event. The Excel IF function allows this. Consider the following example and Excel cell entries:

Cell C2 describes an event - the possible entry of a competitor in a given market. There is a 50-50 chance of entry. If entry occurs, your sales level will be 65; if entry does not occur, your sales level will be 100.

C2: RiskDiscrete({0,1},{50,50})

D2: IF(C2=1,100,65)

In the above example, the IF function in cell D2 will return a value of 100 if the outcome in cell C2 is 1 (no entry) and will return 65 if the outcome is 0 (entry). This simple example can be extended throughout your @RISK models. Each iteration of a simulation the DISCRETE function will return one of the possible values. Depending on the value returned, the calculations in your worksheet can change.

Caution

People used to working with single estimates in spreadsheets often substitute discrete distributions where a continuous distribution should be used. For example, they use a discrete distribution to put in three possible price levels where, in reality, price could take any value in a range.

This common mistake occurs because many people are used to manual "what-if" modeling which necessarily limits the user to a small number of discrete estimates. Use a continuous form when any value in a range is possible, and save DISCRETE for event modeling and variables that truly are discrete.

Oil Wells and Insurance Claims

Modeling an Uncertain Number of Events, Each with Uncertain Parameters

Example Model: CLAIMS.XLS

In real life situations, uncertainty often has two or more dimensions. The situation you face may have an uncertain number of events, each of which has an uncertain value. Think, for example, of the insurance industry. An uncertain number of claims may be filed on a new policy. And each claim filed has an uncertain dollar amount. How do you simulate the total claim amount possible? The oil industry faces a similar problem. When drilling oil wells, an uncertain number of wells will be successful. But the amount of oil discovered by each successful well is uncertain. How do you simulate the total amount of oil discovered?

Risk Analysis is very useful in modeling situations such as this. @RISK, combined with the Excel IF function, can provide an easy means of performing such an analysis. It is helpful to review the example simulation, CLAIMS.XLS, while proceeding through this modeling technique.

To model situations where 2 or more levels of uncertainty are present, a worksheet is set up which includes a column of calculations for each of the possible events. For example, if the maximum number of possible claims is 100, 100 columns would be used, each to calculate the results from a single possible claim. A claim # (1-100) would be given at the top of each column.

To run the analysis:

- 1) First, a cell is used to sample the number of events which occur in a given iteration
- 2) The number of events is compared to the # at the top of each column which refers to the calculation of results for a specific event

The number of events sampled is compared to the "claim #" cell at the top of each column. The Excel IF function is used to make the comparison.

For example, if cell A1 holds the *Number of Claims* (sampled from a probability distribution function) and B5 has the *Claim #*:

Number of Claims **A1: 6**

Claim # **B5: 3**

Comparison **B6: IF(B5 <= A1, RiskNormal(100,10), 0)**

The formula in cell B6 says "if the *Claim #* is less than or equal to the *Number of Claims*, return a sample from the normal distribution - otherwise return a value of zero". In the example above, the value in cell B5 is less than the value in A1 so a sample from the normal distribution is returned.

Using this structure, it is possible to evaluate a changing number of events each iteration. And because it is very simple in Excel to copy the calculation for one event across a worksheet, you just have to set up one column of calculations for a single event, then copy it. In the example here, you would set up a column of calculations for Claim #1, and then copy that column across the worksheet to end up with a total number of columns equal to the maximum possible number of events.

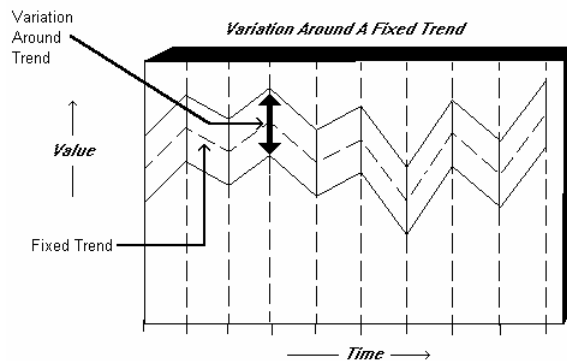
Adding Uncertainty Around a Fixed Trend

I Have To Use This Projection But Don't Trust It

Example Model: ERROR.XLS

Often Excel modelers are given data from other sources for inclusion in their worksheets. "The economics group has provided this projection for GNP growth, so include it in your worksheet model" - could be the guideline. But how often does the future exactly follow even the best projection?

Recognizing the uncertainty inherent to projections, you still may want to stay true to the basic direction offered by the values in the trend. In this case "error terms" let you put a certain amount of variation around the values in the trend. This allows you to examine how variation in trend value will impact your results.



With @RISK, you can easily append an error term to a fixed trend you have already entered in your worksheet. Let's say, for example, that worksheet row B contains the fixed trend in your model. An error term is just a factor you'll multiply each worksheet cell value by. (You also could add an error term to each trend value)

Row B - GNP Growth in Percent

*B1: 3.2 * RiskNormal(1,.05)*

*B2: 3.5 * RiskNormal(1,.05)*

*B3: 3.4 * RiskNormal(1,.05)*

*B4: 4.2 * RiskNormal(1,.05)*

*B5: 4.5 * RiskNormal(1,.05)*

*B6: 3.5 * RiskNormal(1,.05)*

*B7: 3.0 * RiskNormal(1,.05)*

In this example taken from Excel, the error term for all trend values is drawn from a normal distribution with a mean of 1 and a standard deviation of .05. For each iteration of a simulation, a new error term will be sampled for each cell and will be used to multiply the fixed trend estimate in that cell, allowing variation around the fixed estimate.

An added bonus of an error term is the expected value generated in normal Excel recalculations. Because the expected values of the error terms in the example equal one, they will not affect your normal worksheet recalculations. Thus you can leave the error terms in your formulas and only see their effects when you run your simulations. The same comment holds true if you add, instead of multiply by, the error term. If you add the error term to the fixed estimate, the mean of the error term probability distribution should be zero.

Dependency Relationships

Using Variable Arguments and Correlations — These Values Will Be Affected By What Happens Somewhere Else

Example Models: DEP.XLS, CORRMAT.XLS

Many times you will not know precise argument values for a distribution function in your worksheet. Often the range for a worksheet cell will depend on a value calculated or sampled elsewhere in your model. "If price is low, the range for sales volume is one million to 2 million - but if price is high, the range is only 500,000 to 750,000" is an illustration of this type of dilemma.

Two modeling techniques in @RISK help you to resolve problems such as this, namely (1) *variable arguments for distribution functions* and (2) *correlations in sampling*.

Variable Arguments

The first technique - variable arguments for distribution functions - relies on a standard Excel capability most modelers are familiar with. The referencing of cell addresses in functions is allowed in @RISK just as in Excel. For example:

Minimum *A1: RiskTriang(10,20,30)*

Maximum *B1: RiskNormal(80,10)*

Final Price *C1: RiskUniform(A1,B1)*

This example shows how the range for the uniform distribution for *Final Price* will change with the value sampled for *Minimum* and *Maximum*. Here the range for *Final Price* will change with each iteration of the simulation. *Final Price* thus depends on the variables *Minimum* and *Maximum*.

Correlations in Sampling

The second modeling technique which can be used to affect sampled values based on other calculations in a worksheet is *correlations in sampling*. The @RISK function CORRMAT is used to correlate the values sampled in different distribution functions. This correlation allows you to specify a relationship between the values sampled in different worksheet cells while still maintaining a degree of uncertainty for each.

Correlating Interest Rates and Housing Starts

Interest Rate

A1: RiskUniform(6,14, RiskCorrmat (D1: E2,1))

Housing Starts:

B1: RiskUniform(100000,200000, RiskCorrmat (D1: E2,2))

The CORRMAT function is used when you want to have the value sampled for one cell influence the value which is sampled for another. The variable *Interest Rate* - described by the distribution RiskUniform(6,14) - is the distribution that the *Housing Starts* distribution - RiskUniform(100000,200000) - should be correlated with. The range D1:E2 contains a matrix of 4 cells and a single correlation coefficient, -.75. The -.75 is the coefficient which specifies how the two sampled values are correlated. Coefficients range from -1 to 1. A value of -.75 is a negative correlation - as *Interest Rate* goes up, *Housing Starts* go down.

When you are using sampled, uncertain variables in your worksheet models it is important to recognize correlations in sampling. If you don't use methods such as the two presented here, all uncertain variables will be sampled as if they were completely independent of other variables in the model. This can cause erroneous results. Think about what could happen if *Interest Rate* and *Housing Starts* in the above example were completely independent. *Interest Rate* and *Housing Starts* would be sampled entirely independent of one another. A high *Interest Rate* and a high value for *Housing Starts* would be an entirely possible scenario during sampling. But could this happen in real life? Not in this economy!

Correlating Multiple Distributions

Correlating multiple distribution functions can be accomplished using the CORRMAT function or by selecting the cells containing the distributions in the @RISK Model Window and selecting the **Model menu Correlate Distributions command**. Either of these methods allows you to enter a matrix of correlation coefficients. @RISK then uses the coefficients to correlate the sampling of distribution functions. This is especially useful when pre-existing correlation coefficients (calculated from actual collected data) are available and you want sampling to be governed by those coefficients. Excel can calculate correlations from existing datasets using the CORREL function. For more information on using Correlate or CORRMAT, see the example simulation CORRMAT.XLS.

Sensitivity Simulation

How Do Changes in Model Variables Affect My Simulation Results?

Example Model: SENSIM.XLS

@RISK lets you see the impact of uncertain model parameters on your results. But what if some of the uncertain model parameters are under your control? In this case the value a variable will take is not random, but can be set by you. For example, you might need to choose between some possible prices you could charge, different possible raw materials you could use or from a set of possible bids or bets. To properly analyze your model, you need to run a simulation at each possible value for the "user-controlled" variables and compare the results. A Sensitivity Simulation in @RISK allows you to quickly and easily do this - offering a powerful analysis technique for selecting between available alternatives.

The benefits of Sensitivity Simulation are not limited to evaluating the impacts of user-controlled variables on simulation results. A sensitivity analysis can be run on the probability distributions which describe uncertain variables in your model. You may wish to repetitively re-run a simulation, each time changing the parameters of one (or several) of the distributions in your model. After all the individual simulations are complete, you can then compare the results from each.

The key to a Sensitivity Simulation is the repetitive simulation of the same model while making selected changes to the model each simulation. In @RISK any number of simulations can be included in a single Sensitivity Simulation. The SIMTABLE function is used to enter lists of values, which will be used in the individual simulations, into your worksheet cells and formulas. @RISK will automatically process and display the results from each of the individual simulations together, allowing easy comparison.

To run a Sensitivity Simulation:

- 1) First, enter the lists of values you want used in each of the individual simulations into your cells and formulas using SIMTABLE. For example, possible price levels might be entered into Cell B2:

B2: RiskSimtable({100,200,300,400})

will cause simulation #1 to use a value of 100 for price, simulation #2 to use a value of 200, simulation #3 to use a value of 300 and simulation #4 to use a value of 400.

- 2) Set the number of simulations in the Settings dialog box and run the Sensitivity Simulation using the Start Simulation command.

Each simulation executes the same number of iterations and collects data from the same specified output ranges. Each simulation, however, uses a different value from the SIMTABLE functions in your worksheet.

@RISK processes Sensitivity Simulation data just as it processes data from a single simulation. Each output cell for which data was collected has a distribution for each simulation. Using the functions of @RISK you can compare the results of the different alternatives or "scenarios" described by each individual simulation. The Distribution Summary graph summarizes how the results for an output range change. There is a different summary graph for each output range in each simulation, and these graphs can be compared to show the differences between individual simulations. In addition, the Simulation Summary report is useful for comparing results across multiple simulations.

You also can use Sensitivity Simulation to see how different distribution functions affect your results. The values entered in the SIMTABLE function can be distribution functions. For example, you may wish to see how your results change if you alternately try TRIANG, NORMAL, or LOGNORM as the distribution type in a given cell.

Caution

It is important to distinguish between 1) controlled changes by simulation (which are modeled with the SIMTABLE function) and 2) random variation within a single simulation (which is modeled with distribution functions). SIMTABLE should not be substituted for DISCRETE when evaluating different possible random discrete events. Most modeling situations are a combination of random, uncertain variables and uncertain but "controllable" variables. Typically, the controllable variables will eventually be set to a specific value by the user, based on the comparison conducted with a Sensitivity Simulation.

Advanced Sensitivity Analysis

@RISK 4.5 Professional and Industrial versions include an advanced analysis tool called **Advanced Sensitivity Analysis**. This analysis greatly expands on the sensitivity simulation capabilities described here. For more information on Advanced Sensitivity Analysis, see the Advanced Analyses command in the Reference: @RISK Add-in Menu section of this manual.

Simulating a New Product

The Hippo Example

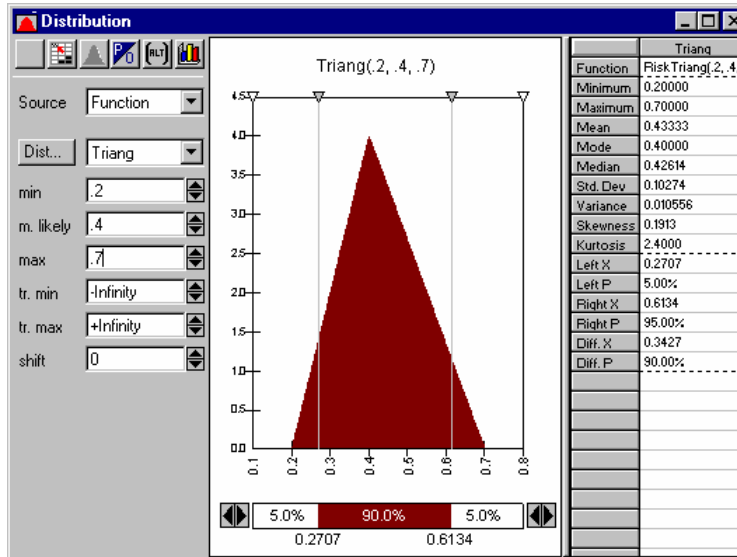
(Chapter 28, Financial Models Using Simulation and Optimization)

When a company develops a new product, the profitability of the product is highly uncertain. Simulation is an excellent tool to estimate the average profitability and riskiness of new products. The following example illustrates how simulation can be used to evaluate a new product.

Example 28.1

ZooCo is thinking of marketing a new drug used to make hippos healthier. At the beginning of the current year there are 1,000,000 hippos that may use the product. Each hippo will use the drug (or a competitor's drug) at most once a year. The number of hippos is forecasted to grow by an average of 5% per year, and we are 95% sure that the number of hippos will grow each year by between 3% and 7%. We are not sure what use of the drug will be during year 1, but our worst case guess is 20% use, most likely use is 40% and best case use is 70%. In later years, we feel the fraction of hippos using our drug (or a competitor's) will remain the same, but in the year after a competitor enters, we lose 20% of our share for each competitor who enters. We will model Year 1 market use with a **triangular random variable**. See Figure 28.1. Basically, @RISK will generate Year 1 market use by making the likelihood of a given market use proportional to the height of the "triangle" in Figure 28.1. Thus a 40% Year 1 market use is most likely; a 30% market use occurs half as often as a Year 1 40% market use, etc. The maximum height of the triangle is 4, because that makes the total area under the triangle equal to one. Probability of market use being in a given range is equal to area in that range under the triangle. For example, the chance of a market use being at most 40% is $.5 * (4) * (.4 - .2) = .4$ or 40%.

Figure 28.1



There are three potential entrants (in addition to ZooCo). At the beginning of each year each entrant who has not already entered the market has a 40% chance of entering the market. The year after a competitor enters, our market use drops by 20% for each competitor who entered. Thus if in Year 1 two competitors enter the market, in Year 2 our market use will be reduced by 40%. To model the number of entrants you can use the **binomial random variable** (in @RISK this requires us to use the =RISKBinomial function). The formula

$$= \text{RISKBinomial}(n, p)$$

generates n independent binomial trials (each a success or failure) having probability of success p and keeps track of the number of successes.

We consider a "success" to be a competitor entering the market. Then the formula

$$= \text{RISKBinomial}(2, .4)$$

will simulate the number of entrants during a year in which two competitors have yet to enter the market. Make sure that if all three entrants have entered, no more entrants may enter.

Each unit of the drug is sold for \$2.20 and incurs a variable cost of \$0.40. Profits are discounted by 10% (risk adjusted rate) annually.

Find a 95% CI for risk-adjusted NPV of project. For now we ignore the fixed cost of developing the drug.

Recall that **risk-adjusted** NPV is expected discounted value of cash flows (discounted at risk-adjusted rate).

Solution

Our spreadsheet is in Figure 28.2 (*file hippo.xls*).

Figure 28.2

| | A | B | C | D | E | F |
|----|---------------------------------|--------------|------------------|--------------|--------------|--------------|
| 1 | Pigco | | | | | |
| 2 | Price | \$ 2.20 | Compet %age | 0.2 | | |
| 3 | Unit Var Cost | \$ 0.40 | Year 1 Market S | 1000000 | | |
| 4 | Interest Rate | 0.1 | Year 1 worst sha | 0.2 | | |
| 5 | Entrant Prob | 0.4 | Year 1 most like | 0.4 | | |
| 6 | | | Year 1 best shar | 0.7 | | |
| 7 | Year | 1 | 2 | 3 | 4 | 5 |
| 8 | Market Size | 1000000 | 1050000 | 1102500 | 1157625 | 1215506.25 |
| 9 | Use per hippo of our drug | 0.4333333333 | 0.3466666667 | 0.2773333333 | 0.2773333333 | 0.2773333333 |
| 10 | Competitors (beginning of year) | 0 | 1 | 2 | 2 | 2 |
| 11 | Entrants | 1 | 1 | 0 | 0 | 0 |
| 12 | Unit Sales | 433333.3333 | 364000 | 305760 | 321048 | 337100.4 |
| 13 | Revenues | \$ 953,333 | \$ 800,800 | \$ 672,672 | \$ 706,306 | \$ 741,621 |
| 14 | Costs | \$ 173,333 | \$ 145,600 | \$ 122,304 | \$ 128,419 | \$ 134,840 |
| 15 | Profits | \$ 780,000 | \$ 655,200 | \$ 550,368 | \$ 577,886 | \$ 606,781 |
| 16 | | | | | | |
| 17 | NPV | \$2,435,545 | | | | |

Step by Step

Step 1: In row 8 we determine the market size during each of the next five years. In B8 we enter =D3. Assuming year to year growth in market size is normally distributed, the given information tells us that the number of pigs grows from year to year by a percentage which is a normal random variable having mean .05 and standard deviation of .01. This follows because 95% of the time a normal random variable is within 2 standard deviations of its mean. Therefore we may conclude $2\sigma = .02$ or $\sigma = .01$. Thus in C8 we determine the Year 2 Market Size with the formula

$$=B8*RISKNormal(1.05,0.01).$$

Essentially, this formula ensures that each year there is a 68% chance that the size of the hippo market grows by between 4% and 6%, a 95% chance that the hippo market grows by between 3% and 7%, and a 99.7% chance that the hippo market grows by between 2% and 8%. Copying this formula to D8:F8 generates the market size for Years 3-5.

Step 2: In row 9 we determine our market use/hippo for each year. Year 1 market use/hippo is computed in B9 with the formula

$$=RISKTriang (D4,D5,D6).$$

In C9:F9 we account for the fact that the year after entry, each entrant takes away 20% of our market share. Thus in C9 we compute our Year 2 Market use/hippo with the formula

$$= B9*(1-B11*$$D$2).$$

Copying this formula to D9:F9 computes Years 3-5 market share.

Step 3: In Row 11 we determine the number of entrants during each year. If less than 3 competitors have entered, then each competitor who has not yet entered has a 40% chance of entering during the current year. If all three competitors have entered, then nobody can enter. In B11 we compute the number of Year 1 entrants with the formula

$$=If(B10<3, RISKBinomial(3-B10,B5),0).$$

Copying this to C11:F11 computes Years 2-5 entrants. If we do not use the =If statement then in a year after all 3 competitors have entered we will obtain an error message because =RISKBinomial cannot take 0 trials as the first argument.

Step 4: In Row 10 we compute the number of competitors present at the beginning of each Year by adding the number of new entrants to the number of competitors already present. In B10 we enter 0 and in C10 we enter

$$= B10 + B11.$$

Copying this formula to D10:F10 computes the number of competitors present at the beginning of each year.

Step 5: In row 12 we compute each year's unit sales = (use/hippo)*market size by copying the formula

$$= B8*B9$$

from B12 to C12:F12.

Step 6: In row 13 we compute our annual revenues by copying the formula

$$=B\$2*B12$$

from B13 to C13:F13.

Step 7: In row 14 we compute our annual variable costs by copying the formula

$$= \$B\$3*B12$$

from B14 to C14:F14.

Step 8: In row 15 we compute our annual profits by copying the formula

$$=B13-B14$$

from B15 to C15:F15.

Step 9: In B17 we compute the NPV of our 5-year profits with the formula

$$= NPV(B4,B15:F15).$$

Step 10: We now run a simulation with cell B17 (NPV) being our forecast cell. We used 500 trials. Our results follow.

Figure 28.3

| | |
|---------------------------|-------------|
| Minimum | 966350.625 |
| Maximum | 4416799 |
| Mean | 2312372.866 |
| Standard Deviation | 633418.2967 |
| Variance | 4.01219E+11 |
| Skewness | 0.319535665 |
| Kurtosis | 2.669867094 |
| NumErrs | 0 |
| Mode | 2212183.8 |
| 5% | 1347058.879 |
| 10% | 1469903.378 |
| 15% | 1630650.378 |
| 20% | 1751069.643 |
| 25% | 1858571 |
| 30% | 1933889.883 |
| 35% | 2012388.996 |
| 40% | 2096331.542 |
| 45% | 2181581.218 |
| 50% | 2265116.5 |
| 55% | 2332394.792 |
| 60% | 2432266.516 |
| 65% | 2535186.715 |
| 70% | 2636504.196 |
| 75% | 2765366.75 |
| 80% | 2877606.315 |
| 85% | 2985581.684 |
| 90% | 3165375.941 |
| 95% | 3399469.707 |

Our point estimate of risk adjusted NPV is the sample mean of NPV's from simulation (\$2,312372.866). To find a 95% confidence interval for the mean in a simulation use the fact that we are 95% sure actual mean NPV is between

$$(\text{Sample Mean of NPV}) \pm 2 * (\text{Sample Standard deviation}) / \sqrt{n},$$

where n = number of iterations.

For example, we are 95% sure the mean NPV (or risk adjusted NPV) is between

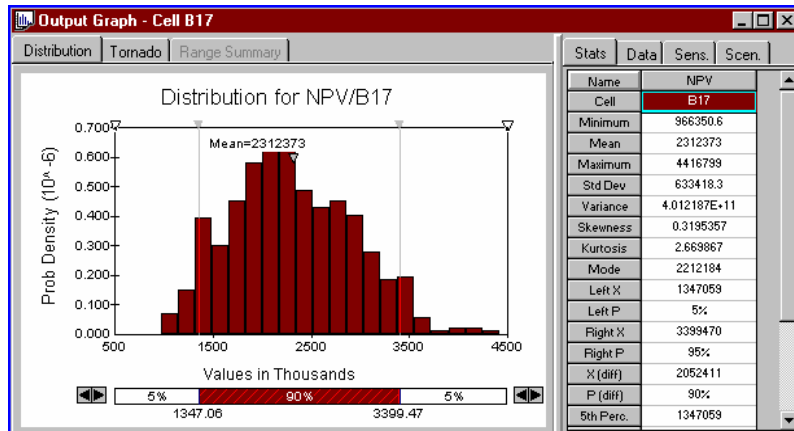
$$2,312373 \pm 2 * (633418) / \sqrt{500} \text{ or}$$

\$2,255,718 and \$2,369,028.

Thus we are pretty sure risk adjusted NPV is between 2.26 and 2.37 million. Since 95% of the time we are accurate within \$50,000 (which is 2% of sample mean) we feel comfortable that we have run enough iterations.

The **actual** discounted (at 10% rate) value of cash flows has much more variability than our confidence interval for risk-adjusted NPV would indicate. To show this look at the following histogram.

Figure 28.4



Note: If you are going to use distribution of NPV's as a tool to compare projects, then you must discount all the company's projects at the same rate (probably obtained from CAPM). Otherwise you will be double counting risk.

Tornado Graphs and Scenarios

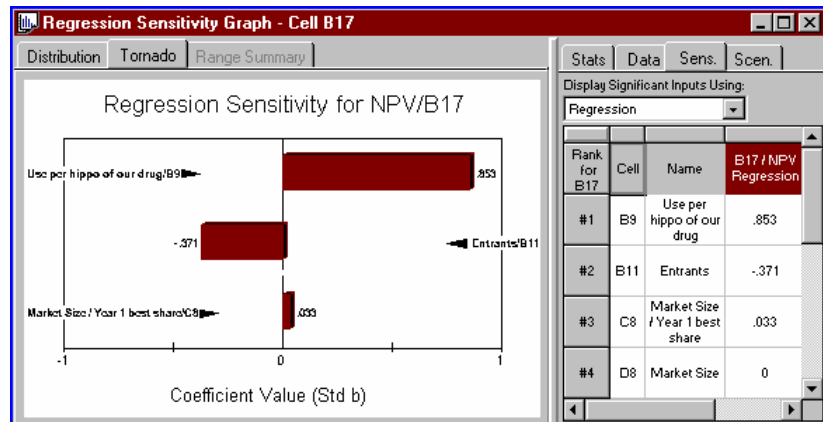
A natural question is what factors have the most influence on the success of the project? Does market growth matter more than the timing of the entrance of competitors? Using @RISK Tornado graphs and Scenario Analysis we can easily answer questions such as:

- What factors appear to have the most influence on the NPV earned by the drug?
- When NPV is in the top 10%ile of all possible NPV, what seems to be going on?

Solution - Part a

Here we utilize a **Tornado Graph**. Make sure that in “*Simulation Settings*” you have checked “*Collect Distribution Samples*”. Then right-click on NPV/B17 in the Explorer list and choose “*Tornado Graph.*” You have two options: A Regression Tornado Graph (see Figure 28.5) or a Correlation Tornado Graph (see Figure 28.6).

Figure 28.5

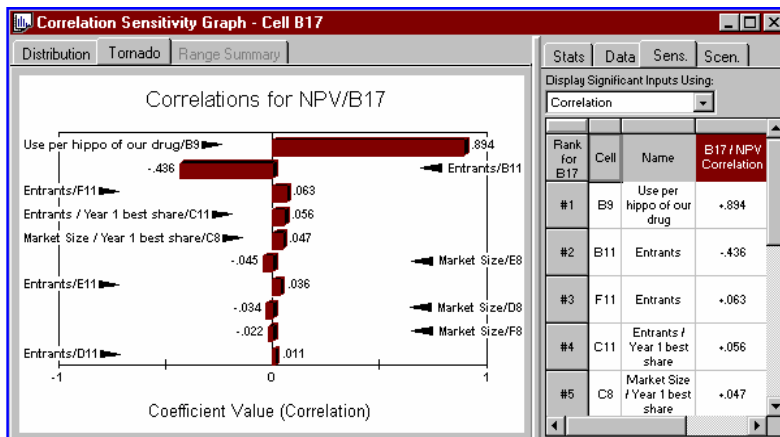


We find (*ceteris paribus*) from Regression Tornado Graph (obtained by selecting "*Regression*" under the "*Display Significant Inputs Using*" entry on the Sens. Tab on the graph) that

- A one standard deviation increase in Year 1 use increases NPV by .853 standard deviations.
- A one standard deviation increase in number of Year 1 entrants decreases NPV by .371 standard deviations.
- Not much else matters!

Basically, when running a Tornado Graph @RISK runs a regression where each iteration represents an observation. The dependent variable is the output cell (NPV) and the independent variables are each "random" @RISK function in the spreadsheet. Then the .853 coefficient for Year 1 Use is the standardized, or beta weight coefficient of Year 1 Use in this regression.

Figure 28.6



From Tornado Correlation Graph in Figure 28.6 (obtained by a change similar to the one above, except with *Correlation* instead of *Regression*) we find

- Year 1 use is most highly correlated (.89) with NPV
- Next is Year 1 Entrants (-.44)
- Rest of the random cells in the spreadsheet do not matter much!

These correlations are **rank correlations**; for example, for all iterations the values of Year 1 use are ranked, as are values of NPV. Then these **ranks** (not actual values) are correlated.

If you check *“Collect Distribution Samples”* under Simulation Settings you can obtain a **Scenario Analysis**. For a given scenario, such as all iterations where NPV is in top 10% of all iterations, the Scenario Analysis identifies random variables whose values differ significantly from their median values.¹

Solution - Part b

We find from Scenario Approach (see Figure 28.7) (click on the *“Insert New Scenarios Window”* icon or click on *“Insert”* and then *“Scenarios”*) that in the iterations yielding the top 10% of all NPV’s the following variables differ significantly from their overall medians:

- Year 1 Use (median is .596, 1.66 sigma above average)
- Year 2 Entrants (median is 0, 1.53 sigma below average)

To change the scenario settings just click on the *“Scenario=”* row in the Scenario Analysis box. Figure 28.7 contains a listing of three Scenario settings (the top 25%, the bottom 25%, and the top 10% of NPV’s) along with the random variables that differ significantly from their average values when the given scenario occurs. For example, for iterations in which the NPV is in the bottom 25% of all iterations, Year 1 Market Use averaged out to 13.9%.

Figure 28.7

| Scenario Analysis:4 | | | | | | | | | | | |
|---|------|---------------------------------|------------------|----------------------|---------------------|------------------|----------------------|---------------------|------------------|----------------------|---------------------|
| Display Groups of Input Values causing Output Scenarios, using: | | | | | | | | | | | |
| All - %Perc's, Actuals and Ratios | | | | | | | | | | | |
| | Cell | Name | B17 / NPV Actual | B17 / NPV Percentile | B17 / NPV Median/SD | B17 / NPV Actual | B17 / NPV Percentile | B17 / NPV Median/SD | B17 / NPV Actual | B17 / NPV Percentile | B17 / NPV Median/SD |
| | | Scenario= | >75% | >75% | >75% | <25% | <25% | <25% | >90% | >90% | >90% |
| #1 | C8 | Market Size / Year 1 best share | .. | .. | .. | .. | .. | .. | .. | .. | .. |
| #2 | D8 | Market Size | .. | .. | .. | .. | .. | .. | .. | .. | .. |
| #3 | E8 | Market Size | .. | .. | .. | .. | .. | .. | .. | .. | .. |
| #4 | F8 | Market Size | .. | .. | .. | .. | .. | .. | .. | .. | .. |
| #5 | B9 | Use per hippo of our drug | 0.55023 | 85.2% | +1.211851 | 0.3170602 | 13.88287% | -1.056794 | 0.5964633 | 92.98451% | +1.661682 |
| #6 | B11 | Entrants | .. | .. | .. | 2 | 93.6% | +1.177332 | .. | .. | .. |
| #7 | C11 | Entrants / Year 1 best share | .. | .. | .. | .. | .. | .. | .. | .. | .. |
| #8 | D11 | Entrants | 0 | 49.33334% | -1.53496 | .. | .. | .. | 0 | 49.33334% | -1.53496 |
| #9 | E11 | Entrants | .. | .. | .. | .. | .. | .. | 1 | 95.76923% | +1.687376 |
| #10 | F11 | Entrants | .. | .. | .. | .. | .. | .. | .. | .. | .. |

¹ @RISK will identify any random variable whose median value in iterations satisfying the scenario condition differs by more than .5 standard deviations from the median value of the random variable in all iterations.

Finding Value at Risk (VAR) of a Portfolio

VAR

(Chapter 45, Financial Models Using Simulation and Optimization)

Anybody who owns a portfolio of investments knows there is a great deal of uncertainty about the future worth of the portfolio. Recently the concept of **value at risk (VAR)** has been used to help describe a portfolio's uncertainty. Simply stated, **value at risk** of a portfolio at a future point in time is usually considered to be the fifth percentile of the loss in the portfolio's value at that point in time. In short, there is considered to be only one chance in 20 that the portfolio's loss will exceed the VAR. To illustrate the idea suppose a portfolio today is worth \$100. We simulate the portfolio's value one year from now and find there is a 5% chance that the portfolio's value will be \$80 or less. Then the portfolio's VAR is \$20 or 20%. The following example shows how @RISK can be used to measure VAR. The example also demonstrates how buying puts can greatly reduce the risk, or **hedge**, a long position in a stock.

Example 45.1

Let's suppose we own one share of Dell computer on June 30, 1998. The current price is \$94. From historical data (see Chapter 41) we have estimated that the growth of the price of Dell stock can be modeled as a Lognormal random variable with $\mu = 57\%$ and $\sigma = 55.7\%$. To hedge the risk involved in owning Dell we are considering buying (for \$5.25) a European put on Dell with exercise price \$80 and expiration date November 22, 1998. Here you will:

- a) Compute the VAR on November 22, 1998 if we own Dell computer and do not buy a put.
- b) Compute the VAR on November 22, 1998 if we own Dell computer and buy the put.

Solution

The key idea is to realize that in valuing the put we let Dell price grow at the risk-free rate, but when doing a VAR calculation we should let Dell price grow at the rate at which we expect it to grow. Our work is in file *var.xls*. See Figure 45.1.

Figure 45.1

| | A | B | C |
|----|----------------------------|-------------------|----------|
| 2 | Stress Testing Dell | Range Name | |
| 3 | Current price | S | \$ 94.00 |
| 4 | Put exercise price | x | \$ 80.00 |
| 5 | put duration | d | 0.39726 |
| 6 | risk free rate | r_ | 0.053541 |
| 7 | actual growth rate | g | 0.57 |
| 8 | volatility | v | 0.557 |
| 9 | put price | p | \$ 5.25 |
| 10 | | | |
| 11 | Dell price at expiration | 110.8422852 | |
| 12 | put value at expiration | 0 | |
| 13 | | | |
| 14 | % age Gain without put | 17.9% | |
| 15 | %age Gain with put | 11.7% | |

We have created range names as indicated in Figure 45.1.

Step by Step

Step 1: In cell B11 we generate Dell's price on November 22, 1998 with the formula

$$=S*EXP((g-0.5*v^2)*d+RISKNormal(0,1)*v*SQRT(d)).$$

Step 2: In cell B12 we compute the payments from the put at expiration with the formula

$$=If(B11>x,0,x-B11).$$

Step 3: The percentage gain on our portfolio if we just own Dell is given by

$$\frac{\text{Ending Dell Price} - \text{Beginning Dell Price}}{\text{Beginning Dell Price}}.$$

In B14 we compute the percentage gain on our portfolio if we do not buy a put with the formula

$$=(B11-S)/S.$$

Step 4: The percentage gain on our portfolio if we own Dell and a put is

$$\frac{\text{Ending Dell Price} + \text{Cash Flows from Put} - \text{Beginning Dell Price} - \text{Put Price}}{\text{Beginning Dell Price} + \text{Put Price}}$$

In cell B15 we compute the percentage gain on our portfolio if we buy the put with the formula

$$=((B12+B11)-(S+p))/(S+p).$$

Step 5: After selecting B14 and B15 as output cells and running 1600 iterations we obtained the @RISK output in Figure 45.2.

Figure 45.2

| | A | B | C | D |
|----|---|--------------------|----------------------|-----------------------|
| 5 | | Outputs | age Gain without put | age Gain with put / p |
| 6 | | Simulations | 1 | 1 |
| 7 | | Statistics / Cell | \$B\$14 | \$B\$15 |
| 8 | | Minimum | -0.642035902 | -0.193954661 |
| 9 | | Maximum | 3.177449226 | 2.956475973 |
| 10 | | Mean | 0.254243646 | 0.211036404 |
| 11 | | Standard Deviation | 0.454892722 | 0.403309986 |
| 12 | | Variance | 0.206927388 | 0.162658945 |
| 13 | | Skewness | 1.143237318 | 1.476801364 |
| 14 | | Kurtosis | 5.403351991 | 6.242985756 |
| 15 | | NumErrs | 0 | 0 |
| 16 | | Mode | 0.08372049 | -0.193954661 |
| 17 | | 5% | -0.339069603 | -0.193954661 |
| 18 | | 10% | -0.248756616 | -0.193954661 |
| 19 | | 15% | -0.180831468 | -0.193954661 |
| 20 | | 20% | -0.122562966 | -0.168636009 |
| 21 | | 25% | -0.069621287 | -0.11883527 |
| 22 | | 30% | -0.019603747 | -0.070426121 |
| 23 | | 35% | 0.029372538 | -0.025077894 |
| 24 | | 40% | 0.07877676 | 0.022255717 |
| 25 | | 45% | 0.12772527 | 0.068072297 |
| 26 | | 50% | 0.178600937 | 0.116256803 |
| 27 | | 55% | 0.232349445 | 0.167411774 |
| 28 | | 60% | 0.288710708 | 0.221145466 |
| 29 | | 65% | 0.349292601 | 0.277919441 |
| 30 | | 70% | 0.416892367 | 0.341943413 |
| 31 | | 75% | 0.493522018 | 0.414519578 |
| 32 | | 80% | 0.583930341 | 0.501486897 |
| 33 | | 85% | 0.696615509 | 0.608136356 |
| 34 | | 90% | 0.847094699 | 0.74938947 |
| 35 | | 95% | 1.100368295 | 0.98926574 |

We find our VAR if we do not buy the put to be 33.9% of our invested cash while if we buy the put our VAR drops to 19.4% of the invested cash. The reason for this is, of course, that if Dell stock drops below \$80, every one dollar decrease in the value of Dell is countered by a one dollar increase in the value of the put. Also note that if we do not buy the put, Dell (despite its high growth rate) might lose up to 64% of its value.

The following histograms give the distribution of the percentage gain on our portfolio with and without the put.

Figure 45.3

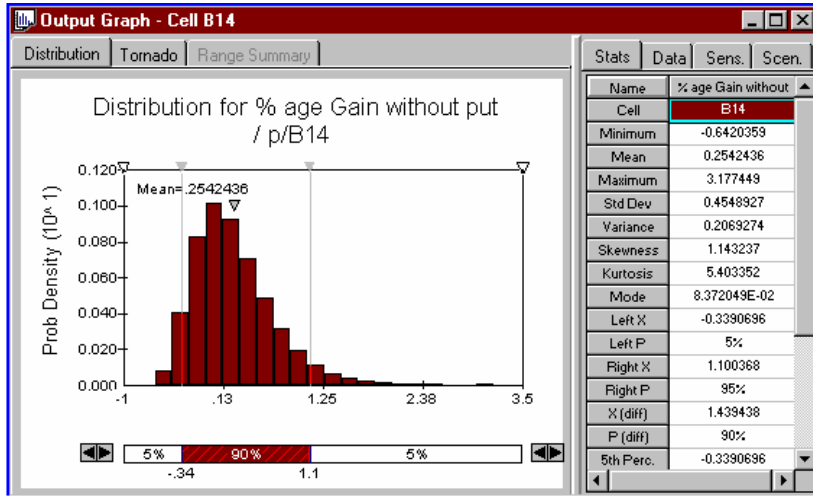
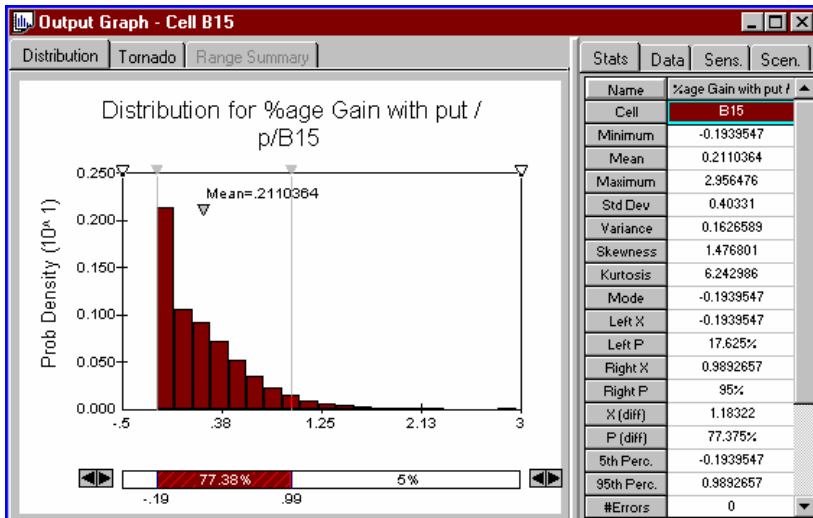


Figure 45.4



From Figures 45.3 and 45.4 we see that there is a much greater chance of a big loss if we do not buy the put. Note, however, that our average return without the put is 25.4% while our average return with the put is 21.1%. In effect, buying the put is a form of **portfolio insurance**, and we must pay for this insurance.

Simulating the NCAA Tournament

NCAA

(Chapter 62, Financial Models Using Simulation and Optimization)

The file *NCAA.xls* lets you play out the NCAA tournament as many times as you want. We factor in the abilities (through the SAGARIN ratings published in *USA Today*) of each team. Extensive data analysis has indicated that teams play on average to SAGARIN ratings and perform according to a standard deviation of 7 points about that level. For example, in 1997 SAGARIN rated NC a 94 and Fairfield a 70. Thus we would model NC's play by a $RISKNormal(94,7)$ and Fairfield by a $RISKNormal(70,7)$ and declare the team with the higher performance the winner. Our simulation of the 1996 NCAA tournament is in file *NCAA96.xls*

To begin we label the EAST teams 1-16 in the order they are listed in bracket. Then teams 17-32 are the SOUTHEAST, teams 33-48 the WEST and teams 49-64 the MIDWEST. It is important that we list things so that winner of 1 and 2 play winner of 3 and 4, etc.

Step by Step

Step 1: We enter the ratings, numerical codes, and team names in rows 2-4. We name the range A3:BL4 Ratings.

| | A | B | C | D |
|---|------|-----------|------|-----|
| 1 | 1 | 2 | 3 | 4 |
| 2 | UNC | Fairfield | Ind | Col |
| 3 | 1 | 2 | 3 | 4 |
| 4 | 94.4 | 70.3 | 85.3 | 82 |

| | A | B | C |
|---|------|---|------|
| 5 | East | | |
| 6 | 1 | | 2 |
| 7 | 94.4 | 1 | 70.3 |

Step 2: We model the UNC Fairfield game in A6:C7. In A7 we generate UNC's performance with the formula

$$=RISKNormal(HLOOKUP(A6,Ratings,2),7).$$

This looks up UNC's rating and generates a performance with that mean and a standard deviation of 7.

Similarly, in C7 we generate Fairfield's performance. In B7 we determine who wins the game with the formula

`=If(A7>C7,A6,C6).`

After "playing" the Colorado-Indiana game in E6:G7 (see Figure 62.1) we play the winners of these two games in A9:C10.

Figure 62.1

| | E | F | G |
|---|------|---|----|
| 6 | 3 | | 4 |
| 7 | 85.3 | 3 | 82 |

We ensure that the entry in A9 is the winner of UNC-Fairfield Game and the entry in C9 is winner of Indiana-Colorado game. Then in Row 10 we "play" this game. See Figure 62.2.

Figure 62.2

| | A | B | C |
|----|------|---|------|
| 9 | 1 | | 3 |
| 10 | 94.4 | 1 | 85.3 |

You can follow this logic down to Row 57. Here the Final Four begins! See Figure 62.3.

Figure 62.3

| | A | B | C | D | E | F | G | H |
|----|---------------|------|--------|------|---|---------|----|---------|
| 57 | | East | | West | | Midwest | | Mideast |
| 58 | | 1 | | 17 | | 33 | | 49 |
| 59 | | 94.4 | 17 | 97.9 | | 97.4 | 33 | 93.6 |
| 60 | | | | | | | | |
| 61 | | | | | | | | |
| 62 | Finals | | | | | | | |
| 63 | | 17 | | 33 | | | | |
| 64 | | 97.9 | 17 | 97.4 | | | | |
| 65 | | | Kansas | | | | | |

In 1997 East played West and Midwest played Mideast. Each year the final four matchups will change and you will need to adjust this part of spreadsheet. In C65 we print out the winner with the formula

`=HLOOKUP(C64,A1:BL2,2).`

This formula finds the Team Name corresponding to the code number of the winner. Hit the F9 key several times to see what happens.

We used cell C64 as our output cell and ran the tournament 5000 times. The teams having at least a 5% chance of winning were

- UNC: 13%
- Kansas: 26%
- Kentucky: 27%
- Duke: 8%
- Minnesota: 9%

Of course, Arizona won (we gave them a .0084 chance!). That's what makes sports great!

Remarks

Remember each year the Final Four brackets change. This will require you to rearrange the rows where the East, Midwest, Mideast and West regions are located.