

CRYSTAL BALL PROFESSIONAL INTRODUCTORY TUTORIAL

Lawrence I. Goldman

1515 Arapahoe Street
Suite 1311
Decisioneering, Inc.
Denver, CO 80202, U.S.A.

ABSTRACT

Crystal Ball® 2000 Professional Edition is a suite of easy-to-use Microsoft® Excel® add-in software that helps you analyze the risks and uncertainties associated with your spreadsheet models. The suite includes analysis tools for Monte Carlo simulation (Crystal Ball), time-series forecasting (CB Predictor), and optimization (OptQuest) as well as developer kits for building custom interfaces and processes. Spreadsheets alone are inadequate for assessing the probability of an event because they lack the ability to generate and analyze alternative scenarios in a sophisticated way. Spreadsheet add-ins such as Crystal Ball can provide this functionality and help modelers gain new insights into the mechanisms that drive their models and affect positive outcomes. This tutorial uses the example of an emerging media product to discuss how the analytical tools of Monte Carlo simulation and time-series forecasting can provide a greater understanding and quantification of the risks inherent in a spreadsheet-based business decision.

1 INTRODUCTION

1.1 Spreadsheet Models and Risk Analysis

Uncertainty about a situation can often indicate risk, which is the possibility of loss, damage, or any other undesirable event. Most people desire low risk, which would translate to a high probability of success, profit, or some form of gain. For example, if sales for next month are above a certain amount (a desirable event), then orders will reduce the inventory, and there will be a delay in shipping orders (an undesirable event). If a shipping delay means losing orders, then that possibility presents a risk.

When using spreadsheets, analysts traditionally input average or best-guess values for uncertain variables because Excel only allows them to enter a single value or formula in a cell. These “deterministic” models provide a single outcome upon which a business or technical decision is made. To capture uncertainty, analysts can perform

simple “what if analysis” or “scenario analysis” by manually changing model variables and analyzing their effect on the key outputs. This approach provides a range of possible outcomes but does not impart an understanding of the likelihood of any particular outcome.

More often, managers will perform a best case, worse case, and most likely case analysis where all uncertain variables are at their best, worse, or likely values at the same time. This approach has limited benefit because most real-world problems involving elements of uncertainty are too complex to solve analytically. There are simply too many combinations of input values to calculate every possible result.

1.2 Monte Carlo Simulation

One established solution to the limitations of spreadsheet risk analysis is Monte Carlo simulation. Since Excel alone does not have the ability to run and analyze simulations, modelers must rely on third-party programs like Crystal Ball that add in and expand the features of Excel. Crystal Ball adds two techniques to Excel: the replacement of single values with probability distributions and the random simulation of a model. The result is a probability-based spreadsheet with quantifiable outcomes, such as a 75% probability of staying under budget or a 90% certainty of 100 million barrels of oil within a geologic reservoir.

Monte Carlo simulation is a proven, efficient technique that only requires a random number table or a random number generator on a computer. A random number is a mathematically selected value that is generated to conform to a probability distribution. Monte Carlo simulation was named for Monte Carlo, Monaco, where the primary attractions are casinos containing games of chance, such as roulette wheels, dice, and slot machines, that exhibit random behavior.

The random behavior in games of chance is similar to how Monte Carlo simulation selects variable values at random to simulate a model. When you roll a die, you know that a 1, 2, 3, 4, 5, or 6 will come up, but you don't know

which for any particular roll. It's the same with the variables that have a known range of values but an uncertain value for any particular time or event (e.g. interest rates, staffing needs, stock prices, inventory, phone calls per minute). The multiple scenarios created through simulation can be analyzed to give more insight into the risks and mechanisms of the spreadsheet model. When used correctly, Monte Carlo simulation can provide valuable insights not available through deterministic models.

2 CRYSTAL BALL 2000 PROFESSIONAL EDITION

Crystal Ball 2000 Professional Edition is a desktop software suite that features spreadsheet-based analysis tools for Monte Carlo simulation (Crystal Ball), time-series forecasting (CB Predictor), and optimization (OptQuest). The suite also includes Crystal Ball and CB Predictor Developer Kits for building custom interfaces and processes using Visual Basic for Applications (VBA).

2.1 Crystal Ball

Crystal Ball is an easy-to-use add-in designed to help all levels of Excel spreadsheet modelers perform Monte Carlo simulation. Crystal Ball allows users to define probability distributions on uncertain model variables, and then uses simulation to generate random values from within the defined probability ranges. Spreadsheet modelers can create and analyze thousands of alternate scenarios and quantify their level of risk for any given scenario. Crystal Ball can be applied to existing or new spreadsheet models, and the enhancements provided by Crystal Ball do not alter the formulae or functions of the original spreadsheet. Crystal Ball also includes the CB tools, a series of seven wizard-driven add-ins that can assist in setting up and analyzing models. These tools include Bootstrap, Tornado Analysis, Scenario Analysis, 2D simulation, and Batch Fit.

2.2 CB Predictor

CB Predictor is a wizard-driven Excel add-in that guides you through the process of time-series forecasting. CB Predictor analyzes your data series and uses the level, trend, seasonality, and error of the data to project the series into the future. CB Predictor applies eight separate seasonal and nonseasonal time-series forecasting methodologies to data and can apply multiple linear regression in cases where the primary data series is dependent on other independent data series. The output forecasts can be defined as Crystal Ball assumptions and used in Monte Carlo simulation.

2.3 OptQuest

OptQuest is a global optimization add-in created specifically for Crystal Ball. The optimizer enhances Crystal Ball by automatically searching for and finding optimal solutions to simulation models. OptQuest uses a mixture of techniques including scatter search and advanced tabu search to find the right combination of decision variables for the best possible results. As the program runs, adaptive and neural network technologies help it learn from past optimizations so that it achieves better results in less time. A setup wizard helps users define constraints, objectives, and requirements, and even to set an efficient frontier option. OptQuest will not be applied in this tutorial.

2.4 Developer Kits

The Crystal Ball and CB Predictor Developer Kits allow users to completely automate and control Crystal Ball simulations and CB Predictor forecasts from within a VBA program or any other language outside of Excel supporting OLE 2 automation. The kits consist of libraries of macro commands and functions that unlock the programmability of Crystal Ball and CB Predictor. With the kits, users can automatically run multiple simulations to test different sets of assumptions, integrate Crystal Ball with other software tools, create turnkey applications that shield users from program intricacies, and even build custom reports or automate post-simulation analysis. The Developer Kits will not be applied in this tutorial.

3 COLORADO CABLE EXAMPLE

This tutorial reviews how simulation and forecasting can improve a business forecast. In this example, you will forecast from a series of historical data, define probability distributions, run a Monte Carlo simulation, and analyze and report the simulation results. The model described here is available for free on the Decisioneering Web site (see the Appendix). At the end of the example, you will have a breakdown of the key success factors, an analysis of the bottom-line performance of the new product, and a report with a clear presentation of your findings.

Colorado Cable, a local cable provider, is evaluating a new technology known as Interactive TV (ITV). ITV will provide content – movies, sports games, and news – on demand. Colorado Cable believes that the local audience will embrace this type of service, but the company is concerned that there might be some unanticipated risks because of the down economy.

Management has requested a forecast model for ITV from its introduction in 2003 through to 2008. They want to better understand the sales and marketing potential of this new technology before making a hefty investment. Does ITV have enough potential to merit bringing it to market? What is the Net Present Value (NPV) over six-years?

3.1 Creating the Model, Addressing Uncertainty

Your first task is to create a valid and verifiable deterministic model that represents the most likely ITV sales scenario. You carefully gather data and construct the model, relying on colleagues to help create a realistic base-case model of the market potential of ITV. Your model (Figure 1) looks at the market shares and sizes of Colorado Cable's three existing products – cable, satellite dish, and broadcast TV – and adds what you expect for ITV.

You forecast the revenues and expenses of the ITV technology and determine the NPV for a six-year period. Based on your deterministic estimations, you expect a \$51 million dollar NPV (with a 10% discount rate). But how likely is this outcome? What is the probability you will make more than \$51 million? Break even? Lose money? CB Predictor and Crystal Ball will help you to answer these questions.

	2003	2004	2005	2006	2007	2008
Households with TV	3,200,000	3,250,000	3,300,000	3,350,000	3,400,000	3,450,000
Market Shares						
Cable	78.0%	72.0%	68.0%	62.0%	56.0%	50.0%
Satellite Dish	12.0%	14.0%	16.0%	18.0%	20.0%	22.0%
Interactive TV	2.0%	7.0%	10.0%	15.0%	20.0%	25.0%
Broadcast TV	8.0%	7.0%	6.0%	5.0%	4.0%	3.0%
TOTAL	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Market Size						
Cable	2,496,000	2,340,000	2,244,000	2,077,000	1,904,000	1,725,000
Satellite Dish	384,000	455,000	528,000	603,000	680,000	759,000
Interactive TV	64,000	227,500	330,000	502,500	680,000	862,500
Broadcast TV	256,000	237,500	198,000	167,500	136,000	103,500
TOTAL	3,200,000	3,250,000	3,300,000	3,350,000	3,400,000	3,450,000
Revenues						
Annual Revenue per Household	\$ 120	\$ 125	\$ 130	\$ 135	\$ 140	\$ 145
Total Revenues	\$ 7,680,000	\$ 28,437,500	\$ 42,900,000	\$ 67,837,500	\$ 95,200,000	\$ 125,062,500
Expenses						
Initial Investment	\$ 100,000,000	\$ 21,000,000	\$ 22,000,000	\$ 23,000,000	\$ 24,000,000	\$ 25,000,000
Annual Operating Costs	\$ 20,000,000	\$ 21,000,000	\$ 22,000,000	\$ 23,000,000	\$ 24,000,000	\$ 25,000,000
Total Expenses	\$ 120,000,000	\$ 21,000,000	\$ 22,000,000	\$ 23,000,000	\$ 24,000,000	\$ 25,000,000
Net Profit	\$ (112,320,000)	\$ 7,437,500	\$ 20,900,000	\$ 44,837,500	\$ 71,200,000	\$ 100,062,500
NPV (10% discount rate)	\$ 51,056,363					

Figure 1: Colorado Cable Spreadsheet Model

3.2 Launching Crystal Ball

After your model is designed and tested, you can launch Crystal Ball and Excel through the Windows Start menu. Crystal Ball adds a new tool bar and three new menus to Excel. The tool bar (Figure 2) has been designed from left to right to follow the modeling process from setup to simulation to the analysis and presentation of results. The final button is for online help.

The three Crystal Ball menus contain all of the functions on the tool bar plus a few additional features. The Cell menu contains all of the model setup functions, and

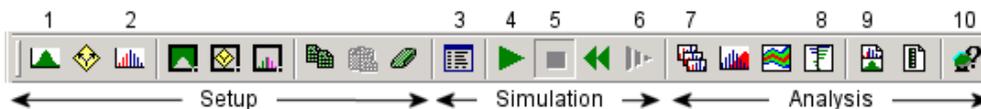


Figure 2: Tool Bar Buttons: (1) Define Assumption (2) Define Forecast (3) Run Preferences (4) Run (5) Stop (6) Single Step (7) Forecast Windows (8) Sensitivity Analysis (9) Create Report (10) Help

the *Run* menu lists all of the simulation and analysis functions. The final menu, *CBTools*, opens other tools in the Professional Edition suite.

3.3 Forecasting Households from Historical Data

The numbers of future households in your area, those that might want services from Colorado Cable, are uncertain past 2002. Based on twenty years of historical data, you have estimated a constant growth rate of 50,000 new households with TV per year, although you know that the actual rate of increase is somewhat less predictable. Because this data has a time component, you can use CB Predictor to replace your estimated values with more accurate time-series forecasts.

To run CB Predictor, select any cell within the data series you plan to forecast and open the program from the *CBTools* menu. The first four steps help you to define, organize, and view the data. An autocorrelation feature determines if any seasonality (shown as lags) is present in the data. In this annual data, there is no seasonality, just an increase in number of households.

The Method Gallery allows you to select any or all of the eight available time-series methods for comparison to your data. They are divided into Nonseasonal and Seasonal, and you can learn more about each by double-clicking on its gallery image. Through the gallery you can also override the default error statistics or parameters for any methodology. Because you are forecasting a single, independent data series, you do not apply multiple linear regression.

You select a forecast of five periods (years) and a desired confidence interval of 5% and 95% around the value. CB Predictor will let you preview your forecast at any time throughout the setup procedure (Figure 3).

For the household data, CB Predictor has selected Double Exponential Smoothing as the best forecasting method. CB Predictor examined the historical data, fit each nonseasonal method to it, selected the method with the best fit and smallest error, forecast the trend five years into the future, and calculated the 90% confidence interval.

Following the preview, you select the output charts and data format. CB Predictor can create Crystal Ball normal probability distributions as forecasts. After you run the forecast, you can copy and paste these distributions into the household estimates variables of the spreadsheet. As a result of the more accurate time-series forecasts, the base-case NPV has improved to almost \$53 million.

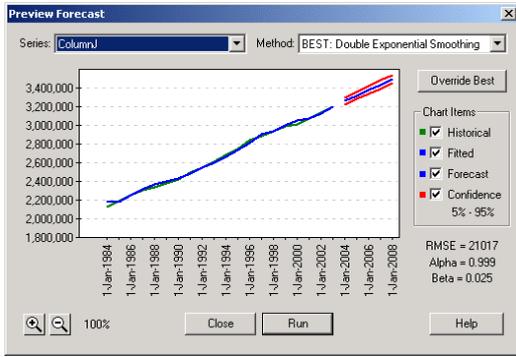


Figure 3: Time-Series Forecast of Available Households

3.4 Defining Crystal Ball Assumptions

In Crystal Ball, probability distributions are referred to as “assumptions” and are the basic inputs you use to define the uncertainty in any model variable. The best way to identify which variables to convert from single values to probability distributions is to decide which variables are known and which are “soft” and uncertain. In this model, you know that Market Share, Initial Investment, and Annual Operating Costs are all prime candidates for Crystal Ball assumptions.

To define an assumption, select a spreadsheet variable and click on the Define Assumption button on the tool bar. In the Distribution Gallery (Figure 4), Crystal Ball offers sixteen pre-defined distributions and one customizable distribution. The four most commonly used assumptions are the Normal, Triangular, Uniform, and Lognormal.

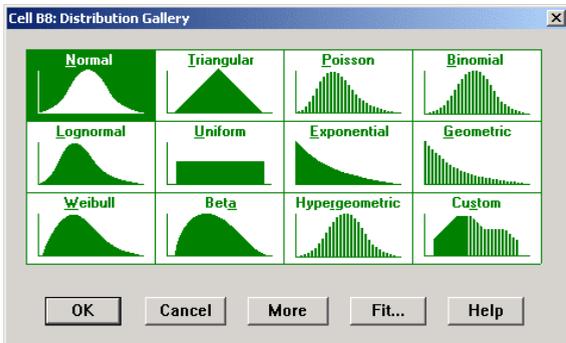


Figure 4: Distribution Gallery

Crystal Ball can also fit continuous distributions to available data and can define correlation coefficients between related pairs of assumptions. In this example, you define assumptions based on your own knowledge and intuition. The more you know about a variable, the more accurate the distribution.

3.4.1 Market Share

For Market Share for 2003, you have estimated 2%, although you know it could be as low as 0% or as high as

3%. Click on the Define Assumptions button, and select the Triangular distribution. Triangular distributions are excellent for describing basic situations where the minimum, likeliest, and maximum values are known. Figure 5 shows a Triangular assumption for the 2003 Market Share of ITV. You can also define this distribution using percentiles should you receive the data in that format.

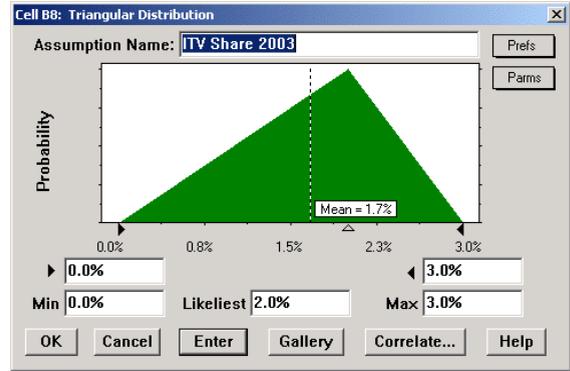


Figure 5: Market Share of ITV as a Triangular Distribution

You also define the 2004-2008 estimates of ITV Market Share as Triangular distributions, with the data in Table 1. While you expect the market share to grow with each year, you may not see the steady growth predicted in the base-case model. Assumptions let you build in the compounding uncertainty as time passes.

Table 1: ITV Market Share Uncertainty for 2004-2008

Year	Minimum	Likeliest	Maximum
2004	0%	7%	10%
2005	0%	10%	15%
2006	0%	15%	20%
2007	0%	20%	30%
2008	0%	25%	35%

3.4.2 Initial Investment

The Normal distribution, often referred to as the bell curve, describes naturally occurring quantities, quantities that have a most likely value, are symmetrical, and are more likely to have values near to the mean. In your base case, you estimated an Initial Investment of \$100 million. You know that while the true investment amount will be close to the mean value of \$100 million, it could be as much as \$130 million or as little as \$70 million. A Normal distribution with a mean of \$100 million and a standard deviation of \$10 million is a more precise description of this variable (Figure 6).

3.4.3 Annual Operating Costs

The final set of unknown variables is the operating costs for each of the six years. In the base case, you have antici-

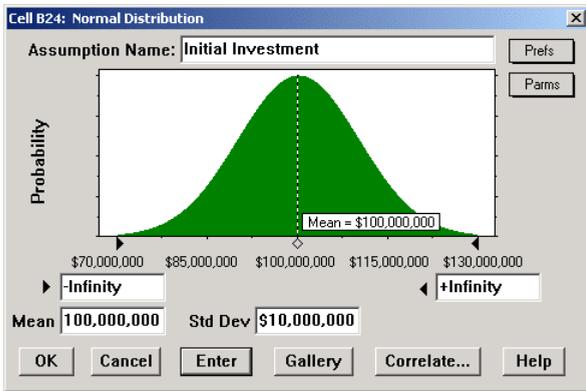


Figure 6: Initial Investment as a Normal Distribution

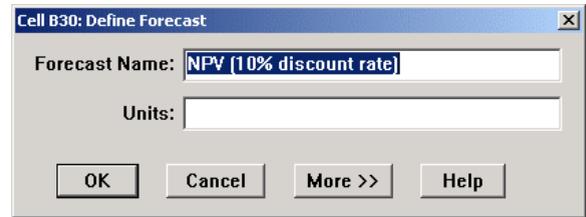


Figure 8: Defining a Forecast Cell

puted a \$1 million growth in expenses each year. With Crystal Ball, you define these variables as Normal distributions with a standard deviation of \$1 million. For each mean value, you can use cell referencing to include data from your spreadsheet (Figure 7).

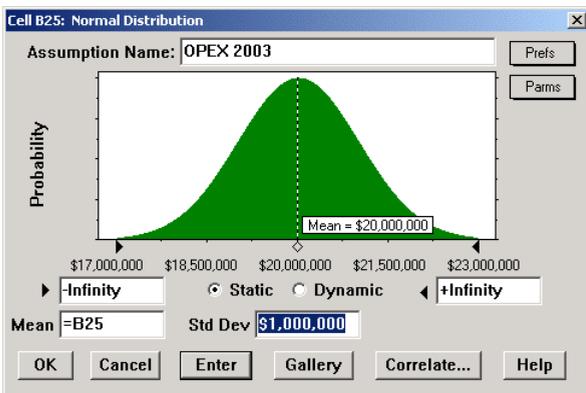


Figure 7: Operating Expenses for 2003 as a Normal Distribution with Cell Referencing

3.5 Defining a Forecast

After defining all of your inputs (assumptions), you need to define the output, termed a “forecast” in Crystal Ball. A forecast is any formula cell that you assign Crystal Ball to track during a simulation. As with assumptions, there is no limit to the number of forecasts you can define, although more memory is required to store the simulated values. This model has only one forecast, the NPV calculation. You select this cell, click on the Define Forecast button to open the Define Forecast dialog (Figure 8), enter a unique name, and click on OK.

3.6 Running a Monte Carlo Simulation

3.6.1 Setting the Run Preferences

Crystal Ball gives you full access to the simulation controls. These settings include number of trials to run, what sampling method to use, where and when during the simulation cycle to run macros, ways to increase the speed of simulations, and Precision Control, a method of confidence testing that lets you achieve a desired level of accuracy in your forecasts based on simulation statistics.

There are two sampling methods: Monte Carlo and Latin Hypercube. With Monte Carlo, the generated random numbers are completely independent of one another. With Latin Hypercube sampling, Crystal Ball divides the assumption’s probability distribution into segments of equal probability that you define, and during a simulation generates random numbers for each segment. In this example, you use Monte Carlo simulation for 2000 trials.

3.6.2 Testing and Running the Simulation

Before running the simulation, you can check that the model is functioning properly by clicking on the Single Step button on the tool bar. Each single step is a random trial in Crystal Ball. In one step, Crystal Ball generates a random number for each assumption, and Excel automatically recalculates the model. You may discover calculation errors with this test as the model recalculates for alternate scenarios.

During a simulation, Crystal Ball runs as many trials as you request and saves the forecast values for later analysis. To run the simulation, click the Start Simulation button on the tool bar. As the simulation begins and the NPV forecast (or frequency) chart builds, you can watch the spreadsheet values change as Crystal Ball inserts randomly generated values (from the probability distributions) into the assumption cells.

3.7 The Forecast Chart

When the 2,000 trials are completed, you can analyze these results using the forecast chart, an interactive histogram that contains all of the statistics for the 2000 trials. How likely was your original estimate of roughly \$51 million? Based on the simulation results (Figure 9), you only have about an 11% certainty of achieving or exceeding your

original NPV value. Your probability of achieving \$53 million, which you had calculated after forecasting the historical household data, is even less. You can also determine that you have about a 60% chance of breaking even by entering 0 in the left range field of the chart.

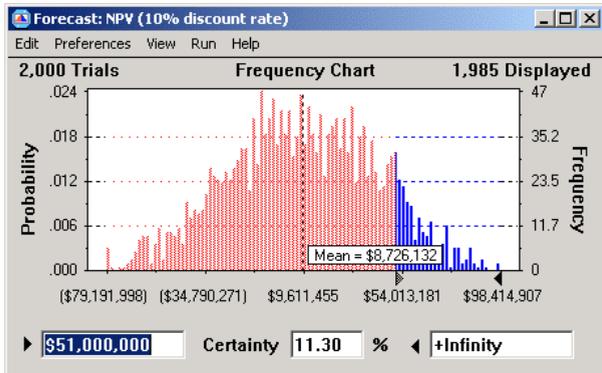


Figure 9: Forecast Chart for NPV

The Statistics view (Figure 10) can help you gain additional insight. The results appear to be close to normally distributed, as shown by the kurtosis and skewness values. Your standard deviation is over \$34 million, and the range of results spans \$210 million. Why is the range of values so wide? Crystal Ball's sensitivity analysis can point to the key factors responsible for the variation in the results.

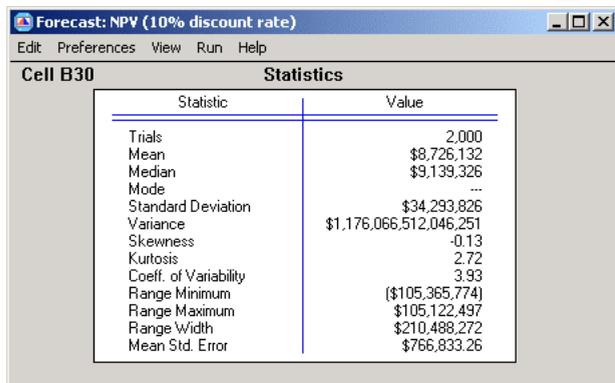


Figure 10: Statistics for NPV Forecast

3.8 Sensitivity Analysis

As the simulation runs, sensitivity analysis uses Rank Correlation to dynamically calculate the relationships among the 13 assumptions and the forecast. The sensitivity chart displays these relationships as correlation coefficients or as percentage values. The top listed assumptions have the greatest effect on the NPV, and the direction of the bar indicates a direct or inverse relationship.

In this example, the market share variables for 2005 to 2008 have the greatest impact on NPV (Figure 11). In the final years of this study, you selected larger market shares

(Table 1). Were these assumptions reasonable? What will happen if you reduce the ranges of these distributions and re-run your simulation? As you begin to ask these questions, you can see how simulation and sensitivity analysis help you focus on the most important drivers in your model. Simulation modeling is an iterative process, and though your initial results may indicate ITV will not perform to expectations, knowing what drives the model can help you to improve it and communicate the inherent risks to others.

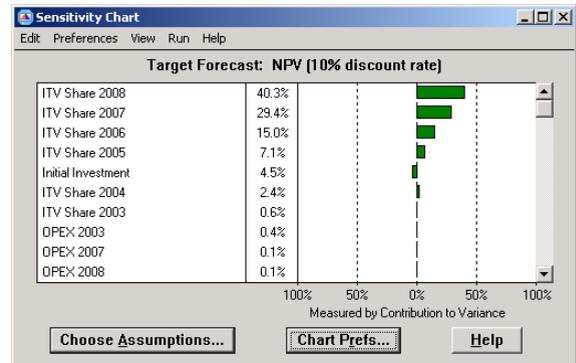


Figure 11: Sensitivity Chart for NPV

3.9 Saving and Documenting Your Results

When you are satisfied with your model and simulation results, you have two ways to document your efforts. First, you can create a printable report of the simulation results, charts, and variables. The report file, which can include Excel charts, can be saved as a new worksheet or as a worksheet in the same workbook as your model. You can also extract the raw forecast data such as forecast values, statistics, percentiles, or frequency counts to a new worksheet or as a worksheet in the same workbook as your model.

4 CONCLUSION

Performing simulation analysis for future outcomes can help spreadsheet modelers to reduce their risks and dramatically improve the quality of their decisions. Crystal Ball 2000 Professional Edition overcomes the limitations of spreadsheets by enhancing Excel with user-friendly tools for Monte Carlo simulation, global optimization, and time-series forecasting. By moving to a probabilistic approach to spreadsheet forecasting, analysts can better quantify the risks inherent in their models and gain insights not available through traditional deterministic approaches.

APPENDIX: CRYSTAL BALL RESOURCES

The model described in this tutorial is available for free on the Decisioneering web site's example model library <www.crystalball.com/models/model_index

[.html](#)>. Other free resources on the site include white papers, articles, tutorials, case studies, and Internet links. Free downloadable trial versions of the software and tutorials are available to all site visitors.

AUTHOR BIOGRAPHY

LAWRENCE GOLDMAN is a Product Marketing Manager at Decisioneering in Denver, CO. He joined the company in June 1997 as a program manager with the product development group and has held positions as Webmaster, Web Producer, and Trainer. His current focus is educational product materials for Crystal Ball. Prior to working for Decisioneering, Lawrence was employed as a trainer and technical support assistant for Techbase, a company specializing in mining and engineering database software. He has an M.S. in Geology from the University of Cincinnati and a B.A. in Geology from Cornell University. His email address and web address are lgoldman@decisioneering.com and www.crystalball.com.