Crystal Ball®

2000

User Manual

Written, designed, and published in the United States of America.

To purchase additional copies of this document, contact the Technical Services or Sales Department at the address below:

Decisioneering, Inc.
1515 Arapahoe St., Suite 1311
Denver, Colorado, USA 80202

Phone: (303) 534-1515 Toll-free sales: 1-800-289-2550 Fax: (303) 534-4818

Contents

Contents

Contents

Contents

Contents

Chapter 6: Crystal Ball Tools

Appendix A: Error Recovery

Appendix B: Equations and Methods

Contents

Contents

Welcome to Crystal Ball® 2000

Welcome to Crystal Ball 2000!

Crystal Ball is a user-friendly, graphically oriented forecasting and risk analysis program that takes the uncertainty out of decision-making.

Through the power of simulation, Crystal Ball becomes an effective tool in the hands of the decision-maker. You can answer questions such as, “Will we stay under budget if we build this facility?” or, “What are the chances this project will finish on time?” or, “How likely are we to achieve this level of profitability?” With Crystal Ball, you will become a more confident, efficient, and accurate decision-maker.

Crystal Ball is easy to learn and easy to use. Unlike other forecasting and risk analysis programs, you do not have to learn unfamiliar formats or special modeling languages. To get started, all you have to do is create a spreadsheet. From there, this manual guides you step by step, explaining Crystal Ball terms, procedures, and results.

And you do get results from Crystal Ball. Through a technique known as Monte Carlo simulation, Crystal Ball forecasts the entire range of results possible for a given situation. It also shows you confidence levels, so you will know the likelihood of any specific event taking place.

Who this program is for

Crystal Ball is for the decision-maker—from the businessperson analyzing the potential for new markets to the scientist evaluating experiments and hypotheses. Crystal Ball is easy to learn and easy to use, and has been developed with a wide range of spreadsheet uses and users in mind.

You don’t need highly advanced statistical or computer knowledge to use Crystal Ball to its full potential. All you need is a basic working knowledge of your personal computer and the ability to create a spreadsheet model.

What you will need

To run Crystal Ball, you will need the following:

- Microsoft Excel 95 (or later)
- Windows 95, Windows 98, Windows 2000, or Windows NT 4.0
- Personal computer with Pentium microprocessor and at least 16 MB RAM
- Hard disk drive with at least 16 MB free
- 2X CD-ROM drive or faster
- VGA or SVGA graphics card or compatible video graphics adapter and monitor

Recommended system requirements

- Pentium II, 200 MHz or faster
- 24 MB RAM or more

How this manual is organized

This manual includes the following:

- **Chapter 1 - "Getting Started With Crystal Ball"**

  This chapter contains two tutorials designed to give you a quick overview of Crystal Ball's features and to show you how to use it. Read this chapter if you need a basic understanding of Crystal Ball.

- **Chapter 2 - "Understanding the Terminology"**

  This chapter contains a thorough description of risk analysis, the seventeen probability distributions, and other statistical terms used in Crystal Ball. Read this chapter carefully if your statistical knowledge is limited or if you need a review.

- **Chapter 3 - "Setting Up and Running a Simulation"**

  This chapter provides step-by-step instructions for setting up and running a simulation in Crystal Ball.

- **Chapter 4 - "Interpreting the Results"**

  This chapter explains how to use Crystal Ball's powerful features to interpret the results of a simulation.

- **Chapter 5 - "Maximizing Your Use of Crystal Ball"**

  This chapter describes different aspects that enhance the performance of the program's features.

- **Chapter 6 - "Crystal Ball Tools"**

  This chapter describes tools that extend the functionality of Crystal Ball, such as the Tornado Chart and Decision Table tools.

- **Appendices**

  - A - "Error Recovery"

    A list of the most important Crystal Ball error messages with directions for error recovery.

  - B- "Equations and Methods"

    A list of the mathematical formulas used in calculating the distributions and the descriptive statistics. It also describes the type of random number generator used in Crystal Ball. This appendix is designed for the statistically sophisticated user.

  - C- "Keyboard Commands"

    A list of the commands you can execute directly from the keyboard.

  - D- "Index of Commands"

    A brief description of Crystal Ball's pull-down menus.

  - E- "Default Names and Distribution Parameters"

    A brief description of Crystal Ball's default values.

- **Bibliography**

  A list of related publications, including statistics textbooks.

- **Glossary**

  A compilation of terms specific to Crystal Ball and other statistical terms used in this manual.

- **Index**

  An alphabetical list of subjects and corresponding page numbers.

What's new in Crystal Ball 2000

If you are already familiar with Crystal Ball, read this section for an overview of Crystal Ball 2000's new features.

Microsoft Office 2000 compatibility

This version of Crystal Ball is fully functional with Excel 2000 included as part of the Office 2000 suite of products.

If you have configured Crystal Ball to start automatically with a previous version Excel, and then you installed Office 2000 over the previous version, you must reset this option in the Add-ins dialog. For more information on doing this, see the FAQ on our website at:

www.decisioneering.com

New tools and menu

Crystal Ball has additional tools that enhance and extend the functionality of Crystal Ball. These tools are in the CBTools menu and include:

- Batch Fit tool
- Bootstrap tool
- Correlation Matrix tool
- Decision Table tool
- Tornado Chart tool
- Two-dimensional Simulation tool

Precision control

You can now specify a level of precision for your simulation results, minimizing the number of trials required to achieve a specified precision with a particular confidence level. When the selected statistics reach the specified precision and confidence level, the simulation stops.

Decision variables

There is now a new Crystal Ball data type: decision variable. Decision variables define variables within your control, and are used with the Decision Table tool and OptQuest (OptQuest is only available with Crystal Ball 2000, Professional Edition).

Enhanced cell preferences

There are two additional cell preferences that let you display:

- The mean of the assumption distribution in each assumption cell.
- The midpoint of the decision variable range in each decision variable cell

Additional alternate parameters

You can now define parameters for most continuous distributions using either the standard parameter set or one of several alternate sets, such as percentiles (e.g., the 10th and 90th percentiles), under the Parms button on each Assumption dialog.

New triangular alternate parameter set

You can now use three percentiles to define an alternate parameter set for a triangular distribution. This lets you define asymmetric triangular distributions using three percentiles.

Assumption preference menu

All the distribution dialogs have a Prefs menu to let you set preferences for the distribution dialog, such as whether to:

- View the graph in cumulative format (CDF)
- Show mouse coordinates
- Show the mean line on the chart

Reset single step

You can now reset after using the single step command.

Update simulation results

You can specify how often to update the forecast windows and check precision control during simulations. During simulations, the forecast results repeatedly update after the specified number of trials. If you use the Precision Control feature, after the specified number of trials, the Precision Control feature repeatedly checks to see if the needed precision is met so that it can stop the simulation.

Reversing percentiles

This option reverses whether the percentiles that appear in the Define Assumption dialogs and the forecast results windows represent the percentage of the distribution above or below the value. Below is standard, but you can reverse the percentile representations by changing this option in the Statistics Preferences dialog.

New default location

Crystal Ball used to be installed under the Excel installation folder by default. The default installation is now directly under the Program Files folder in a folder called Crystal Ball.

Also, because the default location is in a different folder, the installer will automatically uninstall previous versions of Crystal Ball.

Additional resources

Decisioneering, Inc. offers these additional resources to increase the effectiveness with which you can use our product.

Technical support

Decisioneering, Inc. offers free technical support for 90 days. If you have a technical support question or would like to comment on Crystal Ball, there are a number of ways to reach technical support described in the README file in your Crystal Ball installation folder.

Consulting referral service

Decisioneering, Inc. provides referrals to individuals and companies alike. The primary focus of this service is to provide a clearinghouse for consultants in specific industries who can provide specialized services to the Crystal Ball user community.

If you wish to learn more about this referral service, call 800-289-2550 Monday through Friday, between 8:00 A.M. and 5:00 P.M. Mountain Standard Time or see our website at: www.decisioneering.com.

Conventions used in this manual

This manual uses the following conventions:

Text separated by > symbols means you select menu options in the sequence shown, starting from the left. The following example means that you select the Exit option from the File menu:

1. Select File > Exit.

Steps with attached icons mean you can click on the icon instead of manually selecting the menu options in the text. For example:

2. Select Cell > Define Assumption.

Notes provide additional information, expanding on the text. There are four categories of notes:

Crystal Ball Note: *Notes that provide additional directions or information about using Crystal Ball.*

Excel Note: *Notes that provide additional information about using the program with Microsoft Excel.*

OptQuest Note: *Notes that provide additional directions or information about using OptQuest.*

Statistical Note: *Notes that provide additional information about statistics.*

Conventions used in this manual

Words in angle brackets represent keys on the keyboard, such as **<Ctrl>** or **<Return>**. Press the key; *don't* type the letters, words, or brackets.

Sometimes you have to press two or more keys simultaneously. For example, **<Ctrl>-c** means that you hold down the **<Ctrl>** key and type **c**. Capitalization is important; **<Ctrl>-c** and **<Ctrl>-C** are two different key sequences.

A key sequence without hyphens means you type the sequence in the order shown but not simultaneously. For example, **<Ctrl>-q N** means that you press the **<Ctrl>** key and type **q** simultaneously, and then type **N**.

Screen capture notes

All the screen captures in this document were taken in Excel 97 for Windows NT, using a seed value of 999.

Due to round-off differences between various system configurations, you might obtain slightly different calculated results than those shown in the examples.

Chapter 1

Getting Started With Crystal Ball

In this chapter

- What Crystal Ball Does
- How Crystal Ball uses Monte Carlo Simulation
- "Futura Apartments" spreadsheet tutorial
- "Vision Research" spreadsheet tutorial
- Defining assumptions
- Selecting and defining distributions
- Defining forecasts
- Running a simulation
- Interpreting the results

In this chapter are two tutorials, one short, one long, providing an overview of Crystal Ball's features. The first tutorial, the "Futura Apartments" spreadsheet, simulates profit/loss projections from apartment rentals. This tutorial is ready to run so you can quickly see how Crystal Ball works. If you work regularly with statistics and forecasting techniques, this might be all the introduction you need before running your own spreadsheets with Crystal Ball.

The second tutorial, the "Vision Research" spreadsheet, gives you a chance to enter data and set up a complete simulation for a major corporate expenditure decision. As you work through the second tutorial, do not worry about making mistakes; recovery is as easy as backing up and repeating the steps. If you need additional help, see "Error Recovery" on page 313, or the Crystal Ball on-line help.

Now, spend a few moments learning how Crystal Ball can help you make better decisions under conditions of uncertainty.

What Crystal Ball does

Glossary Term:
spreadsheet model—
Any spreadsheet that represents an actual or hypothetical system or set of relationships.

Crystal Ball extends the forecasting capability of your **spreadsheet model** and provides the information you need to become a more accurate, efficient, and confident decision maker. As a spreadsheet user, you know that spreadsheets have two major limitations:

- You can change only one spreadsheet cell at a time. As a result, exploring the entire range of possible outcomes is next to impossible; you cannot realistically determine the amount of risk that is impacting your bottom line.
- "What-if" analysis always results in single-point estimates which do not indicate the likelihood of achieving any particular outcome. While single-point estimates might tell you what is *possible*, they do not tell you what is *probable*.

Crystal Ball overcomes both of these limitations:

Glossary Term:
assumption—
An estimated value or input to a spreadsheet model.

Glossary Term:
Monte Carlo simulation—
A system which uses random numbers to measure the effects of uncertainty in a spreadsheet model.

- You can describe a range of possible values for each uncertain cell in your spreadsheet. Everything you know about each **assumption** is expressed all at once.
- Using a process called **Monte Carlo simulation**, Crystal Ball displays results in a forecast chart that shows the entire range of possible outcomes and the likelihood of achieving each of them.

Starting the first tutorial

The best way to quickly understand this process is to start Crystal Ball and work on the first tutorial: The Futura Apartments spreadsheet.

1. Start Crystal Ball.

Crystal Ball Note: *For directions on starting Crystal Ball, see* *"Starting Crystal Ball manually" on page 129.*

2. **Open the Futura Apartments spreadsheet file from the Crystal Ball Examples folder. You can find this folder by selecting Start > Programs > Crystal Ball > Examples.**

   The Futura Apartments spreadsheet appears, as in Figure 1.1.

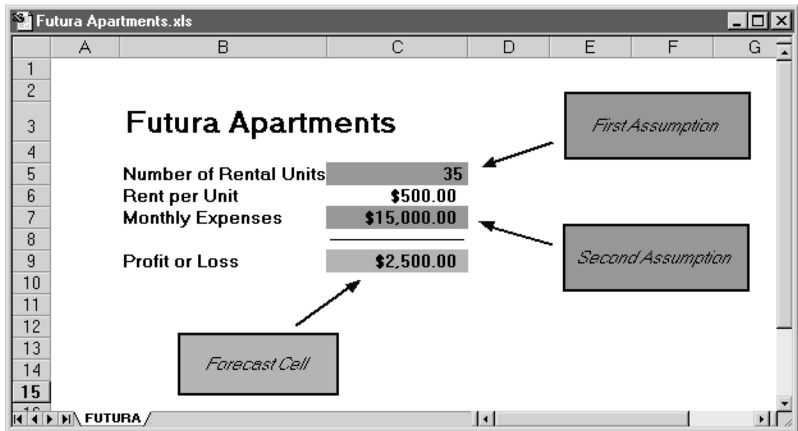


Figure 1.1 Futura Apartments spreadsheet

In this example, you are a potential purchaser of the Futura Apartments complex. You have researched the situation and created the above spreadsheet to help you make a knowledgeable decision. Your work has led you to make the following assumptions:

- $500 per month is the going rent for the area.
- The number of units rented during any given month will be somewhere between 30 and 40.
- Operating costs will average around $15,000 per month for the entire complex, but might vary slightly from month to month.

Based on these assumptions, you want to know how profitable the complex will be for various combinations of rented units and operating costs. This would be difficult to determine using a spreadsheet alone. As useful as spreadsheets are, you cannot reduce the last two assumptions to single values as required by the spreadsheet format. If you tried all the combinations, you would need to spend a great deal of time working through what-if scenarios, entering single values and recording the results. Even then, you would likely be left with a mountain of data instead of the overall profit and loss picture.

With Crystal Ball, this kind of analysis is easy.

Running the simulation

Use the following steps to run the simulation:

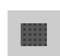
1. Select Run > Run.

Crystal Ball runs a simulation for the situation in the Futura Apartments spreadsheet and displays a **forecast** chart as it creates it.

Glossary Term: **forecast**— A statistical summary of the simulation results in a spreadsheet model, displayed graphically or numerically.

After the simulation has run for at least 500 **trials**, as displayed on your spreadsheet status bar.

Glossary Term: **trial,** ***also*** **iteration**— A three-step process in which Crystal Ball generates random numbers for assumption cells, recalculates the spreadsheet modes, and displays the results in a forecast chart.



2. In the forecast window, select Run > Stop.

Excel Note: *If the forecast window disappears behind Excel's window during a simulation, you can bring it back to the front by pressing <Alt>-<Tab>.*

The forecast window appears, as in Figure 1.2.

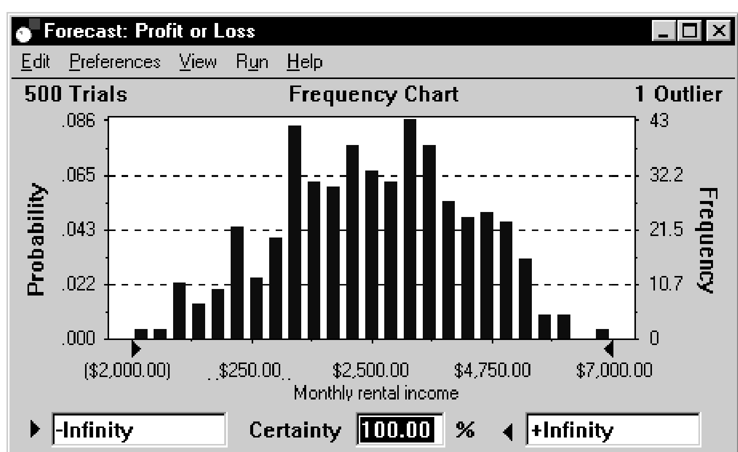


Figure 1.2 Futura Apartments profit/loss forecast

Glossary Term: **probability**— (Classical Theory) The likelihood of an event occurring.

The forecast chart reveals the total range of profit and loss outcomes predicted for the Futura Apartments situation. Each bar on the chart represents the likelihood, or probability, of earning a given income. The cluster of columns near the center indicates that the most likely income level is between $250 and $4,750 per month. Crystal Ball also forecasts that the worst case is a $2,000 loss and the best case is nearly a $7,000 gain.

Determining profit

Now you can use Crystal Ball to determine the statistical likelihood of making a profit and what is the most likely range.

To have Crystal Ball determine a statistical likelihood:

1. **Press <Tab> twice or select the left field in the forecast window.**
2. **Type 0 in the field.**
3. **Press <Enter>.**

   The value in the Certainty field changes to reflect the probability of an income level ranging from $0 to positive infinity—the probability of making a profit. With this information, you are in a much better position to make a decision on whether to purchase the Futura Apartments. As shown in Figure 1.3, there is roughly a 90% chance that you will make a profit.

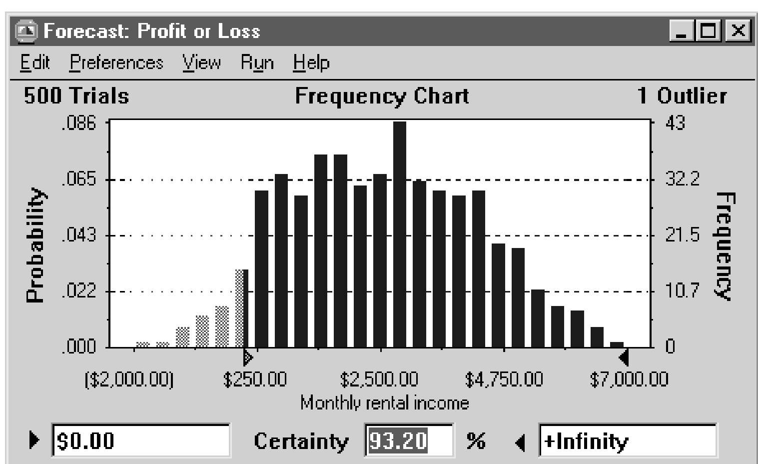


Figure 1.3 Chance of profit

Note that there is also a small chance of losing as much as $2000 per month.

How Crystal Ball uses Monte Carlo simulation

Glossary Term:
random number—
A mathematically selected value which is generated (by a formula or selected from a table) to conform to a probability distribution.

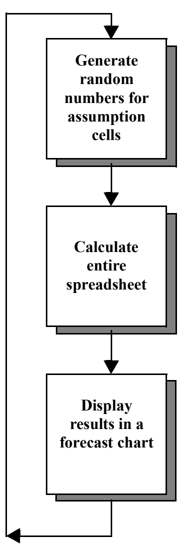
Glossary Term:
random number generator—
A method implemented in a computer program that is capable of producing a series of independent, random numbers.

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. Monte Carlo simulation is an efficient technique that requires only a **random number** table or a **random number generator** on a computer.

Crystal Ball Note: *For information on the random number generator that Crystal Ball uses, see "Random number generator" on page 342.*

This is an iterative process that continues until either:

- The simulation reaches a stopping criterion
- You stop the simulation manually



The final forecast chart reflects the combined uncertainty of the assumption cells on the model's output. Keep in mind that Monte Carlo simulation can only approximate a real-world situation. When you build and simulate your own spreadsheet models, you need to carefully examine the nature of the problem and continually refine the models until they approximate your situation as closely as possible.

Crystal Ball also provides statistics that describe the forecast results. For more information on statistics and forecast results, see Chapter **2**, "Understanding the Terminology," and Chapter **4**, "Interpreting the Results."

Closing the first tutorial

To close the first tutorial:

1. **Select Run > Reset.**

   A dialog appears, confirming the reset.

2. **Click on OK.**

   The simulation resets and the forecast window disappears.

3. **Close the Futura Apartments spreadsheet.**

Starting the second tutorial

The remainder of this chapter contains a tutorial for the Vision Research spreadsheet. This tutorial provides a more realistic situation to let you examine Crystal Ball's features in greater depth. However, if you feel comfortable running Crystal Ball now, you can turn to Chapter **2**, "Understanding the Terminology," for some background on Crystal Ball terminology. Then you can read Chapter **3**, "Setting Up and Running a Simulation," to start analyzing your own spreadsheets.

The Vision Research spreadsheet models a business situation filled with uncertainty. Vision Research has completed preliminary development of a new drug, code-named ClearView, that corrects nearsightedness. This revolutionary new product could be completely developed and tested in time for release next year if the FDA approves the product. Although the drug works well for some patients, the overall success rate is marginal, and Vision Research is uncertain whether the FDA will approve the product.

Vision Research can use Crystal Ball to help decide whether to scrap the project or to proceed to develop and market this potentially profitable new drug. The ClearView project is a multimillion-dollar risk. Crystal Ball is a powerful decision-support program designed to take the mystery out of decisions like this.

To see how Crystal Ball works in a typical business decision:

1. **Open the Vision Research spreadsheet from the Crystal Ball Examples folder. You can find this folder by selecting (in Windows) Start > Programs > Crystal Ball > Examples.**

   The Vision Research spreadsheet for the ClearView project appears, as in Figure 1.4.

Take a look at the Vision Research spreadsheet. This spreadsheet models the problem that Vision Research is trying to solve.

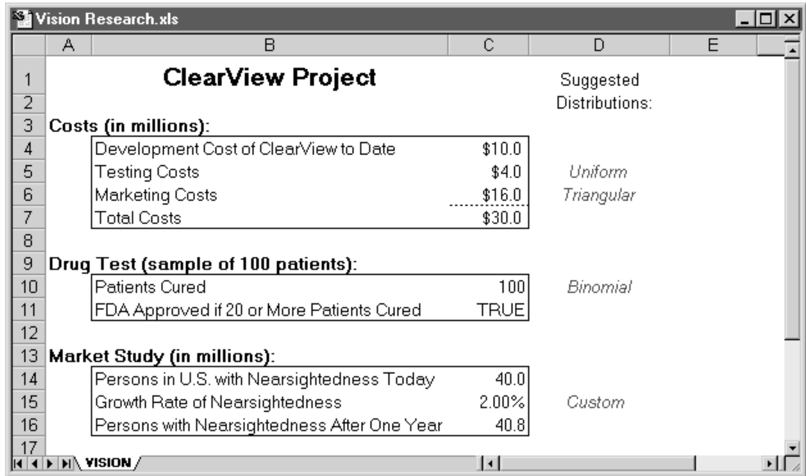


Figure 1.4 Vision Research's ClearView project spreadsheet

Defining assumptions

Glossary Term:
probability distribution— A set of all possible events and their associated probabilities.

In Crystal Ball, you define an assumption for a value cell by choosing a **probability distribution** that describes the uncertainty of the data in the cell. To accomplish this, you select from the 17 distribution types in the Distribution Gallery (see Figure 1.5).

How do you know which distribution type to choose? This portion of the tutorial will help you understand how to select a distribution type based on the answer you are looking for. In the following examples, you select the assumption cells in the Vision Research spreadsheet and choose the probability distributions that most accurately describe the uncertainties of the ClearView project.

This tutorial explains the reasons for choosing a particular distribution for each assumption. Detailed descriptions of how to select distributions are in Chapter **2**, "Understanding the Terminology," and Chapter **3**, "Setting Up and Running a Simulation."

Defining testing costs: the uniform distribution

So far, Vision Research has spent $10,000,000 developing ClearView and expects to spend an additional $3,000,000 to $5,000,000 to test it based on the cost of previous tests. For this variable, “testing costs,” Vision Research thinks that any value between $3,000,000 and $5,000,000 has an equal chance of being the actual cost of testing.

Using Crystal Ball, Vision Research chooses the uniform distribution to describe the testing costs. The uniform distribution describes a situation where all values between the minimum and maximum values are equally likely to occur, so this distribution best describes the company’s best guess for the cost of testing ClearView.

Once you choose the correct distribution type, you are ready to define the assumption cell.

To define the assumption cell for testing costs:

1. **Click on cell C5.**

2. **Select Cell > Define Assumption.**

   The Distribution Gallery dialog appears.

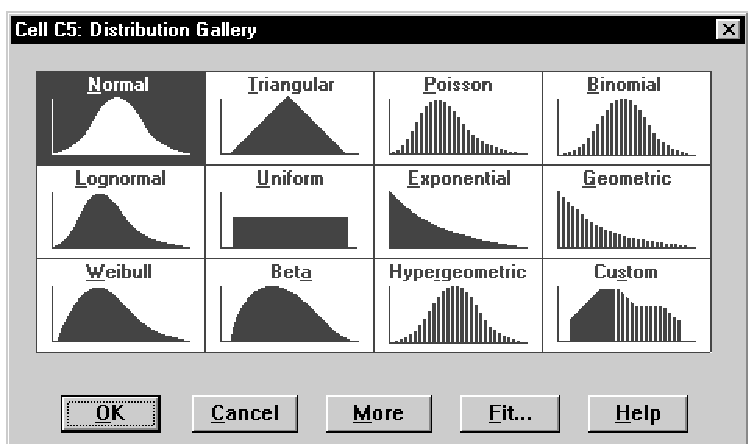


Figure 1.5 Distribution Gallery dialog

3. **Click on the Uniform distribution.**

4. **Click on OK.**

   The Uniform Distribution dialog appears.

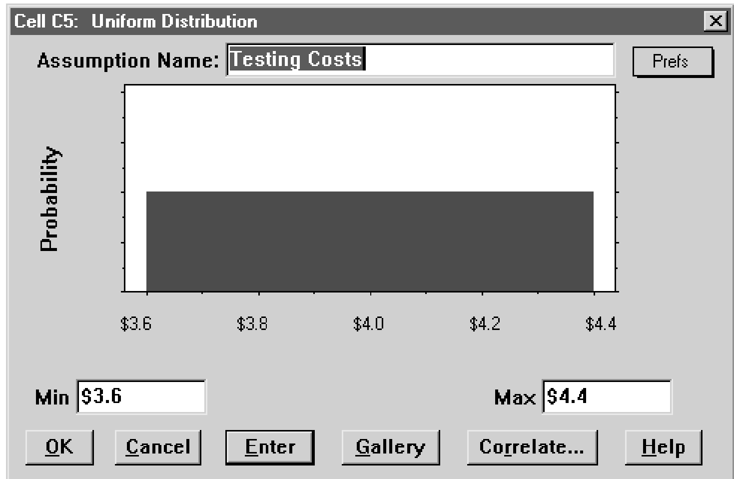


Figure 1.6 Uniform distribution for C5

Since cell C5 already has a name next to it on the spreadsheet, that name appears in the Assumption Name field. Use this name, rather than typing a new one. Also, notice that Crystal Ball assigns default values to the distribution. For more information on the method Crystal Ball uses to assign these default values to each distribution, see Appendix E, "Default Names and Distribution Parameters."

The uniform distribution has two parameters—minimum and maximum. Vision Research expects to spend a minimum of $3,000,000 and a maximum of $5,000,000 on testing. Use these values in place of the defaults to specify the parameters of the uniform distribution in Crystal Ball, as described in the following steps:

5. **Type 3 in the Min field (remember that the numbers on the spreadsheet represent millions of dollars).**

   This represents $3,000,000, the minimum amount Vision Research estimates for testing costs.

6. **Press <Tab>.**

7. **Type 5 in the Max field.**

   This represents $5,000,000, the maximum estimate for testing costs.

8. **Click on Enter.**

   The distribution changes to reflect the values you entered, as shown in Figure 1.7.

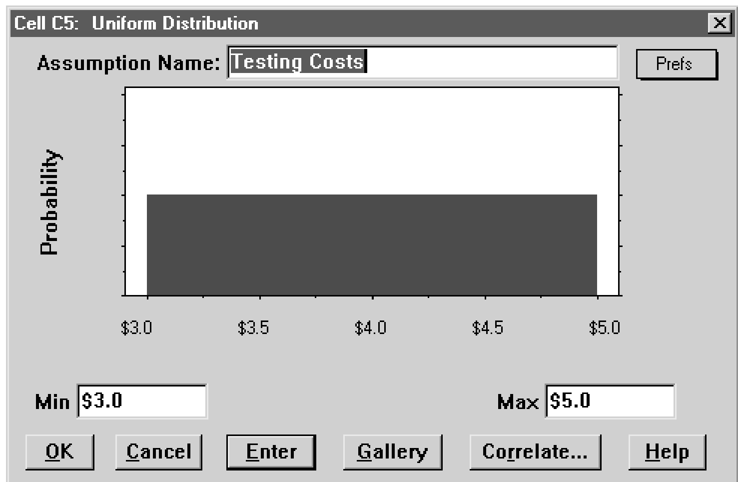


Figure 1.7 Changed distribution values

With the values from steps 5 and 7 entered correctly, your distribution looks like Figure 1.7. If not, repeat those steps. Later, when you run the simulation, Crystal Ball generates random values for cell C5 that are evenly spread between 3 and 5 million dollars.

9. **Click on OK to return to the spreadsheet.**

   The assumption cell is now green.

Defining marketing costs: the triangular distribution

Vision Research plans to spend a sizeable amount marketing ClearView if the FDA approves it. They expect to hire a large sales force and kick off an extensive advertising campaign to educate the public about this exciting new product. Including sales commissions and advertising costs, Vision Research expects to spend between $12,000,000 and $18,000,000, most likely $16,000,000.

Vision Research chooses the triangular distribution to describe marketing costs because the triangular distribution describes a situation where you can estimate the minimum, maximum, and most likely values to occur.

To define the assumption cell for marketing costs:

1. **Click on cell C6.**

2. **Select Cell > Define Assumption.**

   The Distribution Gallery dialog appears.

3. **Click on the Triangular distribution.**

4. **Click on OK.**

   The Triangular Distribution dialog appears for cell C6.

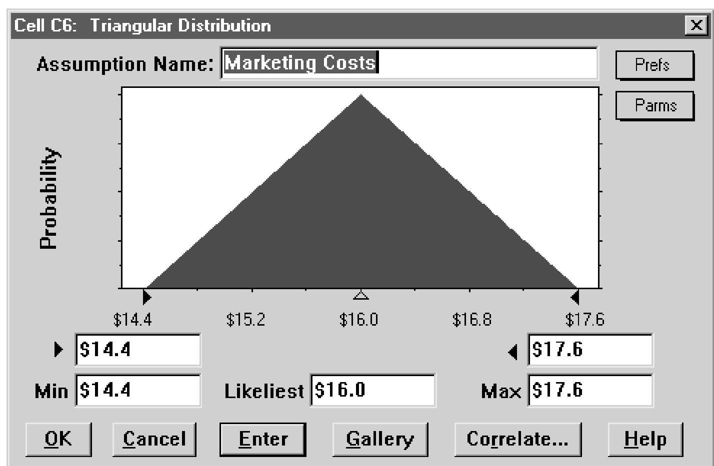


Figure 1.8 Triangular distribution for cell C6

Now specify the parameters for the triangular distribution. As you can see in Figure 1.8, the parameters for the triangular distribution are different from those specified earlier for the uniform distribution. The triangular distribution has three parameters—minimum, maximum, and likeliest. The following steps explain how to enter the parameters of the triangular distribution.

5. **Type 12 in the Min field.**

   This represents $12,000,000, the minimum amount Vision Research estimates for marketing costs.

6. **Press <Tab> to access the Likeliest field. If it doesn't contain the value 16, type 16.**

   This represents $16,000,000, the most likely amount for marketing costs.

7. **Press <Tab> and type 18 in the Max field.**

   This represents $18,000,000, the maximum estimate for marketing costs.

8. **Click on Enter.**

   The distribution changes to reflect the values you entered.

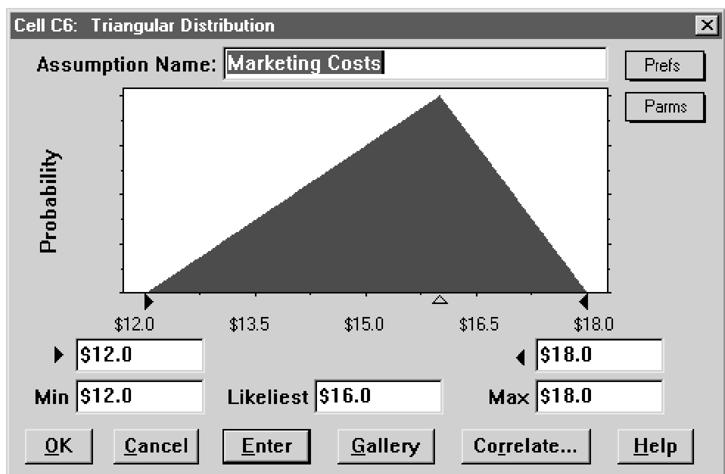


Figure 1.9 Distribution changed for new values

When you run the simulation, Crystal Ball generates random values that fall around 16, with fewer values near 12 and 18.

9. **Click on OK to return to the spreadsheet.**

Defining patients cured: the binomial distribution

Before the FDA will approve ClearView, Vision Research must conduct a controlled test on a sample of 100 patients for one year. Vision Research expects that the FDA will grant an approval if ClearView completely corrects the nearsightedness of 20 or more of these patients without any significant side-effects. In other words, 20% or more of the patients tested must show corrected vision after taking ClearView for one year. Vision Research is very encouraged by their preliminary testing, which shows a success rate of around 25%.

For this variable, “patients cured,” Vision Research only knows that their preliminary testing shows a cure rate of 25%. Will ClearView meet the FDA standards? Using Crystal Ball, Vision Research chooses the binomial distribution to describe the uncertainties in this situation because the binomial distribution describes the random number of successes (25) in a fixed number of trials (100).

To define the assumption cell for patients cured, use the following steps.

1. **Click on cell C10.**

2. **Select Cell > Define Assumption.**

   The Distribution Gallery appears.

3. **Click on the Binomial distribution.**
4. **Click on OK.**

   The Binomial Distribution dialog appears (notice that the default value for the probability parameter is 0.5 or 50%).

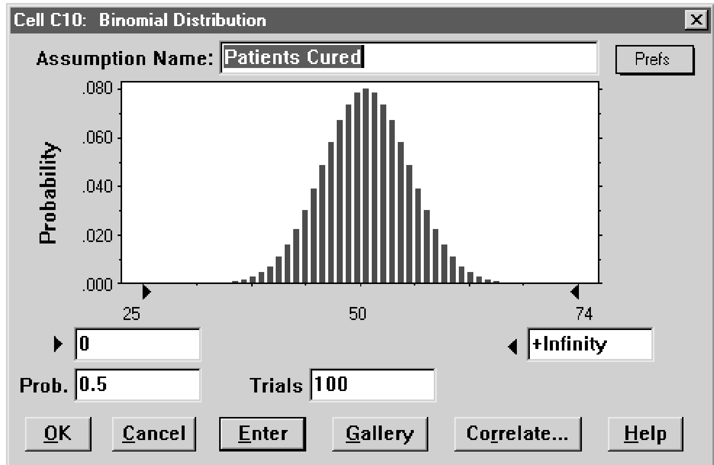


Figure 1.10 Binomial Distribution dialog

The binomial distribution has two parameters: probability (Prob) and trials. You know that Vision Research experienced a 25% success rate during preliminary testing, so use the value 0.25 for the probability parameter to show the likelihood or success.

Crystal Ball Note: *You can express probabilities either as decimal fractions between 0 and 1, such as 0.03, or as whole numbers followed by the percent sign, such as 3%.*

You also know the FDA expects Vision Research to test 100 people, so use the value 100 for the trials parameter. The following steps explain how to enter these parameters in the Binomial Distribution dialog.

5. **Type 0.25 in the Prob field.**

   This represents the 25% likelihood, or probability, of successfully correcting nearsightedness.

6. **If the Trials field doesn't contain the value 100, press <Tab>.**

7. **Type 100 in the Trials field.**

   This represents the 100 patients in the FDA test.

8. **Click on Enter.**

   The distribution changes to reflect the values you entered.

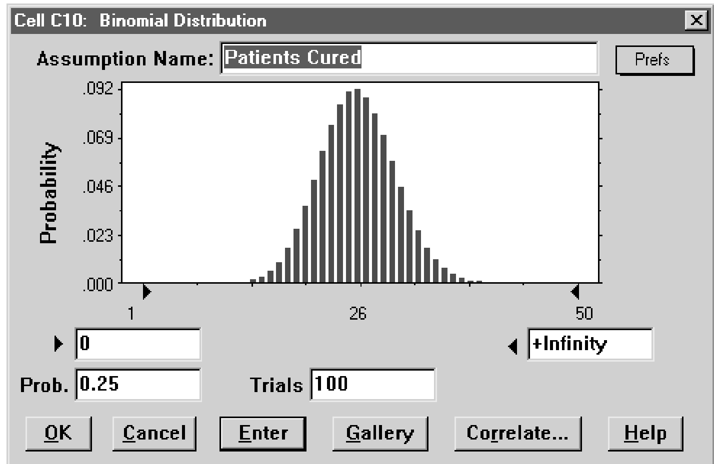


Figure 1.11 Changed binomial distribution

   When you run the simulation, Crystal Ball generates random integers between 0 and 100, simulating the number of patients that would be cured in the FDA test.

9. **Click on OK to return to the spreadsheet.**

Growth rate: the custom distribution

Vision Research has determined that nearsightedness afflicts nearly 40,000,000 people in the United States, and an additional 0% to 5% of these people will develop this condition during the year in which ClearView is tested.

However, the marketing department has learned that a 25% chance exists that a competing product will be released on the market soon. This product would decrease ClearView's potential market by 5% to 15%.

This variable, "growth rate of nearsightedness," cannot be described by any of the standard probability distributions. Since the uncertainties in this situation require a unique approach, Vision Research chooses Crystal Ball's custom distribution to define the growth rate. For the most part, the custom distribution is used to describe situations that other distribution types cannot.

The method for specifying parameters in the custom distribution is quite unlike the other distribution types, so follow the directions carefully. If you make a mistake, click on Gallery to return to the distribution gallery, then start again at step 3.

Use the custom distribution to plot both the potential increase and decrease of ClearView's market.

To define the assumption cell for the growth rate of nearsightedness:

1. **Click on cell C15.**

2. **Select Cell > Define Assumption.**

   The Distribution Gallery dialog appears.

3. **Click on the Custom distribution.**

4. **Click on OK.**

   The Custom Distribution dialog appears. Notice that in Figure 1.12 that the chart area remains empty until you enter the values for the distribution.

The area remains empty until you enter values.

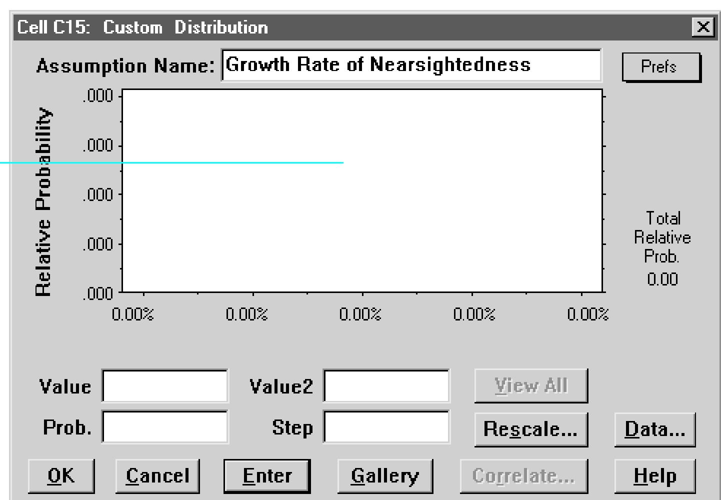


Figure 1.12 Custom Distribution dialog

To enter the first range of values:

5. Type 0% in the Value field.

This represents a 0% increase in the potential market.

6. Press <Tab>.

7. Type 5% in the Value2 field.

This represents a 5% increase in the potential market.

8. Press <Tab>.

9. Type 75% in the Prob field.

This represents the 75% chance that Vision Research's competitor will not enter the market and reduce Vision Research's share.

10. Click on Enter.

A uniform distribution for the range 0% to 5% appears.

Uniform distribution for the first range of values.

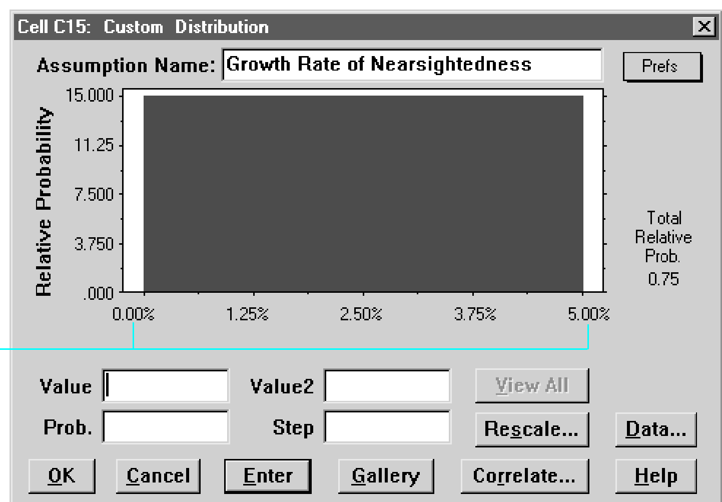


Figure 1.13 Uniform distribution range

To enter a second range of values:

11. Type -15% in the Value field.

This represents a 15% decrease in the potential market.

12. Press <Tab>.

13. Type -5% in the Value2 field.

This represents a 5% decrease in the potential market.

14. Press <Tab>.

15. Type 25% in the Probability field.

This represents the 25% chance that Vision Research's competitor will enter the market place and decrease Vision Research's share by 5% to 15%.

16. Click on Enter.

A uniform distribution for the range -15% to -5% appears. Both ranges now appear in the Custom Distribution dialog.

Crystal Ball displays both ranges.

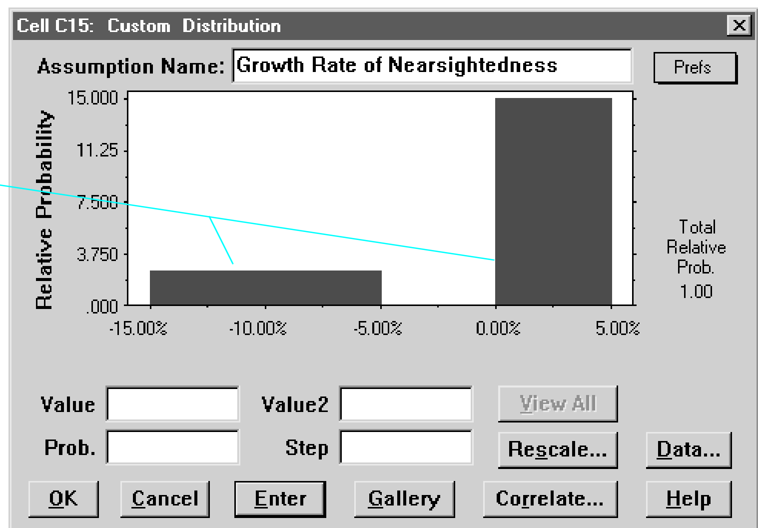


Figure 1.14 Customized uniform distribution

In Chapter **2**, "Understanding the Terminology", you learn to use the data button, a special feature on the Custom Distribution dialog. You can use the Data button to pull numbers from specified cell ranges on the spreadsheet rather than typing them in the Custom Distribution dialog. When you run the simulation, Crystal Ball generates random values within the ranges you specified.

17. Click on OK to return to the spreadsheet.

Glossary Term:
standard deviation—
The square root of the variance for a distribution. A measurement of the dispersion of values around the mean.

Glossary Term:
mean or mean value—
The familiar arithmetic average of a set of numerical observations (the sum of the observations divided by the number of observations).

Defining market penetration: the normal distribution

The marketing department estimates that Vision Research's eventual share of the total market for the product will be normally distributed around a mean value of 8% with a **standard deviation** of 2%. "Normally distributed" means that Vision Research expects to see the familiar bell-shaped curve with about 68% of all possible values for market penetration falling between one standard deviation below the **mean value** and one standard deviation above the mean value, or between 6% and 10%.

The low mean value of 8% is a conservative estimate that takes into account the side effects of the drug that were noted during preliminary testing. In addition, the marketing department estimates a minimum market of 5%, given the interest shown in the product during preliminary testing.

Vision Research chooses the normal distribution to describe the variable "market penetration."

To define the assumption cell for market penetration:

1. **Click on cell C19.**

2. **Select Cell > Define Assumption.**

   The Distribution Gallery dialog appears.

3. **Click on the Normal distribution.**
4. **Click on OK.**

   The Normal Distribution dialog appears.

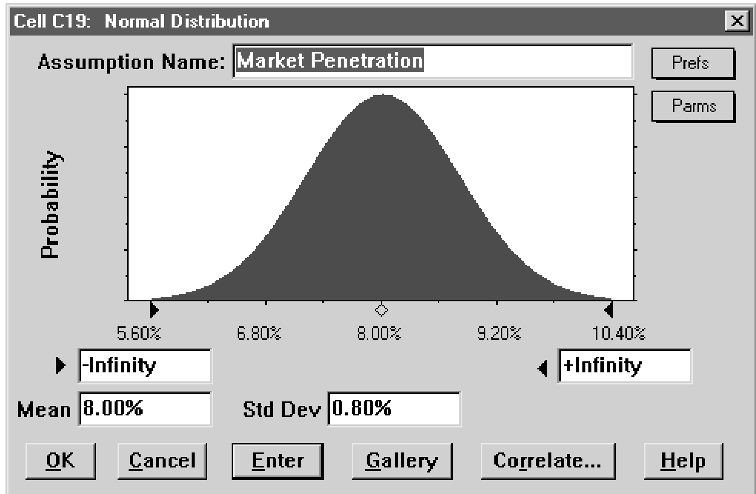


Figure 1.15 Normal distribution for cell C19

Now specify the parameters for the normal distribution: the mean and the standard deviation.

5. **If the Mean field doesn't contain 8.00%, type 8% in the Mean field.**

   This represents an estimated average for market penetration of 8%.

6. **Press <Tab>.**

7. **Type 2% in the StdDev field.**

   This represents an estimated 2% standard deviation from the mean.

8. **Click on Enter.**

   The normal distribution scales to fit the chart area, so the shape of the distribution does not change. However, the percent range at the bottom of the chart does change.

9. **Press <Tab> twice.**

10. **Type 5% in the left truncation grabber field.**

    This represents 5%, the minimum market for the product.

11. Click on Enter.

The distribution changes to reflect the values you entered.

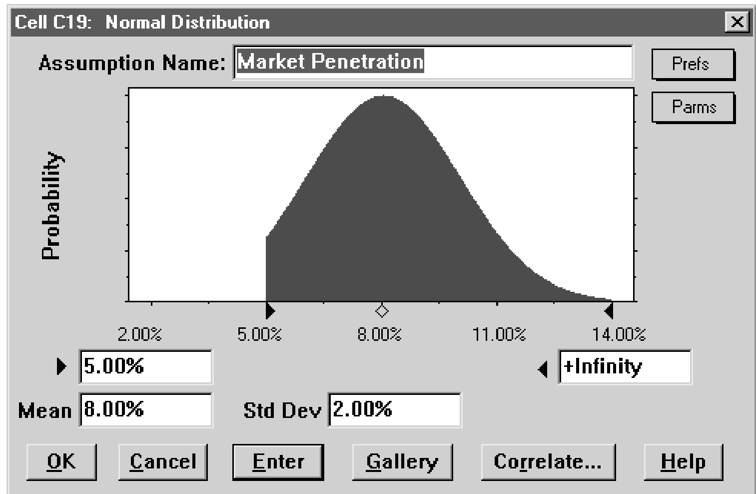


Figure 1.16 Changed distribution for the new values

When you run the simulation, Crystal Ball generates random values that follow a normal distribution around the mean value of 8%, and with no values generated below the 5% minimum limit.

12. Click on OK to return to the spreadsheet.

Defining forecasts

Now that you have defined the assumption cells in your model, you are ready to define the forecast cells. Forecast cells contain formulas that refer to one or more assumption cells.

The president of Vision Research would like to know both the likelihood of achieving a profit on the product and the most likely profit, regardless of cost. These forecasts appear in the gross profit (cell C21) and net profit (cell C23) for the ClearView project.

Calculating gross profit

Crystal Ball can generate more than one forecast during a simulation. In this case, you can define both the gross profit and net profit formulas as forecast cells. First, look at the contents of the cell for gross profit.

1. **Click on cell C21.**

   The cell contents appear in the formula bar near the top of your spreadsheet. The contents are C16*C19*C20. Crystal Ball uses this formula to calculate gross profit by multiplying Persons With Nearsightedness After One Year (C16) by Market Penetration (C19) by Profit Per Customer (C20).

Glossary Term:
forecast formula—
A formula that has been defined as a forecast cell.

Now that you understand the gross profit formula, you are ready to define the **forecast formula** cell for gross profit.

To define the forecast cell for gross profit:

2. **Select Cell > Define Forecast.**

   The Define Forecast dialog appears. You can enter a name for the forecast. Since the forecast cell has a name next to it on the spreadsheet, that name appears in the dialog by default.

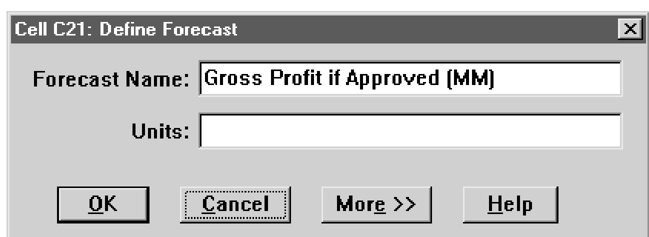


Figure 1.17 Define Forecast dialog - Gross Profit If Approved

   Use the forecast name that appears, rather than typing a new name.

   Since the spreadsheet model involves millions of dollars, indicate that in this dialog.

3. **Press <Tab>.**
4. **Type Millions in the Units field.**

5. Click on OK to return to the spreadsheet.

Calculating net profit

Before defining the forecast cell formula for net profit, look at the contents of the cell for net profit:

1. Click on cell C23.

The contents appear in the formula bar above the spreadsheet. The contents are: IF(C11, C21-C7, -C4-C5).

The formula translates as follows:

If the FDA approves the drug (C11 is true), then calculate net profit by subtracting total costs (C7) from gross profit (C21). However, if the FDA does not approve the drug, (C11 is false), then calculate net profit by deducting both development costs (C4) and testing costs (C5) incurred to date.

To define the forecast cell for net profit:

2. Select Cell > Define Forecast.

The Define Forecast dialog appears, as in Figure 1.18.

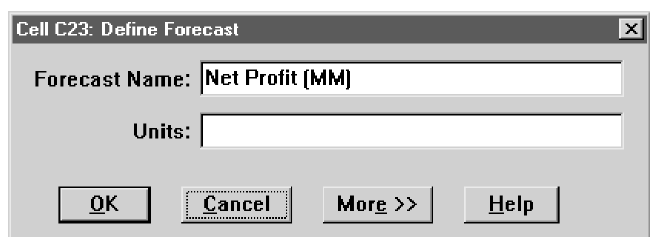


Figure 1.18 Define Forecast dialog - Net Profit

Again, use the forecast name that appears in the Forecast Name field and specify millions in the Units field.

3. Press <Tab>.

4. Type Millions in the Units field.

5. Click on OK to return to the spreadsheet.

You have defined assumptions and forecast cells for the Vision Research spreadsheet, and are now ready to run a simulation.

Running a simulation

When you run a simulation in Crystal Ball, you have the freedom to stop and then continue the simulation at any time. The Run, Stop, and Continue commands appear on the Run menu as you need them. For example, when you are running a simulation, the Stop command appears at the top of the menu. If you stop the simulation, the Continue command takes its place. Practice using these commands when you run the simulation for the ClearView project.

Glossary Term:
seed value—
The first number in a sequence of random numbers. A given seed value produces the same sequence of random numbers every time you run a simulation.

Before you begin the simulation, specify the number of trials and initial **seed value** so your simulation will look like the forecast charts in this tutorial. See Chapter **4**, “Interpreting the Results” for more information on trials and seed values.

To specify the number of trials and initial seed value:

1. **Select Run > Run Preferences > Trials.**

   The Run Preferences Trials dialog appears.

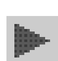
2. **In the Maximum Number Of Trials field, type 500.**
3. **Click on Sampling.**
4. **Click the Use Same Sequence Of Random Numbers option.**
5. **In the Initial Seed Value field, type 999.**
6. **Click on OK.**

Practice routines

In the next few sections, practice a few basic routines.

Run, Stop, and Continue

To practice using the Run, Stop, and Continue commands:

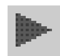


1. **Select Run > Run.**

   The Net Profit forecast window appears, neatly stacked on top of the Gross Profit forecast window. As the simulation proceeds, the forecast charts reflect the changing values in the forecast cells.

2. **Select Run > Stop.**

Crystal Ball Note: *You can also stop the simulation by pressing <Alt>-u,o.*



3. In the front forecast window, select Run > Continue.

Crystal Ball Note: *Each forecast window has its own menu bar.*

The simulation continues.

Crystal Ball Note: *You can also continue the simulation by pressing <Alt>-u,u.*

View the forecast charts

You might not see two complete forecast charts at the same time. However, there are several ways to bring individual forecast windows to the front of the window stack. The easiest way is to click on the forecast window if it is visible.

1. Click on the Gross Profit If Approved forecast window.

The Gross Profit chart appears, as in Figure 1.19.

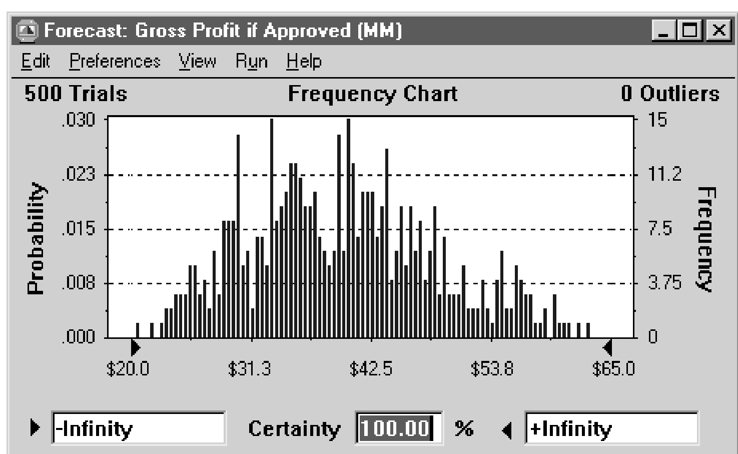


Figure 1.19 Gross Profit If Approved forecast

2. Select Run > Forecast Windows.

3. Click on Open All Forecasts to move the Net Profit forecast window to the front again.

While the simulation is running, a frequency distribution for each forecast appears to reflect the changing values in the forecast cell. The frequency distribution appears as columns.

4. Continue to run the simulation until it stops at 500 trials.

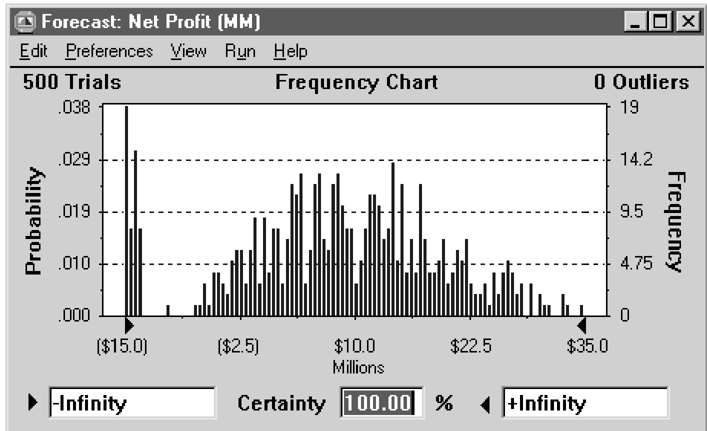


Figure 1.20 Net Profit forecast

A frequency distribution shows the number or frequency of values occurring in a given group interval. In Figure 1.20, the frequency distribution on the Net Profit forecast chart shows a frequency of 19 for the group interval that contains the most values. That means 19 values occurred in the group interval.

Chapter **3**, "Setting Up and Running a Simulation" and Chapter **4**, "Interpreting the Results" describe the forecast chart in more detail. For now, remember that the forecast chart graphs the forecast results and shows how the forecast values are distributed. As the simulation progresses, Crystal Ball continues to update the frequency distribution for each forecast cell, and the forecast results become more accurate.

Interpreting the results

Now that you have run the simulation, you are ready to interpret the forecast results in more depth. The president of Vision Research faces a difficult decision—should the company scrap the ClearView project or proceed to develop and market this revolutionary new drug? To examine this question in more detail you need to look at the forecast chart in more detail.

Excel Note: *Crystal Ball windows are separate from Excel windows. If Crystal Ball's windows disappear from your screen, they are usually simply behind the main Excel window. To bring them to the front, click on the Crystal Ball icon in the Windows taskbar or press <Alt>-<Tab>.*

Understanding the forecast chart

Crystal Ball forecasts the entire range of results for the Vision Research project. However, the forecast charts show only the display range that, by default, is set to exclude the most extreme values (outliers). Chapter **4**, "Interpreting the Results" describes how to change the display range to examine specific sections of the forecast in greater detail.

In Figure 1.21, the display range includes the values from minus $15.00 to $35.00, as shown on the Net Profit forecast chart.

The forecast chart also shows the certainty range for the forecast. By default, the certainty range includes all values from negative infinity to positive infinity.

Crystal Ball compares the number of values in the certainty range with the number of values in the entire range to calculate the certainty level. The example above shows a certainty level of 100%, since the initial certainty range includes all possible values.

Remember that the certainty level is an approximation, since the spreadsheet model can only approximate the elements of the real world.

In the upper left corner of the forecast chart, Crystal Ball shows the total number of trials that were run for this forecast. In the upper right corner, the number of trials outside of the display range appear. Since the display range is initially set by default to exclude outlying values, the number of "outliers" is usually greater than zero when your simulation stops.

Determining the certainty level

Now the Vision Research president wants to know how certain Vision Research can be of achieving a profit and what are the chances of a loss.

To determine the certainty level of a specific value range:

1. **In the Net Profit forecast chart, press <Tab> twice.**
2. **Type 0 in the range minimum field.**
3. **Press <Enter>.**

   Crystal Ball moves the left certainty grabber to the break-even value of $0.0 and recalculates the certainty level.

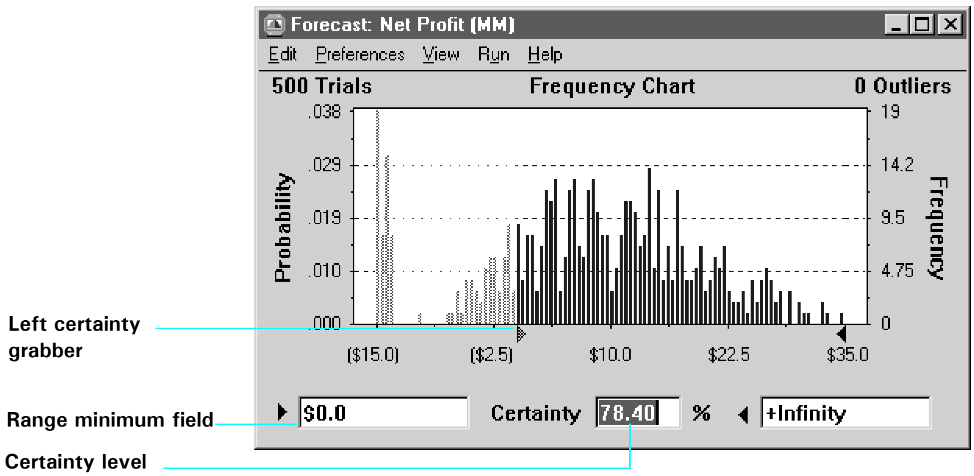


Figure 1.21 Net Profit forecast—$0 minimum

Analyzing the Net Profit forecast chart again, you can see that the value range between the certainty grabbers shows a certainty level of 78.4%. That means that Vision Research can be 78.4% certain of achieving a net profit. You can therefore calculate a 21.6% chance of suffering a net loss (100% minus 78.4%).

Now the president of Vision Research would like to know the certainty of achieving a minimum profit of $2,000,000. With Crystal Ball, you can easily answer this question.

4. **Type 2 in the range minimum field.**
5. **Press <Enter>.**

   A Figure 1.22 shows, Crystal Ball moves the left certainty grabber to $2.0 and recalculates the certainty level.

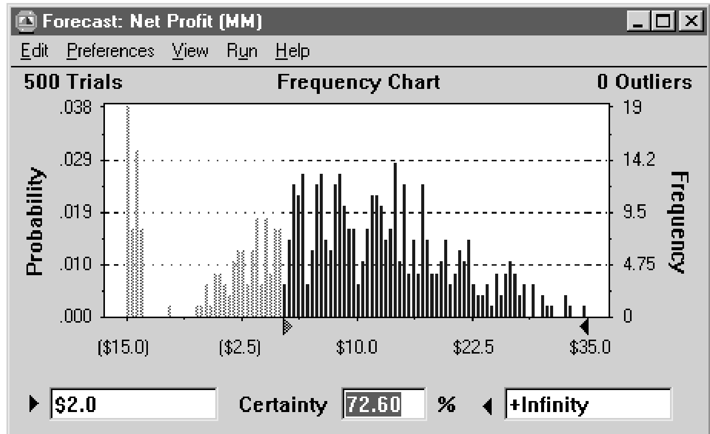


Figure 1.22 Recalculated certainty level

Vision Research can be 72.6% certain of achieving a minimum net profit of $2,000,000.

Vision Research is very encouraged by the forecast result. The president now wants to know how certain Vision Research can be of achieving a minimum net profit of $4,000,000. If Crystal Ball shows that Vision Research can be at least two-thirds certain of a $4,000,000 net profit, the president is ready to go ahead with the ClearView project. Again, Crystal Ball can easily answer this question.

6. **Type 4 in the range minimum field.**
7. **Press <Enter>.**

   Crystal Ball moves the left certainty grabber to $4.0 and recalculates the certainty level.

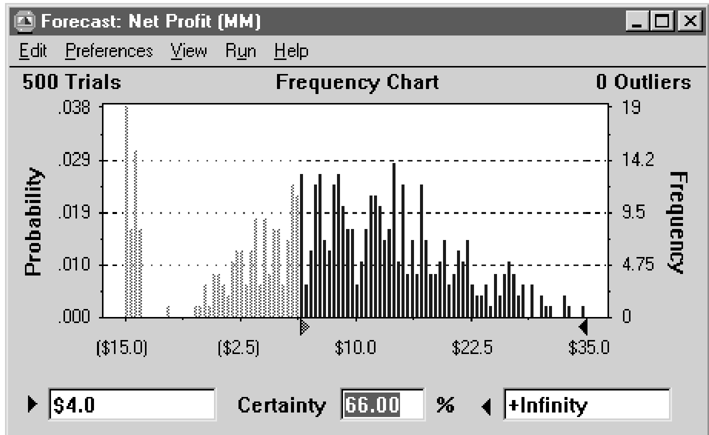


Figure 1.23 net Profit forecast—$4.0 minimum

The Net Profit forecast chart in Figure 1.23 shows a certainty level of 66%. With a two-thirds certainty of a minimum net profit of $4,000,000, Vision Research decides to go ahead with the ClearView project and proceed to develop and market this revolutionary new drug.

The president of Vision Research also is interested in the most likely profit regardless of cost. You now can analyze the gross profit forecast chart as you did the net profit chart.

Summary

In this tutorial, you explored only a few questions that Vision Research might ask as they analyze the results of the simulation. As you read through this manual, you will learn to explore the forecast results in more depth. For example, you can customize the forecast charts, create overlay charts, create trend charts, analyze the sensitivity of the model, interpret the descriptive statistics, and print comprehensive reports for any simulation. Crystal Ball provides these features so that you can quantify the risk inherent in your assumptions.

Crystal Ball keeps your assumption and forecast definitions (but not the forecast values) with the spreadsheet. When you save your spreadsheet, the definitions are saved with it. To learn about saving forecast results, see "Saving and restoring a simulation" on page 178.

Closing Crystal Ball

To close Crystal Ball without exiting Excel, in Excel, select Run > Close Crystal Ball.

A dialog appears asking you to confirm the decision. Click on OK to close Crystal Ball and remove the Cell, Run, and CBTools menus and the Crystal Ball toolbar from Excel. The Vision Research spreadsheet remains open in Excel.

Crystal Ball Note: *Crystal Ball also closes automatically when you exit Excel.*

Chapter 2

Understanding the Terminology

In this chapter

- Model building and risk analysis
- Understanding probability distributions
- What is probability?
- Selecting a probability distribution
- Most commonly used distributions
- Other commonly used distributions
- Less commonly used distributions
- Understanding other statistical terms

This chapter explains probability and probability distributions. Your understanding of both probability and probability distributions will let you select the right probability distribution for your spreadsheet model. The first section describes in detail the seventeen distribution types Crystal Ball uses and demonstrates using real-world examples. The last section of the chapter explains statistical terms used to describe forecast results.

Model building and risk analysis overview

To use Crystal Ball, you need to first *build a model* and then Crystal Ball will help you perform a *risk analysis*. However, if you're a first-time user, you might not be familiar with or know what is meant by models or risk analysis. Or, if you do know, you might want a better understanding of how Crystal Ball performs a risk analysis.

What is a model?

Crystal Ball works with spreadsheet models, specifically Excel spreadsheet models. Your spreadsheet might already contain a model, depending on what type of information you have in your spreadsheet and how you use it.

Data vs. analysis

If you only use spreadsheets to hold data—sales data, inventory data, account data, etc., then you don't have a model. Even if you have formulas that total or subtotal the data, you might not have a model that is useful for simulation.

A model is a spreadsheet that has taken the leap from being a data organizer to an analysis tool. A model represents a process with combinations of data, variables, formulas, and functions. As you add cells that help you better understand and analyze your data, your data spreadsheet becomes a spreadsheet model.

Traditional spreadsheet analysis

So now you have a model, or you have create your first model. For each value in your model, ask yourself, "How certain am I of this value? Will it vary? Is this a best estimate or a known fact?" You might notice that your model has some values in it that you are unsure of. Perhaps you don't have the actual data yet (this month's sales figures) or the value varies unpredictably (individual item cost). This lack of knowledge about particular values or the knowledge that some values may always vary contribute to the model's *uncertainty*.

Traditional spreadsheet analysis tries to capture this uncertainty in one of three ways:

- Point estimates
- Range estimates
- What-if scenarios

Point estimates

Point estimates are when you use what you think are the most likely values (technically referred to as the mode) for the uncertain variables. These estimates are the easiest, but can return very misleading results. For example, try crossing a river with an average depth of three feet. Or, if it takes you an average of 25 minutes to get to the airport, leave 25 minutes before your flight takes off. You will miss your plane 50% of the time.

Range estimates

Range estimates typically calculate three scenarios: the best case, the worst case, and the most likely case. These types of estimates can show you the range of outcomes, but not the probability of any of these outcomes.

What-if scenarios

What-if scenarios are usually based on the range estimates, and calculate as many scenarios as you can think of. What is the worst case? What if sales are best case but expenses are the worst case? What if sales are average, but expenses are the best case? What if sales are average, expenses are average, but sales for the next month are flat?

As you can see, this is extremely time consuming, and results in lots of data, but still doesn't give you the probability of achieving different outcomes.

Risk

Uncertainty, like that described above, can often indicate *risk*, which is the possibility of loss, damage, or any other undesirable event and the severity associated with the event. For example, if sales for next month are above a certain amount (a desirable event), then orders will reduce the inventory. If the reduction in inventory is large enough, there will be a delay in shipping orders (an undesirable event). If a delay in shipping means losing orders (severity), then that possibility presents a risk.

There are two points to keep in mind when analyzing risk:

- Where is the risk?
- How significant is the risk?

Almost any change, good or bad, poses some risk. Your own analysis will usually reveal numerous potential risk areas: overtime costs, inventory shortages, future sales, geological survey results, personnel fluctuations, unpredictable demand, changing labor costs, government approvals, potential mergers, pending legislation.

Once you identify your risks, a model can help you quantify them. Quantifying risk means determining the chances that the risk will occur and the cost if it does, to help you decide whether a risk is worth taking. For example, if there is a 25% chance of running over schedule, costing you $100 out of your own pocket, that might be a risk you are willing to take. But if you have a 5% of running over schedule, knowing that there is a $10,000 penalty, you might be less willing to take that risk.

Risk analysis

You can perform risk analysis in several ways, but one method involves building a spreadsheet model. A good spreadsheet model can be very helpful in identifying where your risk might be, since cells with formulas and cell references identify causal relationships among variables.

One of the drawbacks of conventional spreadsheet models, however, is that you can only enter one value in a cell at a time. Why would you want to put more than one value in a cell, you ask? Remember those uncertain values that you could represent either with point estimates, range estimates, or what-if scenarios? A spreadsheet will not let you put in a range or multiple values; it allows only one value at a time. So calculating the range requires you to replace the uncertain value several times to see what effect the minimum, most likely, and maximum values have.

Calculating more realistic what-if scenarios is the same, except it requires you to change your spreadsheet even more. And don't forget to keep track of all the results somewhere, or you will have to repeat the scenario!

This is where Crystal Ball comes in. Crystal Ball helps you define those uncertain variables in a whole new way: by defining the cell with a range or a set of values. So you can define your business phone bill for future months as any value between $2500 and $3750, instead of using a single point estimate of $3000. It then uses the defined range in a simulation.

In addition, Crystal Ball keeps track of the results of each scenario for you.

Monte Carlo simulation

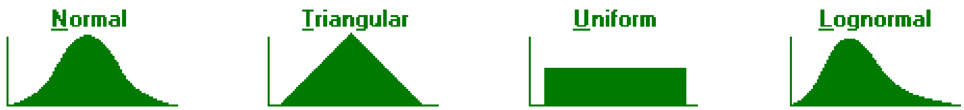
Simulation is any analytical method that is meant to imitate a real-life system, especially when other analyses are too mathematically complex or too difficult to reproduce. Spreadsheet risk analysis uses both a spreadsheet model and simulation to analyze the effect of varying inputs on outputs of the modeled system. One type of spreadsheet simulation is Monte Carlo simulation, which randomly generates values for uncertain variables over and over to simulate a model.

History

Monte Carlo simulation was named for Monte Carlo, Monaco, where the primary attractions are casinos containing games of chance. Games of chance such as roulette wheels, dice, and slot machines, 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 either a 1, 2, 3, 4, 5, or 6 will come up, but you don't know which for any particular trial. 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).

For each variable, you define the possible values with a *probability distribution*. The type of distribution you select depends on the conditions surrounding the variable. For example, some common distribution types are:



During a simulation, the value to use for each variable is selected randomly from the defined possibilities.

Discussion

A simulation calculates numerous scenarios of a model by repeatedly picking values from the probability distribution for the uncertain variables and using those values for the cell. Commonly, a Crystal Ball simulation calculates hundreds or thousands of scenarios in just a few seconds.

Since all those scenarios produce associated results, Crystal Ball also keeps track of the *forecasts* for each scenario. Forecasts are cells (usually with formulas of functions) that you define as important outputs of the model. These usually are cells such as totals, net profit, or gross expenses.

For each forecast, Crystal Ball remembers the cell value for all the trials (scenarios). During the simulation, you can watch a histogram of the results, which shows how they stabilize toward a smooth *frequency distribution* as the simulation progresses. After hundreds or thousands of trials, you can view sets of values, the statistics of the results (such as the mean forecast value), and the certainty of any particular value.

Certainty

The forecast results not only show you the different result values for each forecast, but also the probability of any value. Crystal Ball normalizes these probabilities to calculate another important number: the *certainty*. The certainty is the percent chance that a particular forecast value will fall within a specified range.

For example, the chance of any forecast value falling between -Infinity and Infinity is always 100%. However, the chance of that same forecast being at least zero (which you might want to calculate to make sure that you make a profit) might be only 45%.

For any range you define, Crystal Ball calculates the resulting certainty. This way, you not only know that your company has a chance to make a profit, you can quantify that chance by saying that it has a 45% chance of making a profit on a venture (a venture you might decide to skip).

Benefits of risk analysis

Finding the certainty of achieving a particular result is often the goal of a model analysis. Risk analysis takes a model and sees what effect changing different values has on the bottom line. Risk analysis can:

- Help end "analysis paralysis," and contribute to better decision-making by quickly examining all possible scenarios
- Identify which variables most affect the forecast
- Expose the uncertainty in a model, leading to a better communication of risk

Understanding probability distributions

To begin to understand probability, consider this example: You want to look at the distribution of non-exempt wages within one department of a large company. First, you gather raw data—in this case the wages of each non-exempt employee in the department. Second, you organize the data into a meaningful format and plot the data as a frequency distribution on a chart. To create a frequency distribution, you divide the wages into group intervals and list these intervals on the chart's horizontal axis. Then you list the number or frequency of employees in each interval on the chart's vertical axis. Now you can easily see the distribution of non-exempt wages within the department.

A glance at the chart illustrated in Figure 2.1 reveals that most of the employees (approximately 60 out of a total of 180) earn from $7 to $9 per hour.

Figure 2.1 Raw data for probability

You can chart this data as a probability distribution. A probability distribution shows the number of employees in each interval as a fraction of the total number of employees. To create a probability distribution, you divide the number of employees in each interval by the total number of employees and list the results on the chart's vertical axis.

The chart illustrated in Figure 2.2 shows you the number of employees in each wage group as a fraction of all employees; you can estimate the likelihood or probability that an employee drawn at random from the whole group earns a wage within a given interval. For example, assuming the same conditions exist at the time the sample was taken, the probability is 0.33 (a 1 in 3 chance) that an employee drawn at random from the whole group earns between $8 and $8.50 an hour.

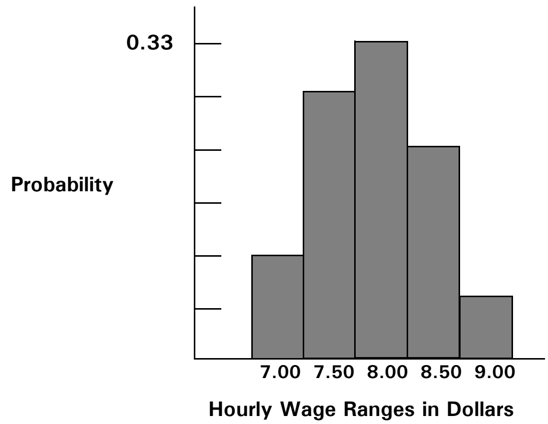


Figure 2.2 Probability distribution of wages

Compare the probability distribution in the example above to the probability distributions in Crystal Ball as shown in Figure 2.3.

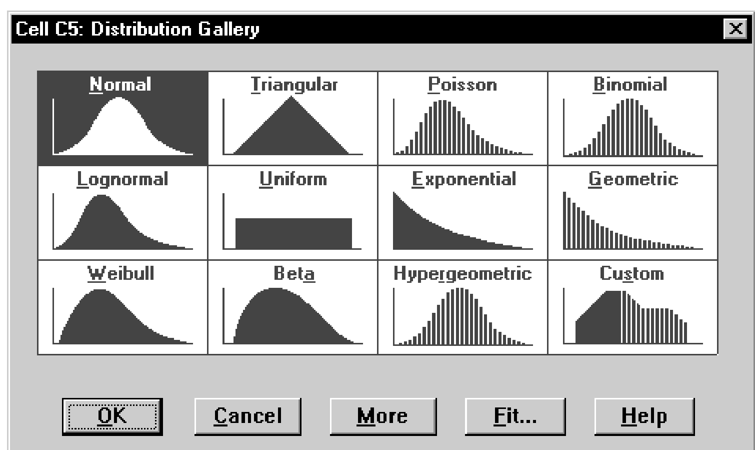


Figure 2.3 Distribution Gallery dialog

The probability distribution in the example in Figure 2.2 has a shape similar to many of the distributions in the Distribution Gallery. This process of plotting data as a frequency distribution and converting it to a probability distribution provides one starting point for selecting a Crystal Ball distribution. Select the distributions in the gallery that appear similar to your probability distribution, then read about those distributions in this chapter to find the correct distribution.

For information about the similarities between distributions, see "Comparing the distributions" on page 111. For a complete discussion of probability distributions, refer to the sources listed in the Bibliography.

Notice that the Distribution Gallery shows whether the probability distributions are discrete or continuous. Discrete probability distributions describe distinct values, usually integers, with no intermediate values and are shown as a series of vertical bars, such as in the binomial distribution in the example above. A discrete distribution, for example, might describe the number of heads in four flips of a coin as 0, 1, 2, 3, or 4.

Continuous probability distributions, such as the normal distribution, describe values over a range or scale and are shown as solid figures in the Distribution Gallery. Continuous distributions are actually mathematical abstractions because they assume the existence of every possible intermediate value between two numbers. For example, a continuous distribution assumes there are an infinite number of values between any two points on the distribution.

However, in many situations, a continuous distribution can be effectively used to approximate a discrete distribution even though the continuous model does not necessarily describe the situation exactly.

In the dialogs for the discrete distributions, Crystal Ball displays the values of the variable on the horizontal axis and the associated probabilities on the vertical axis. For the continuous distributions, Crystal Ball does not display values on the vertical axis since, in this case, probability can only be associated with areas under the curve and not with single values.

Crystal Ball Note: *The precision and format of the displayed numbers in the probability and frequency distributions come from the cell itself. For example, if your cell's precision is set to two decimal places (#.##), the number 0.002946 will be displayed as 0.00. See the user manual for your spreadsheet program for more information about changing the format of the number display.*

Selecting a probability distribution

Plotting data is one guide to selecting a probability distribution. The following steps provide another process for selecting probability distributions that best describe the uncertain variables in your spreadsheets.

To select the correct probability distribution:

- Look at the variable in question. List everything you know about the conditions surrounding this variable.

  You might be able to gather valuable information about the uncertain variable from historical data. If historical data is not available, use your own judgment, based on experience, to list everything you know about the uncertain variable.

- Review the descriptions of the probability distributions.

  This chapter describes each distribution in detail, outlining the conditions underlying the distribution and providing real-world examples of each distribution type. As you review the descriptions, look for a distribution that contains the conditions you have listed for this variable.

- Select the distribution that characterizes this variable.

  A distribution characterizes a variable when the conditions of the distribution match those of the variable.

The conditions of the variable describe the values for the parameters of the distribution in Crystal Ball. Each distribution type has its own set of parameters, which are explained in the following descriptions. In addition to the standard parameter set, each continuous distribution (except uniform) also lets you choose from alternate parameter sets, which substitute percentiles for one or more of the standard parameters. For more information on alternate parameters, see “Alternate parameter sets” on page 136.

Most commonly used distributions

Uniform distribution

In the uniform distribution, all values between the minimum and maximum occur with equal likelihood.

Conditions

The three conditions underlying uniform distribution are:

- The minimum value is fixed.
- The maximum value is fixed.
- All values between the minimum and maximum occur with equal likelihood.

Example one

An investment company interested in purchasing a parcel of prime commercial real estate wants to describe the appraised value of the property. The company expects an appraisal of at least $500,000 but not more than $900,000. They believe that all values between $500,000 and $900,000 have the same likelihood of being the actual appraised value.

The first step in selecting a probability distribution is to match your data with a distribution's conditions. In this case:

- The minimum value is $500,000.
- The maximum value is $900,000.
- All values between $500,000 and $900,000 are equally possible.

These conditions match those of the uniform distribution. The uniform distribution has two parameters: the minimum ($500,000) and the maximum ($900,000). You would enter these values as the parameters of the uniform distribution in Crystal Ball.

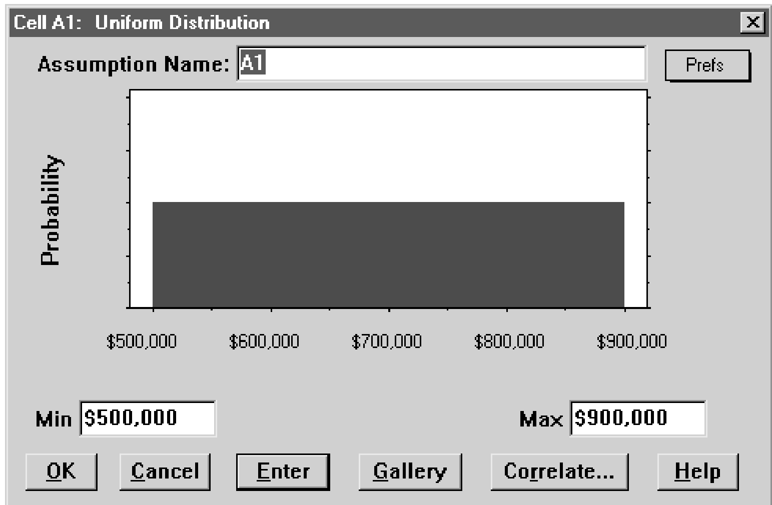


Figure 2.4 Uniform distribution

The distribution in Figure 2.4 shows that all values between $500,000 and $900,000 are equally possible.

Example two

A manufacturer determines that he must receive 10% over production costs—or a minimum of $3 per unit—to make the manufacturing effort worthwhile. He also wants to set the maximum price for the product at $6 per unit, so that he can gain a sales advantage by offering the product for less than his nearest competitor. All values between $3 and $6 per unit have the same likelihood of being the actual product price.

The first step in selecting a probability distribution is to match your data with a distribution's conditions.

Checking the uniform distribution:

- The minimum value is $3 per unit.
- The maximum value is $6 per unit.
- All values between $3 and $6 are equally possible.

You would enter these values in Crystal Ball to produce a uniform distribution showing that all values from $3 to $6 occur with equal likelihood.

Normal distribution

The normal distribution is the most important distribution in probability theory because it describes many natural phenomena, such as people's IQs or heights. Decision-makers can use the normal distribution to describe uncertain variables, such as the inflation rate or the future price of gasoline.

Crystal Ball Note: *You can use the "Parms" or Alternate Parameter menu found in the upper right corner of the dialog to specify a normal distribution using different combinations of percentiles. For more information, see "Alternate parameter sets" on page 136.*

Conditions

The three conditions underlying the normal distribution are:

- Some value of the uncertain variable is the most likely (the mean of the distribution).
- The uncertain variable could as likely be above the mean as it could be below the mean (symmetrical about the mean).
- The uncertain variable is more likely to be in the vicinity of the mean than far away.

Statistical Note: *Of the values of a normal distribution, approximately 68% are within 1 standard deviation on either side of the mean. The standard deviation is the square root of the average squared distance of values from the mean.*

The following example shows a real-world situation that matches (or closely approximates) these conditions. A more detailed discussion of mean and standard deviation follows this example.

Example

The normal distribution can be used to describe future inflation. You believe that 4% is the most likely rate. You are willing to bet that the inflation rate could as likely be above 4% as it could be below. You are also willing to bet that the inflation rate has a 68% chance of falling somewhere within 2% of the 4% rate. That is, you estimate there is approximately a two-thirds chance that the rate of inflation will be between 2% and 6%.

The first step in selecting a probability distribution is to match your data with a distribution's conditions. Checking the normal distribution:

- The mean inflation rate is 4%.
- The inflation rate could as likely be above or below 4%.
- The inflation rate is more likely to be close to 4% than far away. In fact, there is approximately a 68% chance that the rate will lie within 2% of the mean rate of 4%.

These conditions match those of the normal distribution.

The normal distribution uses two parameters—mean and standard deviation. The conditions outlined in the example contain the values for these parameters—a mean of 0.04 (4%) and a standard deviation of 0.02 (2%). You would enter these values as the parameters of the normal distribution in Crystal Ball.

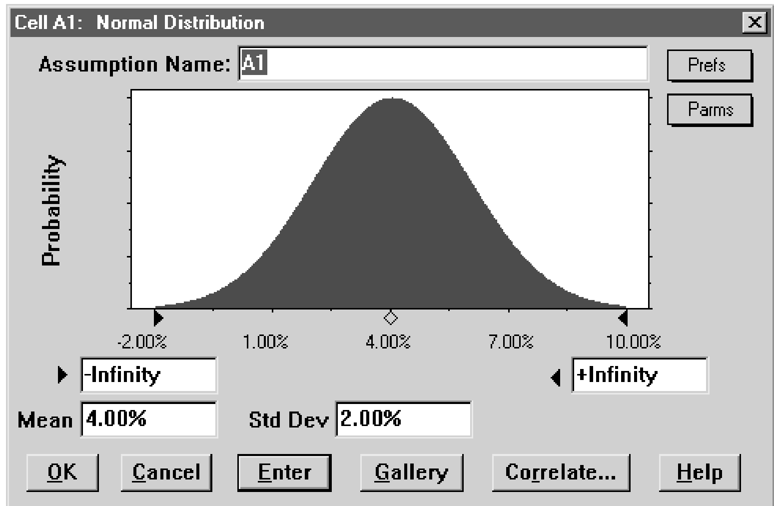


Figure 2.5 Normal distribution

The distribution in Figure 2.5 shows the probability of the inflation rate being a particular percentage.

Calculating standard deviation

This section explains three methods for calculating the standard deviation. However, first you must compute the mean of your data. The mean is the most expected value of the variable. You might already have an intuitive feel for the most expected value, or you can calculate the mean (determine the average) by adding the values in your data and dividing their sum by the number of values.

Crystal Ball automatically displays the normal curve using three standard deviations about the mean (or approximately 99.7% of the values).

You can calculate the standard deviation using one of the following three methods:

1. **If historical data are available, calculate the standard deviation using the basic formula (see page 339).**
2. **Find a centered interval about the mean that contains approximately 68% of the values as described in the normal distribution example. The distance from one end point of the interval to the mean is one standard deviation.**
3. **Drag the diamond truncation grabber in the Normal Distribution dialog to help you determine the standard deviation of your data, as described below:**

   Pick any point on the normal distribution curve and enter a value based on your best estimate. For example, suppose you are using the distribution to predict the interest rate after one year. You estimate the rate to be 0.09 (9%) and enter that number for the mean.

   You then estimate that a 6% interest rate is half as likely to occur as a 9% rate. You drag the diamond grabber to a point on the curve that appears to be half as high as the mean and enter 0.06 (6%) in the dialog that appears. Crystal Ball computes the standard deviation for you by comparing the two values and their locations on the curve.

Triangular distribution

The triangular distribution describes a situation where you know the minimum, maximum, and most likely values to occur. For example, you could describe the number of cars sold per week when past sales show the minimum, maximum, and usual number of cars sold.

Crystal Ball Note: *You can use the "Parms" or Alternate Parameter menu found in the upper right corner of the dialog to specify a triangular distribution using different combinations of percentiles. For more information, see "Alternate parameter sets" on page 136.*

Conditions

The three conditions underlying the triangular distribution are:

- The minimum number of items is fixed.
- The maximum number of items is fixed.
- The most likely number of items falls between the minimum and maximum values, forming a triangular-shaped distribution, which shows that values near the minimum and maximum are less likely to occur than those near the most likely value.

Example one

An owner needs to describe the amount of gasoline sold per week by his filling station. Past sales records show that a minimum of 3,000 gallons to a maximum of 7,000 gallons are sold per week, with most weeks showing sales of 5,000 gallons.

The first step in selecting a probability distribution is to match your data with a distribution's conditions. Checking the triangular distribution:

- The minimum number of gallons is 3,000.
- The maximum number of gallons is 7,000.
- The most likely number of gallons (5,000) falls between 3,000 and 7,000, forming a triangle.

These conditions match those of the triangular distribution.

The triangular distribution has three parameters—minimum, maximum, and likeliest. The conditions outlined in this example contain the values for these parameters—3,000 (minimum), 7,000 (maximum), and 5,000 (likeliest). You would enter these values as the parameters of the triangular distribution in Crystal Ball. You can type in the most likely value or select it with the triangle grabber.

The distribution shown in Figure 2.6 shows the probability of *x* number of gallons being sold per week.

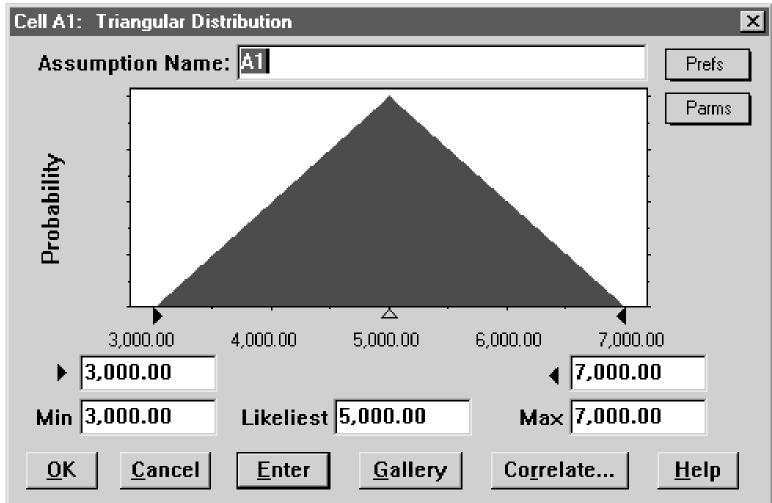


Figure 2.6 Triangular distribution

Example two

The triangular distribution also could be used to approximate a computer-controlled inventory situation. The computer is programmed to keep an ideal supply of 25 items on the shelf, to not let inventory ever drop below 10 items, and not ever let it rise above 30 items.

Check the triangular distribution conditions:

- The minimum inventory is 10 items.
- The maximum inventory is 30 items.
- The ideal level most frequently on the shelf is 25 items.

These conditions match those of the triangular distribution.

The result would be a distribution showing the probability of *x* number of items in inventory.

Other commonly used distributions

Binomial distribution

The binomial distribution describes the number of times a particular event occurs in a fixed number of trials, such as the number of heads in 10 flips of a coin or the number of defective items in 50 items.

Conditions

The parameters for this distribution are the number of trials (n) and the probability (p). The three conditions underlying the binomial distribution are:

- For each trial, only two outcomes are possible.
- The trials are independent. What happens in the first trial does not affect the second trial, and so on.
- The probability of an event occurring remains the same from trial to trial.

Example one

You want to describe the number of defective items in a total of 50 manufactured items, 7% of which (on the average) were found to be defective during preliminary testing.

The first step in selecting a probability distribution is to match your data with a distribution's conditions. Checking the binomial distribution:

Glossary Term: **trial**—(as used to describe a parameter in certain probability distributions)—The number of times a given experiment is repeated.

- There are only two possible outcomes—the manufactured item is either good or defective.
- The **trials** (50) are independent of one another. Any given manufactured item is either defective or not, independent of the condition of any of the others.
- The probability of a defective item (7%) is the same each time an item is tested.

These conditions match those of the binomial distribution.

The parameters for the binomial distribution are trials and probability. In example one, the values for these parameters are 50 (trials) and 0.07 (7% probability of producing defective items). You would enter these values to specify the parameters of the binomial distribution in Crystal Ball.

Statistical Note: *The word "trials," as used to describe a parameter of the binomial distribution, is different from "trials" as it is used when running a simulation in Crystal Ball. Binomial distribution trials describe the number of times a given experiment is repeated, (flipping a coin 50 times would be 50 binomial trials). A simulation trial describes a set of 50 coin flips, (10 simulation trials would simulate flipping 50 coins 10 times).*

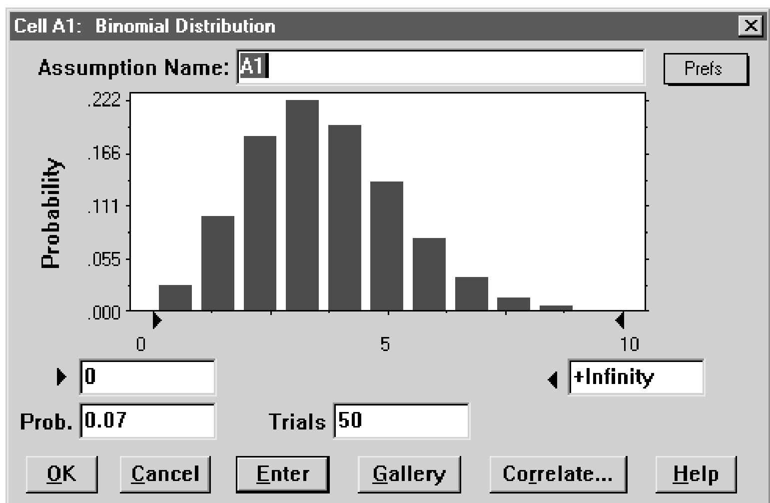


Figure 2.7 Binomial distribution

The distribution illustrated in Figure 2.7 shows the probability of producing *x* number of defective items.

Example two

A company's sales manager wants to describe the number of people who prefer the company's product. The manager conducted a survey of 100 consumers and determined that 60% prefer the company's product over the competitor's.

Again, the conditions fit the Binomial distribution with two important values—100 (trials) and 0.6 (60% probability of success). These values specify the parameters of the binomial distribution in Crystal Ball. The result would be a distribution of the probability that *x* number of people prefer the company's product.

Poisson distribution

The Poisson distribution describes the number of times an event occurs in a given area of opportunity, such as the number of telephone calls per minute or the number of errors per page in a document.

Conditions

The three conditions underlying the Poisson distribution are:

- The number of possible occurrences in any area of opportunity is unlimited.
- The occurrences are independent. The number of occurrences in one area of opportunity does not affect the number of occurrences in other areas.
- The average number of occurrences must remain the same from area to area.

Example one

An aerospace company wants to determine the number of defects per 100 square yards of carbon fiber material when the defects occur an average of 8 times per 100 square yards.

The first step in selecting a probability distribution is to match your data with a distribution's conditions. Checking the Poisson distribution:

- Any number of defects is possible within 100 square yards.
- The occurrences are independent of one another. The number of defects in the first 100 square yards does not affect the number of defects in the second 100 square yards.
- The average number of defects (8) remains the same for each 100 square yards.

These conditions match those of the Poisson distribution.

The Poisson distribution has only one parameter—rate. In this example, the value for this parameter is 8 (defects). You would enter this value to specify the parameter of the Poisson distribution in Crystal Ball.

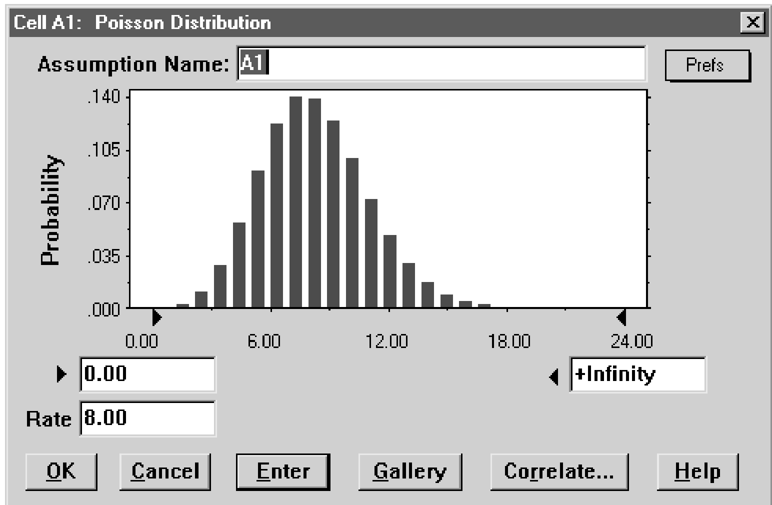


Figure 2.8 Poisson distribution

The distribution illustrated in Figure 2.8 shows the probability of observing *x* number of defects in 100 square yards of the carbon fiber material.

Statistical Note: *The size of the area of opportunity that the rate applies to, 100 square yards in this example, has no bearing on the probability distribution; the rate is the only key factor. If needed for modeling a situation, information on the size of the area of opportunity must be encoded in your spreadsheet formulas.*

Example two

A travel agency wants to describe the number of calls it receives in ten minutes. The average number of calls in 10 minutes is about 35.

Again, you begin by identifying and entering the values to set the parameters of the Poisson distribution in Crystal Ball. In this example, the conditions show one important value—35 calls or a rate of 35. The result would be a distribution showing the probability of receiving *x* number of calls in ten minutes.

Geometric distribution

The geometric distribution describes the number of trials until the first successful occurrence, such as the number of times you need to spin a roulette wheel before you win.

Conditions

The three conditions underlying the geometric distribution are:

- The number of trials is not fixed.
- The trials continue until the first success.
- The probability of success is the same from trial to trial.

Example one

If you are drilling for oil and want to describe the number of dry wells you would drill before the next producing well, you would use the geometric distribution. Assume that in the past you have hit oil about 10% of the time.

The first step in selecting a probability distribution is to match your data with a distribution's conditions. Checking the geometric distribution:

- The number of trials (dry wells) is unlimited.
- You continue to drill wells until you hit the next producing well.
- The probability of success (10%) is the same each time you drill a well.

These conditions match those of the geometric distribution.

The geometric distribution has only one parameter—probability. In this example, the value for this parameter is 0.10, representing the 10% probability of discovering oil. You would enter this value as the parameter of the geometric distribution in Crystal Ball.

The distribution illustrated in Figure 2.9 shows the probability of *x* number of wells drilled before the next producing well.

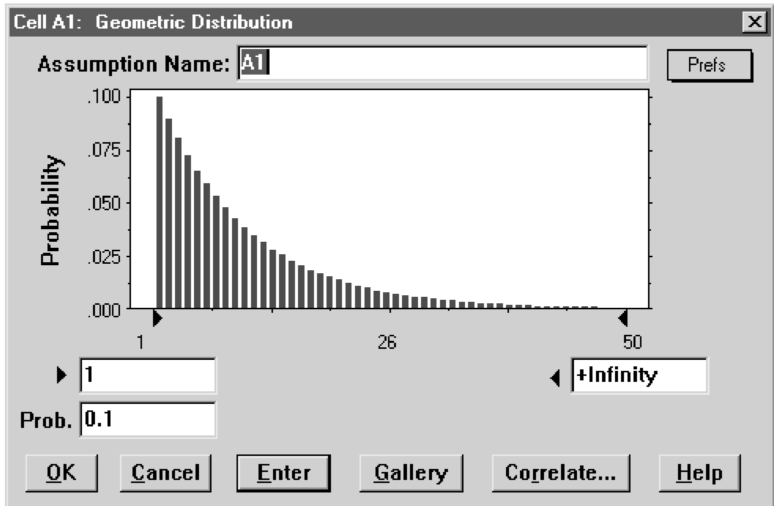


Figure 2.9 Geometric distribution

Example two

An insurance company wants to describe the number of claims received until a “major” claim arrives. Records show that 6% of the submitted claims are equal in dollar amount to all the other claims combined.

Again, identify and enter the parameter values for the geometric distribution in Crystal Ball. In this example, the conditions show one important value—a 0.06 (6%) probability of receiving that “major” claim. The result would be a distribution showing the probability of *x* number of claims occurring between “major” claims.

Hypergeometric distribution

The hypergeometric distribution is similar to the binomial distribution; both describe the number of times a particular event occurs in a fixed number of trials. The difference is that binomial distribution trials are independent, while hypergeometric distribution trials change the probability for each subsequent trial and are called “trials without replacement.”

For example, suppose a box of manufactured parts is known to contain some defective parts. You choose a part from the box, find it is defective, and remove the part from the box. If you choose another part from the box, the probability that it is defective is somewhat lower than for the first part because you have removed a defective part. If you had replaced the defective part, the probabilities would have remained the same, and the process would have satisfied the conditions for a binomial distribution.

Conditions

The conditions underlying hypergeometric distribution are:

- The total number of items or elements (the population size) is a fixed number—a finite population.
- The sample size (the number of trials) represents a portion of the population.
- The known initial probability of success in the population changes slightly after each trial.

Statistical Note: *The word "trials," as used to describe a parameter of the hypergeometric distribution, is different from "trials" as it is used when running a simulation in Crystal Ball. Hypergeometric distribution trials describe the number of times a given experiment is repeated, (removing 20 manufactured parts from a box would be 20 hypergeometric trials). A simulation trial describes the removing of 20 parts, (10 simulation trials would simulate removing 20 manufactured parts 10 times).*

Example one

You want to describe the number of consumers in a fixed population who prefer Brand X. You are dealing with a total population of 40 consumers, of which 30 prefer Brand X and 10 prefer Brand Y. You survey 20 of those consumers.

The first step in selecting a probability distribution is to match your data with a distribution's conditions. Checking the hypergeometric distribution:

- The population size (40) is fixed.
- The sample size (20 consumers) represents a portion of the population.
- Initially, 30 of 40 consumers (75%) preferred Brand X. This percentage changes each time you question one of the 20 consumers, depending on the preference of the previous consumer.

The conditions in this sample match those of the hypergeometric distribution.

The three parameters of this distribution are population size, trials, and initial probability (denoted in Crystal Ball as "Prob"). The conditions outlined in this example contain the values for these parameters—a population size of 40, sample size (trials) of 20, and initial probability of 0.75 (75%) that consumers will prefer Brand X. You would enter these values as the parameters of the hypergeometric distribution in Crystal Ball.

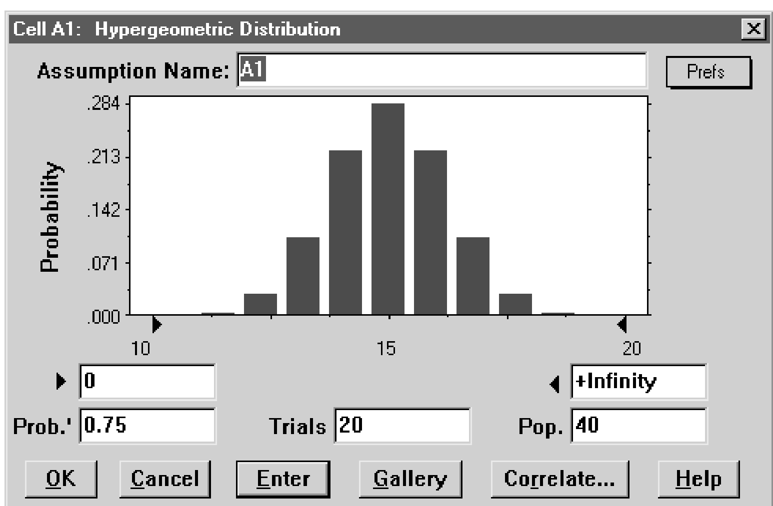


Figure 2.10 Hypergeometric distribution

The distribution illustrated in Figure 2.10 shows the probability that *x* number of consumers prefer Brand X.

Example two

The U.S. Department of the Interior wants to describe the movement of wild horses in Nevada. Researchers in the department travel to a particular area in Nevada to tag 100 horses in a total population of 1,000. Six months later the researchers return to the same area to find out how many horses remained in the area. The researchers look for tagged horses in a sample of 200.

Check the data against the conditions of the hypergeometric distribution. The parameter values for the hypergeometric distribution in Crystal Ball are the population size of 1,000, sample size of 200, and 0.1 (10%) initial probability (100 out of 1,000) of finding tagged horses. The result would be a distribution showing the probability of observing *x* number of tagged horses.

Lognormal distribution

Glossary Term:
skewed, positively—A distribution in which most of the values occur at the lower end of the range.

The lognormal distribution is widely used in situations where values are **positively skewed**, for example in financial analysis for security valuation or in real estate for property valuation.

Stock prices are usually positively skewed, rather than normally (symmetrically) distributed. Stock prices exhibit this trend because they cannot fall below the lower limit of zero but might increase to any price without limit.

Glossary Term:
skewness—The measure of the degree of deviation of a curve from the norm of a symmetric distribution. The greater the degree of skewness, the more points of the curve lie to either side of the peak of the curve. A normal distribution curve, having no skewness, is symmetrical.

Similarly, real estate prices illustrate positive **skewness** since property values cannot become negative.

Crystal Ball Note: *You can use the "Parms" or Alternate Parameter menu found in the upper right corner of the dialog to specify a lognormal distribution using different combinations of percentiles. For more information, see "Alternate parameter sets" on page 136.*

Conditions

The three conditions underlying the lognormal distribution are:

- The uncertain variable can increase without limits but cannot fall below zero.
- The uncertain variable is positively skewed with most of the values near the lower limit.
- The natural logarithm of the uncertain variable yields a normal distribution.

Generally, if the coefficient of variability is greater than 30%, use a lognormal distribution. Otherwise, use the normal distribution.

Example

The lognormal distribution can be used to model the price of a particular stock. You purchase a stock today at $50. You expect that the stock will be worth $70 at the end of the year. If the stock price drops at the end of the year, rather than appreciating, you know that the lowest value it can drop to is $0.

On the other hand, the stock could end up with a price much higher than expected, thus implying no upper limit on the rate of return. In summary, your losses are limited to your original investment, but your gains are unlimited.

Using historical data, you can determine that the standard deviation of the stock's price is $12.

Statistical Note: *If you have historical data available with which to define a lognormal distribution, it is important to calculate the mean and standard deviation of the logarithms of the data and then enter these log parameters using the Parms (Alternate Parameters) menu. Calculating the mean and standard deviation directly on the raw data will not give you the correct lognormal distribution. Alternately, use this distribution fitting feature described on page 138.*

The first step in selecting a probability distribution is to match your data with a distribution's conditions. Checking the lognormal distribution:

- The price of the stock is unlimited at the upper end but cannot drop below $0.
- The distribution of the stock price is positively skewed.
- Mathematically, it can be shown that the natural logarithm of the stock price yields a normal distribution.

These conditions match those of the lognormal distribution.

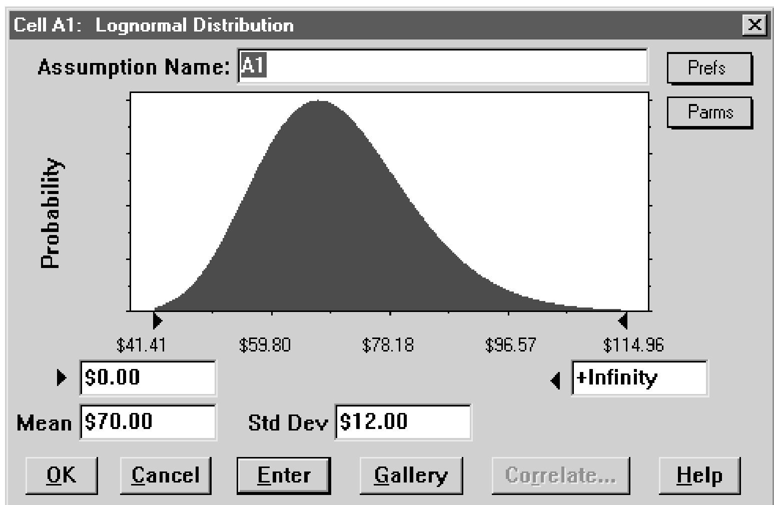


Figure 2.11 Lognormal distribution

In the lognormal distribution shown in Figure 2.11, the mean parameter is set at $70.00 and the standard deviation set at $12.00. This distribution shows the probability that the stock price will be $*x*.

Lognormal parameter sets

By default, the arithmetic mean and standard deviation are used with the lognormal distribution. For applications where historical data is available, it is more appropriate to use the logarithmic mean and standard deviation or the geometric mean and standard deviation. These options are available from the alternate parameter menu in the upper right corner of the dialog, as shown in Figure 2.12

.

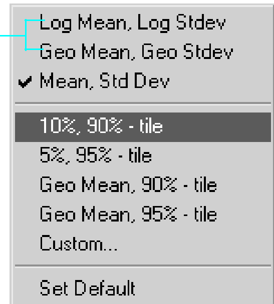


Figure 2.12 Lognormal alternate parameters menu

For more information about this menu, see “Alternate parameter sets” on page 136.

Exponential distribution

The exponential distribution is widely used to describe events recurring at random points in time, such as the time between failures of electronic equipment or the time between arrivals at a service booth. It is related to the Poisson distribution, which describes the number of occurrences of an event in a given interval of time.

An important characteristic of the exponential distribution is the “memoryless” property, which means that the future lifetime of a given object has the same distribution, regardless of the time it existed. In other words, time has no effect on future outcomes.

Crystal Ball Note: *You can use the “Parms” or Alternate Parameter menu found in the upper right corner of the dialog to specify an exponential distribution using different combinations of percentiles. For more information, see “Alternate parameter sets” on page 136.*

Conditions

The two conditions underlying the exponential distribution are:

- The exponential distribution describes the amount of time between occurrences.

- The distribution is not affected by previous events.

Example one

A travel agency wants to describe the time between incoming telephone calls when the calls are averaging about 35 every 10 minutes. This same example was used for the Poisson distribution to describe the number of calls arriving every 10 minutes.

The first step in selecting a probability distribution is to match your data with a distribution's conditions. Checking the exponential distribution:

- The travel agency wants to describe the time between successive telephone calls.
- The phone calls are not affected by previous history. The probability of receiving 35 calls every 10 minutes remains the same.

The conditions in this example match those of the exponential distribution.

The exponential distribution has only one parameter—rate. The conditions outlined in this example contain the value for this parameter—35 (calls) every minute or a rate of 35. You would enter this value to set the parameter of the exponential distribution in Crystal Ball.

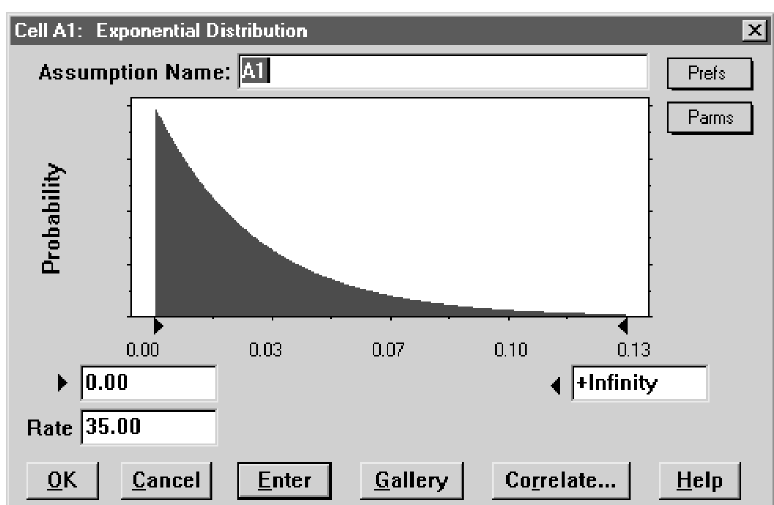


Figure 2.13 Exponential distribution

The distribution shown in Figure 2.13 shows the probability that *x* number of time units (10 minutes in this case) will pass between calls.

Example two

A car dealer needs to know the amount of time between customer drop-ins at his dealership so that he can staff the sales floor more efficiently. The car dealer knows an average of 6 customers visit the dealership every hour.

Checking the exponential distribution:

- The car dealer wants to describe the time between successive customer drop-ins.
- The probabilities of customer drop-ins remain the same from hour to hour.

These conditions fit the exponential distribution.

The resulting distribution would show the probability that *x* number of hours will pass between customer visits.

Weibull distribution (also Rayleigh distribution)

The Weibull distribution describes data resulting from life and fatigue tests. It is commonly used to describe failure time in reliability studies, and the breaking strengths of materials in reliability and quality control tests. Weibull distributions are also used to represent various physical quantities, such as wind speed.

The Weibull distribution is a family of distributions that can assume the properties of several other distributions. For example, depending on the shape parameter you define, the Weibull distribution can be used to model the exponential and Rayleigh distributions, among others.

The Weibull distribution is very flexible. When the Weibull shape parameter is equal to 1.0, the Weibull distribution is identical to the exponential distribution. The Weibull location parameter lets you set up an exponential distribution to start at a location other than 0.0. When the shape parameter is less than

1.0, the Weibull distribution becomes a steeply declining curve. A manufacturer might find this effect useful in describing part failures during a burn-in period.

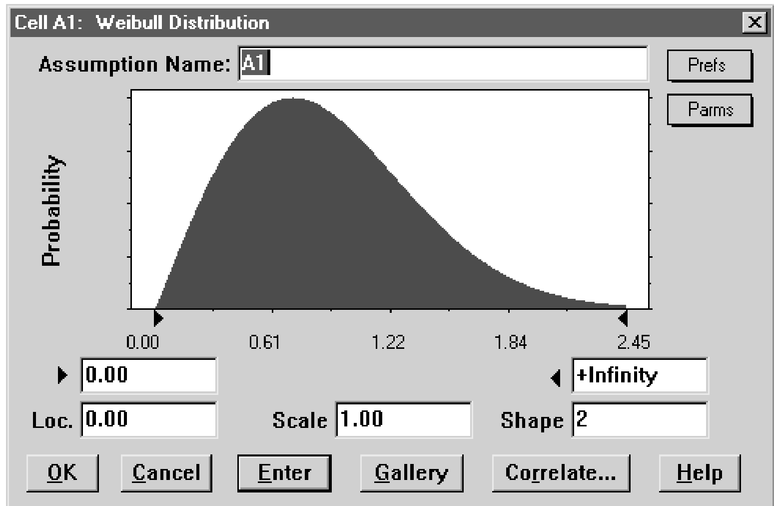


Figure 2.14 Weibull distribution

When the shape parameter is equal to 2.0, a special form of the Weibull distribution, called the Rayleigh distribution, results. A researcher might find the Rayleigh distribution useful for analyzing noise problems in communication systems or for use in reliability studies.

Calculating parameters

There are three standard parameters for the Weibull distribution: location, scale, and shape. The location parameter is the lower bound for the variable.

The shape parameter is a number greater than 0, usually a small number less than 10. When the shape parameter is less than 3, the distribution becomes more and more positively skewed until it starts to resemble an exponential distribution (shape < 1). At 3.25, the distribution is symmetrical, and above that, the distribution becomes more narrow and negatively skewed.

After you select the location and shape parameter, you can estimate the scale parameter. The larger the scale, the larger the width of the distribution.

To calculate a more exact scale, estimate the mean and use the equation:

$$\alpha = \frac{\text{mean} - L}{\Gamma\left(1 + \frac{1}{\beta}\right)}$$

Where α is the scale, β is the shape, *L* is the location, and Γ is the gamma function. You can use the Excel GAMMALN function and Excel Solver to help you calculate this parameter.

Statistical Note: *For this distribution, there is a 63% probability that x falls between* α *and* α*+L.*

Or estimate the mode and use the equation:

$$\alpha = \frac{\text{mode} - L}{\left(1 - \frac{1}{\beta}\right)^{\frac{1}{\beta}}}$$

Where α is the scale, β is the shape, and *L* is the location.

Crystal Ball Note: *You can use the "Parms" or Alternate Parameter menu found in the upper right corner of the dialog to specify a Weibull distribution using different combinations of percentiles. For more information, see "Alternate parameter sets" on page 136.*

Example

A lawn mower company is testing its gas-powered, self-propelled lawn mowers. They run 20 mowers, and keep track of how many hours each mower runs until its first breakdown. They use a Weibull distribution to describe the number of hours until the first failure.

Beta distribution

The beta distribution is a very flexible distribution commonly used to represent variability over a fixed range. One of the more important applications of the beta distribution is its use as a conjugate distribution for the parameter of a Bernoulli distribution. In this application, the beta distribution is used to represent the uncertainty in the probability of occurrence of an event. It is also used to describe empirical data and predict the random behavior of percentages and fractions.

The value of the beta distribution lies in the wide variety of shapes it can assume when you vary the two parameters, alpha and beta. If the parameters are equal, the distribution is symmetrical. If either parameter is 1 and the other parameter is greater than 1, the distribution is J-shaped. If alpha is less than beta, the distribution is said to be positively skewed (most of the values are near the minimum value). If alpha is greater than beta, the distribution is negatively skewed (most of the values are near the maximum value). Because the beta distribution is very complex, the methods for determining the parameters of the distribution are beyond the scope of this manual. However, the example below should help define it. For more information about the beta distribution and Bayesian statistics, refer to the texts listed in the bibliography.

Crystal Ball Note: *You can use the "Parms" or Alternate Parameter menu found in the upper right corner of the dialog to specify a beta distribution using different combinations of percentiles. For more information,* *see "Alternate parameter sets" on page 136.*

Conditions

The two conditions underlying the beta distribution are:

- The uncertain variable is a random value between 0 and a positive value.
- The shape of the distribution can be specified using two positive values.

Example

A company that manufactures electrical devices for custom orders wants to model the reliability of devices it produces. After analyzing the empirical data, the company knows that it can use the beta distribution to describe the reliability of the devices if the parameters are alpha = 10 and beta = 2.

The first step in selecting a probability distribution is to match your data with a distribution's conditions. Checking the beta distribution:

- The reliability rate is a random value somewhere between 0 and 1.
- The shape of the distribution can be specified using two positive values: 10 and 2.

These conditions match those of the beta distribution.

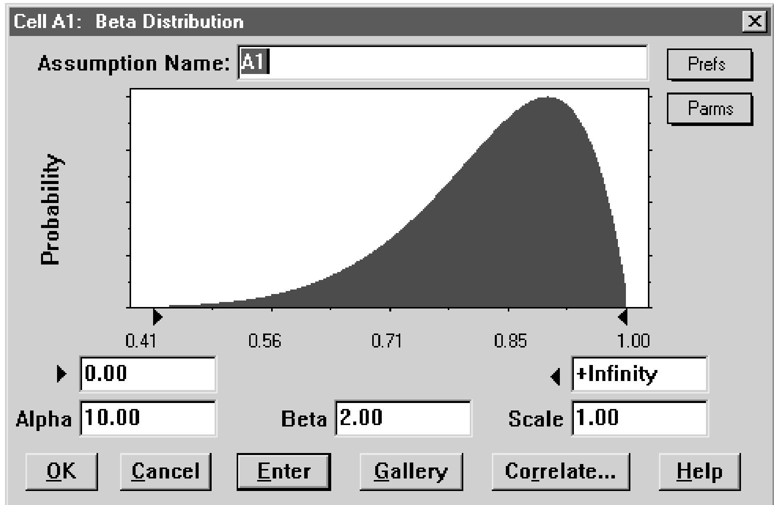


Figure 2.15 Beta distribution

Figure 2.15 shows the beta distribution with the alpha parameter set to 10, the beta parameter set to 2. The reliability rate of the devices will be x.

Custom distribution

With Crystal Ball you can use the custom distribution to represent a unique situation that cannot be described using other distribution types—you can describe a series of single values, discrete ranges, or continuous ranges. This section uses real-world examples to describe the custom distribution. Follow the step-by-step instructions carefully in each example.

Since it is easier to understand how the custom distribution works with a hands-on example, you might want to start Crystal Ball and use it to follow the examples. To follow the custom distribution examples, first create a new Excel spreadsheet and enter any value in cell A1.

Example one

Before beginning example one, open the Custom Distribution dialog as follows:

1. **Click cell A1.**

2. **Select Cell > Define Assumptions.**

   The Distribution Gallery dialog appears.

3. **Click the Custom Distribution.**

4. **Click on OK.**

   Crystal Ball displays the Custom Distribution dialog as in Figure 2.16.

Using the custom distribution, a company can describe the probable retail cost of a new product. The company decides the cost could be $5, $8, or $10. In this example, you will use the custom distribution to describe a series of single values.

To enter the parameters of this custom distribution:

1. **Type 5 in the Value field and click Enter.**

   Since you do not specify a probability, Crystal Ball defaults to a **relative probability** of 1.00 for the value 5. The total relative probability appears as 1.00 on the right side of the screen, and a single value bar displays the value 5.00.

Glossary Term:
Relative Probability — (also referred to as **Relative Frequency**) - A value, not necessarily between 0 and 1, that indicates probability when considered in relation to other relative probabilities.

Statistical Note: *Relative probability means that the sum of the probabilities does not have to add up to 1. So the probability for a given value is meaningless by itself; it makes sense only in relation to the total relative probability. For example, if the total relative probability is 3 and the relative probability for a given value is 1, the value has a probability of 0.33.*

2. **Type 8 in the Value field.**

3. **Click on Enter.**

   Since you did not specify a probability, Crystal Ball defaults to a relative probability of 1.00 (displayed on the scale to the left of the Custom Distribution dialog) for the value 8. The total relative probability for the values 5 and 8 is now displayed as 2.00. A second value bar represents the value 8.

4. **Type 10 in the Value field.**

5. Click on Enter.

Crystal Ball displays a relative probability of 1.00 for the value 10 and a total relative probability of 3.00. A third single value bar represents the value 10.

Figure 2.16 shows the value bars for the values 5, 8, and 10 with a relative probability of 1.00 and a total relative probability of 3.00.

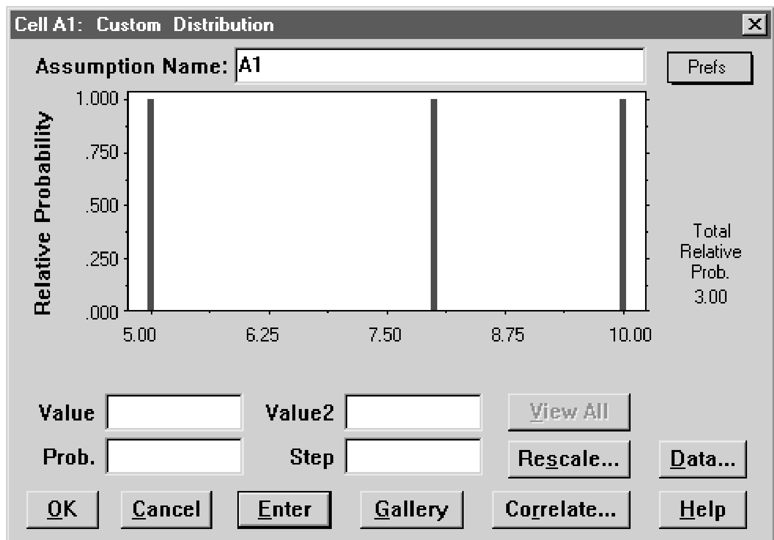


Figure 2.16 Single values

You can leave the total relative probability at 3.00 and run the simulation. However, to gain a better feel for the relationship among these three values, rescale the total relative probability to 1.00, as described in the following steps.

6. Click on Rescale.

The Rescale dialog appears.

7. Click on OK to rescale the total relative probability to 1.00.

Crystal Ball rescales each value to a relative probability of 0.33 on the left side of the screen and a total relative probability of 1.00.

You can check the probability for each value bar as described in the following steps:

8. **Click the value bar for the value 5.00.**

   The bar changes to a lighter shade. The Value field displays the value 5.00 and the Prob field displays a probability of 0.333333.

9. **Click the value bar for the value 8.00, the middle bar.**

   The Value field displays the value 8.00 and the Prob field displays a probability of 0.333333.

10. **Click the value bar for the value 10.00, the right bar.**

    The Value field displays the value 10 and the Prob field displays a probability of 0.333333.

While you are checking the probability for each bar, you can change the probability by typing a new one in the Prob field.

You can also change the probability of each value by dragging the bar, as described in the following steps:

11. **Click the top of the value bar you want to change and drag the bar up or down to the probability you want.**

    Crystal Ball changes the probability in the Prob field as you drag the bar.

Example two

Before beginning example two, clear the values entered in example one as follows:

1. **Click on Gallery to return to the Distribution Gallery.**
2. **Click on OK to return to the Custom Distribution dialog.**

   The previous values are cleared from the dialog.

In this example, you will use the custom distribution to describe a continuous range of values, since the unit cost can take on any value within the specified intervals.

To enter the first range of values:

3. **Type 5 in the Value field.**
4. **Press <Tab> and type 15 in the Value2 field.**
5. **Press <Tab> and type .75 in the Prob field.**

Crystal Ball Note: *You might enter the values in either order, but press Enter only after entering all three values.*

6. Click Enter.

Crystal Ball displays a total relative probability of 0.75, a continuous value bar for the range 5.00 to 15.00, and returns the cursor to the Value field.

To enter the second range of values:

7. Type 15 in the Value field.

8. Press <Tab> and type 20 in the Value2 field.

9. Press <Tab> and type .25 in the Prob field.

10. Click on Enter.

Crystal Ball displays a total relative probability of 1.00 and a continuous value bar for the range 15.00 to 20.00. Both ranges are now displayed in the Custom Distribution dialog.

You can change the probability and slope of a continuous range by dragging the value bar, as described in the following steps:

11. Click anywhere on the value bar for the range 15 to 20.

The value bar changes to a lighter shade.

12. Click on the upper right corner of the value bar for the range 15 to 20, and drag the right side down until it is approximately half as high as the left side.

As Figure 2.17 illustrates, Crystal Ball displays the ratio of the higher side to the lower side (Ratio of Sides) as 1.97. The total relative probability changes to reflect the changes in the value bar.

13. Click on the upper left corner of the value bar for the range 15 to 20, and drag the left side down to the level of the right side.

Crystal Ball displays a ratio of 1.00, and the total relative probability changes.

14. **Click on the center of the value bar for the range 15 to 20, and drag the value bar up to its previous position.**

    The value bar moves as one unit with the ratio of sides remaining at 1.00. Crystal Ball changes the total relative probability back to 1.00.

Drag the right side down until it is half as high as the left

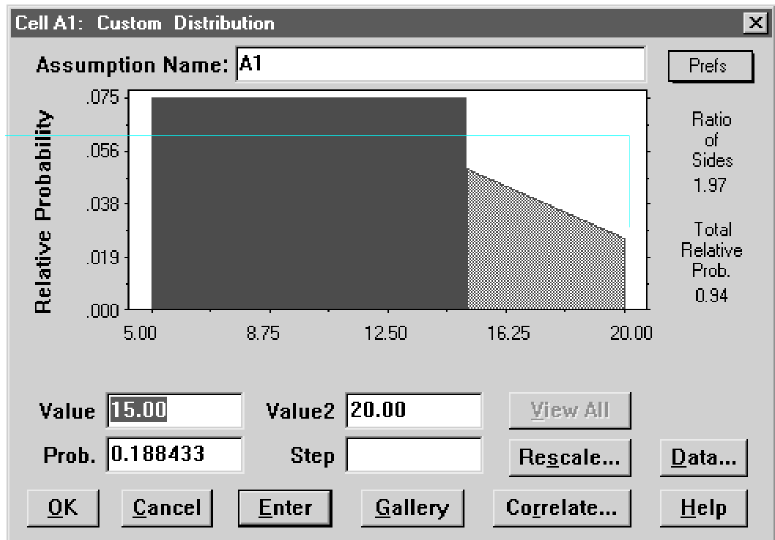


Figure 2.17 Continuous range of values

Example three

Before beginning example three, clear the values entered in example two as follows:

1. **Click on Gallery to return to the Distribution Gallery.**
2. **Click on OK to return to the Custom Distribution dialog.**

    The previous values are cleared from the dialog.

In example three, the same company decides the retail price of its new product has a 75% chance of being a whole dollar amount between $5 and $15 and a 25% chance of being a whole dollar amount between $16 and $20.

Crystal Ball Note: *The second range starts at $16 rather than $15 because, unlike continuous ranges, discrete ranges cannot touch.*

In this example, you will use the custom distribution to describe a discrete range of values, since the retail price is restricted to whole dollar amounts. The Step field on the Custom Distribution dialog lets you specify the distance or "step" between one discrete value and the next.

To enter the first range of values:

1. **Type 5 in the Value field.**
2. **Press <Tab> and type 15 in the Value2 field.**
3. **Press <Tab> and type .75 in the Prob field.**
4. **Press <Tab> and type 1 in the Step field.**

Crystal Ball Note: *You can enter any positive number in the Step field. For this example you enter 1, because the example specifies consecutive integers.*

5. **Click Enter.**

   Crystal Ball displays a total relative probability of 0.75 and a discrete value bar for the range 5 to 15. It also returns the cursor to the Value field.

To enter the second range of values,

1. **Type 16 in the Value field.**
2. **Press <Tab> and type 20 in the Value2 field.**
3. **Press <Tab> and type .25 in the Prob field.**
4. **Press <Tab> and type 1 in the Step field.**
5. **Click Enter.**

   Crystal Ball displays a total relative probability of 1.00 and a discrete value bar made up of a series of individual bars for the range 16 to 20. Figure 2.18 shows the discrete ranges with a step of 1.

You can change the probability of a discrete range by dragging the value bar as you did for the continuous range.

With Crystal Ball, you can enter single values, discrete ranges, or continuous ranges individually. You also can enter any combination of these three types in the same Custom Distribution dialog as long as you follow these guidelines:

ranges and single values cannot overlap one another; however, the ending value of one continuous range can be the starting value of another continuous range.

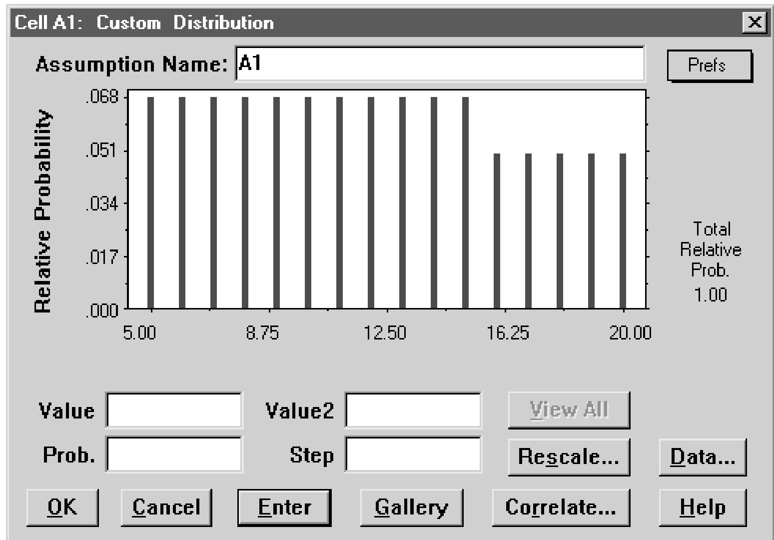
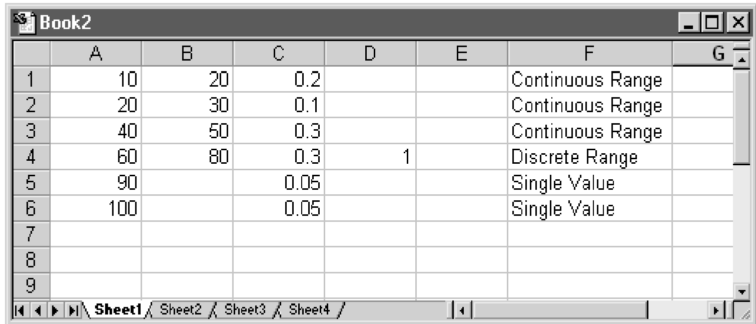


Figure 2.18 Discrete ranges

Example four

This example describes a special feature on the Custom Distribution dialog—the *Data* button, which lets you pull in numbers from specified cell ranges (grouped data) on the spreadsheet. This example is not a hands-on exercise, but the illustrations will guide you through the procedure. After you read this section, you can experiment with your own data by pulling in numbers from specified cell ranges on your spreadsheet.

In this example, the same company decides that the unit cost of the new product can vary widely. The company feels it has a 20% chance of being any number between \$10 and \$20, a 10% chance of being any number between \$20 and \$30, a 30% chance of being any number between \$40 and \$50, a 30% chance of being a whole dollar amount between \$60 and \$80, or there is a 5% chance the value will be either \$90 or \$100. All the values have been entered on the spreadsheet.



	A	B	C	D	E	F	G
1	10	20	0.2			Continuous Range	
2	20	30	0.1			Continuous Range	
3	40	50	0.3			Continuous Range	
4	60	80	0.3	1		Discrete Range	
5	90		0.05			Single Value	
6	100		0.05			Single Value	
7							
8							
9							

Figure 2.19 Four-column custom data range

Since the values are already on the spreadsheet, you can click on Data on the Custom Distribution dialog, then enter the cell range in the field, as in Figure 2.20.

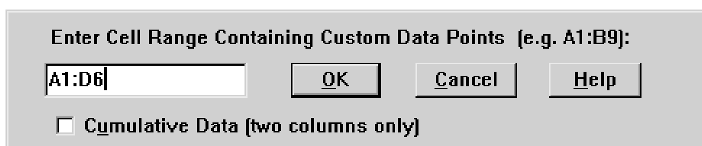


Figure 2.20 Custom Data dialog

In using cell-range data for the custom distribution, follow these guidelines:

- For a series of single values all with equal probabilities, use either one column or five or more columns of data.
- For a series of single values all with different probabilities, use a two-column format: the first column containing single values and the second column containing probabilities.
- For any mixture of continuous ranges, discrete ranges, or single values, use a three or four column format, as in Figure 2.19.
- For a connected series of sloping continuous ranges, use a three-column format: 1) the first column containing the common endpoints of the ranges, 2) the second column left empty, and 3) the third column containing the probabilities of the endpoints.

- For a connected series of continuous ranges specified using cumulative probabilities, use a two-column format: the common endpoints of the ranges in the first column and the cumulative probabilities in the second column. The first row in the cell range should contain the minimum value of the data and should have a cumulative probability of zero.

When you click on OK, Crystal Ball enters the values from the specified cell ranges into the custom distribution and displays the value bars for the specified ranges. In this example, the value bars for the specified ranges would not be completely visible on the screen. Crystal Ball solves this problem by providing a View All button and scroll bars in the Custom Distribution dialog. You can click on View All to view all the value bars for the specified ranges, as shown in Figure 2.21. When you view all, the View All button changes to View Part, allowing you to return to the previous view whenever you want. You can also use the scroll bar to look at any section of the screen.

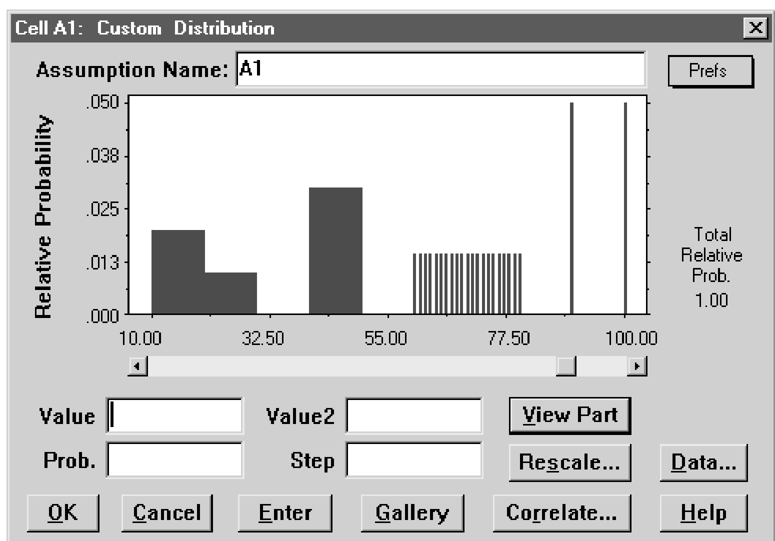


Figure 2.21 Custom data from spreadsheet

Important custom distribution notes

- If you want to delete a value you entered for the custom distribution, type a probability of 0 in the Prob field for the value you want to delete and click on Enter.
- You might change conditions of some of your data. Click any value bar to display its corresponding data boxes.
- If you want to increase the relative probability of a value by 1.00, enter the value again, and the relative probability for that value will increase by 1.00.
- If you want to replace the existing relative probability of a value, enter the value again with a new probability.

Less commonly used distribution types

The following five distributions types are also available: gamma, logistic, Pareto, extreme value, and negative binomial. To access these distributions, click on More in the Distribution Gallery dialog. Figure 2.22 illustrates the Distribution Gallery.

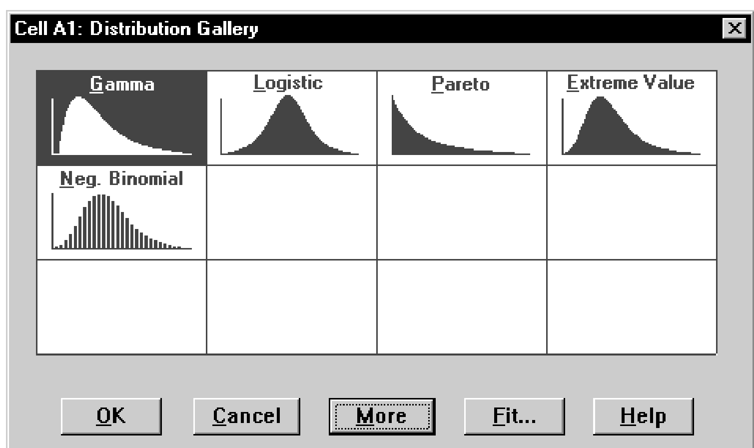


Figure 2.22 More distributions—Distribution Gallery dialog

Gamma distribution (also Erlang and chi-square)

The gamma distribution applies to a wide range of physical quantities and is related to other distributions: lognormal, exponential, Pascal, Erlang, Poisson, and chi-square. It is used in meteorological processes to represent pollutant concentrations and precipitation quantities. The gamma distribution is also used to measure the time between the occurrence of events when the event process is not completely random. Other applications of the gamma distribution include: inventory control, economics theory, and insurance risk theory.

Crystal Ball Note: *You can use the "Parms" or Alternate Parameter menu found in the upper right corner of the dialog to specify a gamma distribution using different combinations of percentiles. For more information, see "Alternate parameter sets" on page 136.*

Conditions

The gamma distribution is most often used as the distribution of the amount of time until the r th occurrence of an event in a Poisson process. When used in this fashion, the conditions underlying the gamma distribution are:

- The number of possible occurrences in any unit of measurement is not limited to a fixed number.
- The occurrences are independent. The number of occurrences in one unit of measurement does not affect the number of occurrences in other units.
- The average number of occurrences must remain the same from unit to unit.

Example one

A computer dealership knows that the lead time for re-ordering their most popular computer system is 4 weeks. Based upon an average demand of 1 unit per day, the dealership wants to model the number of business days it will take to sell 20 systems.

Checking the conditions of the gamma distribution:

- The number of possible customers demanding to buy the computer system is unlimited.
- The decisions of customers to buy the system are independent.
- The demand remains constant from week to week.

These conditions match those of the gamma distribution. (Note that in this example the dealership has made several simplifying assumptions about the conditions. In reality, the total number of computer purchasers is finite and some might have influenced the purchasing decisions of others).

The shape parameter is used to specify the *r*th occurrence of the Poisson event. In this example, you would enter 20 for the shape parameter (5 units per week times 4 weeks). The result is a distribution showing the probability that *x* number of business days will pass until the 20th system is sold.

Figure 2.23 illustrates the gamma distribution.

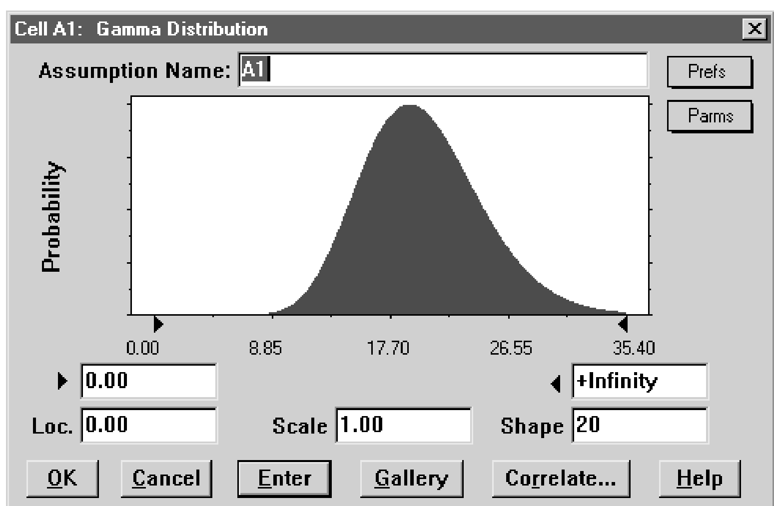


Figure 2.23 Gamma distribution

Example two

Suppose a particular mechanical system fails after receiving exactly 5 shocks to it from an external source. The total time to system failure, defined as the random time occurrence of the 5th shock, follows a gamma distribution with a shape parameter of 5.

Some characteristics of the gamma distribution:

- When shape = 1, gamma becomes a scalable exponential distribution.
- The sum of any two gamma-distributed variables is a gamma variable.
- The product of any two normally distributed variables is a gamma variable.
- If you have historical data that you believe fits the conditions of a gamma distribution, computing the parameters of the distribution is easy. First, compute the mean (μ) and variance (σ^2) of your historical data. Then compute the parameters themselves:

  shape parameter $= \mu^2 / \sigma^2$

  scale parameter $= \sigma^2 / \mu$

Chi-square and Erlang distributions

You can model two additional probability distributions, the chi-square and Erlang distributions, by adjusting the parameters entered in the Gamma Distribution dialog. To model these distributions, enter the parameters as described below:

Chi-square distribution

With parameters *N* and *S*, where *N* = number of degrees of freedom and $S = \text{scale}$, set your parameters as follows:

$$\text{shape} = \frac{N}{2} \qquad \text{scale} = 2S^2$$

The chi-square distribution is the sum of the squares of *N* normal variates.

Erlang distribution

The Erlang distribution is identical to the gamma distribution, except the shape parameter is restricted to integer values. Mathematically, the Erlang distribution is a convolution of *N* exponential distributions.

Logistic distribution

The logistic distribution is commonly used to describe growth (i.e., the size of a population expressed as a function of a time variable). It can also be used to describe chemical reactions and the course of growth for a population or individual.

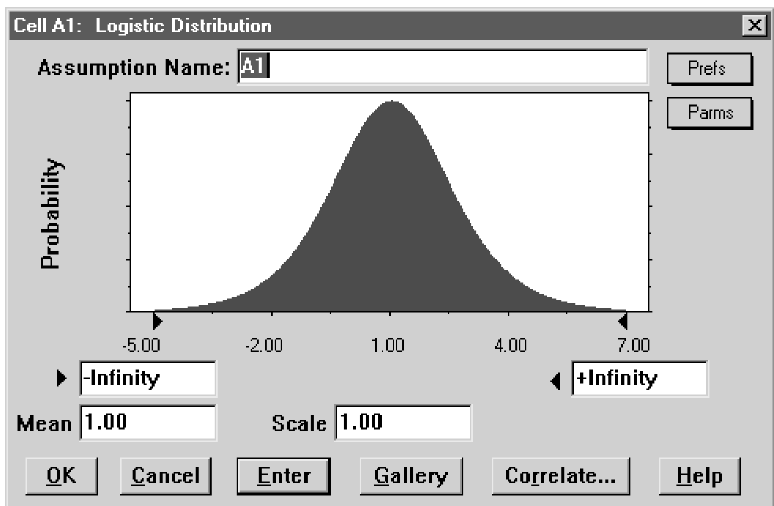


Figure 2.24 Logistic distribution

Calculating parameters

There are two standard parameters for the logistic distribution: mean and scale. The mean parameter is the average value, which for this distribution is the same as the mode, since this is a symmetrical distribution.

After you select the mean parameter, you can estimate the scale parameter. The scale parameter is a number greater than 0. The larger the scale parameter, the greater the variance.

To calculate a more exact scale, you can estimate the variance and use the equation:

$$\alpha = \sqrt{\frac{3 \cdot \text{variance}}{\pi^2}}$$

where α is the scale parameter.

Crystal Ball Note: *You can use the "Parms" or Alternate Parameter menu found in the upper right corner of the dialog to specify a logistic distribution using different combinations of percentiles. For more information, see "Alternate parameter sets" on page 136.*

Pareto distribution

The Pareto distribution, as shown in Figure 2.25, is widely used for the investigation of distributions associated with such empirical phenomena as city population sizes, the occurrence of natural resources, the size of companies, personal incomes, stock price fluctuations, and error clustering in communication circuits.

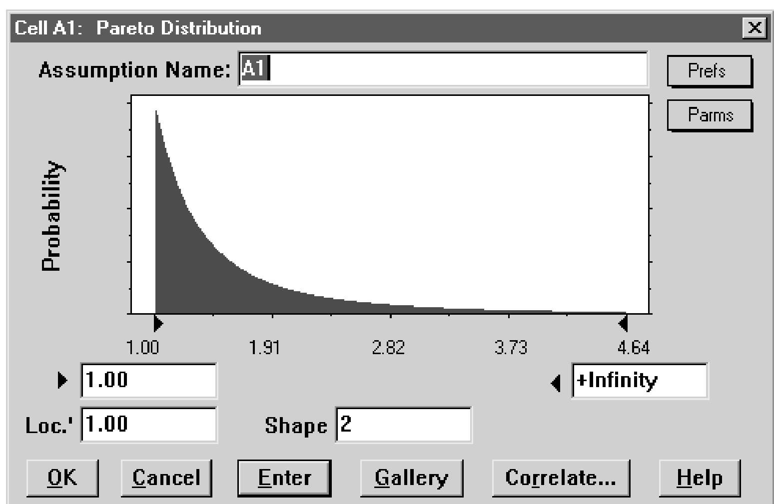


Figure 2.25 Pareto distribution

Calculating parameters

There are two standard parameters for the Pareto distribution: location and shape. The location parameter is the lower bound for the variable.

After you select the location parameter, you can estimate the shape parameter. The shape parameter is a number greater than 0, usually greater than 1. The larger the shape parameter, the smaller the variance and the thicker the right-tail of the distribution appears.

To calculate a more exact shape, you can estimate the mean and use the equation (for shapes greater than 1):

$$\text{mean} = \frac{\beta \cdot L}{\beta - 1}$$

where β is the shape parameter and L is the location parameter. You can use Excel Solver to help you calculate this parameter, setting the constraint of $\beta > 1$.

Or estimate the variance and use the equation (for shapes greater than 2):

$$\text{variance} = \frac{\beta \cdot L^2}{(\beta - 2)(\beta - 1)^2}$$

where β is the shape parameter and L is the location parameter. You can use Excel Solver to help you calculate this parameter, setting the constraint of $\beta > 2$.

Crystal Ball Note: *You can use the "Parms" or Alternate Parameter menu found in the upper right corner of the dialog to specify a Pareto distribution using different combinations of percentiles. For more information, see "Alternate parameter sets" on page 136.*

Extreme value distribution

The extreme value distribution (Type 1), as shown in Figure 2.26, is commonly used to describe the largest value of a response over a period of time: flood flows, rainfall, and earthquakes. Other applications include the breaking strengths of materials, construction design, and aircraft loads and tolerances. The extreme value distribution is also known as the Gumbel distribution.

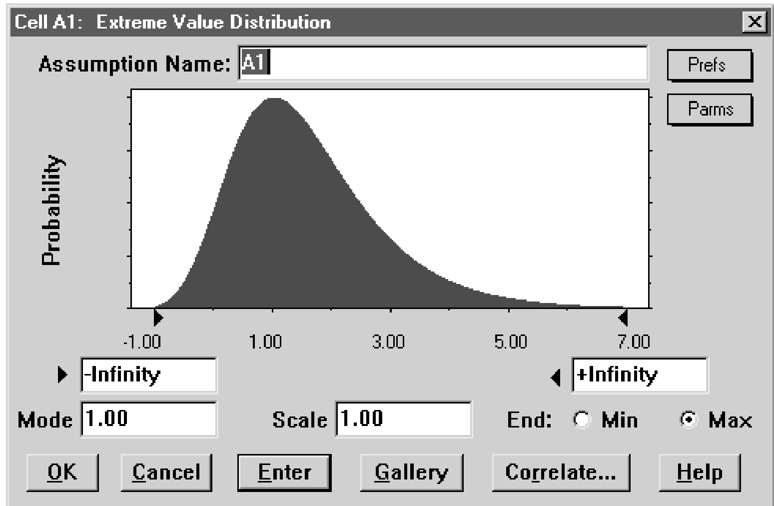


Figure 2.26 Extreme value distribution

Calculating parameters

There are two standard parameters for the extreme value distribution: mode and scale. The mode parameter is the most likely value for the variable (the highest point on the probability distribution).

After you select the mode parameter, you can estimate the scale parameter. The scale parameter is a number greater than 0. The larger the scale parameter, the greater the variance.

To calculate a more exact scale, you can estimate the mean and use the equation:

$$\alpha = \left|\frac{\text{mean} - \text{mode}}{0.57721}\right|$$

where α is the scale parameter.

Or estimate the variance and use the equation:

$$\alpha = \sqrt{\frac{6 \cdot \text{variance}}{\pi^2}}$$

where α is the scale parameter.

In addition to these parameters, you can select whether to skew your distribution. To skew it positively (with most of the values at the lower end of the range), select End: Max. To skew it negatively (with most of the values at the upper end of the range), select End: Min.

Crystal Ball Note: *You can use the "Parms" or Alternate Parameter menu found in the upper right corner of the dialog to specify a extreme value distribution using different combinations of percentiles. For more information, see "Alternate parameter sets" on page 136.*

Negative binomial distribution

The negative binomial distribution is useful for modeling the distribution of the number of trials until the *r*th successful occurrence, such as the number of sales calls you need to make to close a total of 10 orders. It is essentially a *super*-distribution of the geometric distribution.

Conditions

The three conditions underlying the negative binomial distribution are:

- The number of trials is not fixed.
- The trials continue until the *r* th success.
- The probability of success is the same from trial to trial.

Note that the number of trials needed is never less than *r*.

Example

A manufacturer of jet engine turbines has an order to produce 50 turbines. Since about 20% of the turbines do not make it past the high-velocity spin test, the manufacturer will actually have to produce more than 50 turbines.

Matching these conditions with the negative binomial distribution:

- The number of turbines to produce (trials) is not fixed.
- The manufacturer will continue to produce turbines until the 50th one has passed the spin test.
- The probability of success (80%) is the same for each test.

These conditions match those of the negative binomial distribution.

The negative binomial distribution has two parameters—probability and shape. The shape parameter is used to specify the *r*th successful occurrence. In this example you would enter 0.8 for the probability parameter (80% success rate of the spin test) and 50 for the shape parameter, as shown in Figure 2.27.

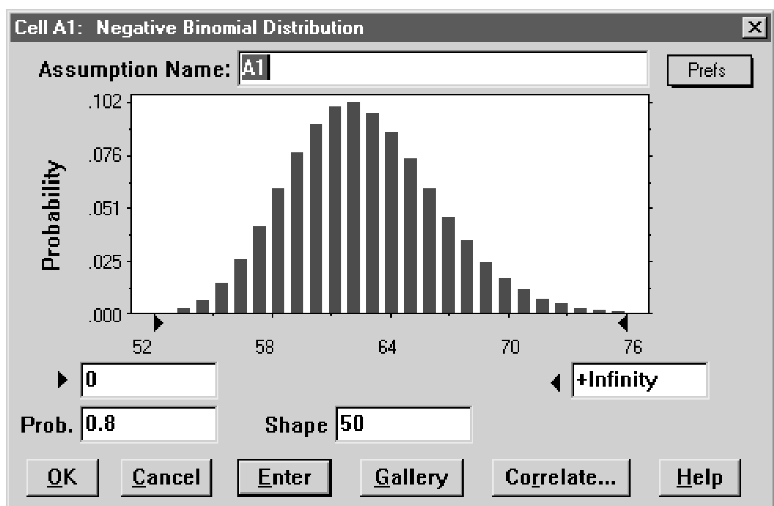


Figure 2.27 Negative binomial distribution

Some characteristics of the negative binomial distribution:

- When shape = 1, the negative binomial distribution becomes the geometric distribution.
- The sum of any two negative binomial distributed variables is a negative binomial variable.

- Another form of the negative binomial distribution, sometimes found in textbooks, considers only the total number of failures until the *r*th successful occurrence, not the total number of trials. To model this form of the distribution, subtract out *r* (the value of the shape parameter) from the assumption value using a formula in your spreadsheet.

Truncating distributions

You can change the bounds or limits of each distribution, except the custom distribution, by adjusting the truncation grabbers. This "truncates" the distribution. You can also exclude a middle area of a distribution by crossing over the truncation grabbers to "white out" the portion you want to exclude.

For example, suppose you want to describe the selling price of a house up for auction after foreclosure. The bank that holds the mortgage will not sell for less than $80,000. They expect the bids to be normally distributed around $100,000 with a standard deviation of $15,000. In Crystal Ball you would specify the mean as 100,000 and the standard deviation as 15,000 and then move the left grabber to set the limit of 80,000. The grabber "whites out" the portion you want to exclude, as shown in Figure 2.28.

Caveat

Each adjustment changes the characteristics of the probability distribution. For example, the truncated normal distribution in Figure 2.28 will no longer have an actual mean of $100,000 and standard deviation of $15,000. The values generated for the truncated distribution will have higher actual percentiles than for the specified values. Similarly, a normal distribution specified with 10/90%-tiles and truncated on either side of the distribution will have actual 10/90%-tiles greater or less than the specified percentiles.

Crystal Ball Note: *Showing the mean line of the distribution is useful when truncating distributions.*

Adjustments are made by moving the end-grabbers or by entering a different numeric endpoint than appears.

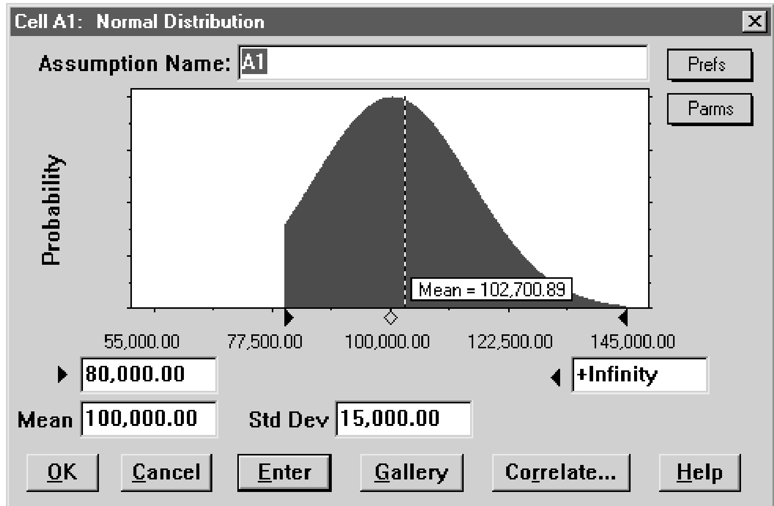


Figure 2.28 Truncated distribution example

Comparing the distributions

Many of the distributions discussed in this chapter are related to one another in various ways. For example, the geometric distribution is related to the binomial distribution. The geometric distribution represents the number of trials until the next success while the binomial represents the number of successes in a fixed number of trials. Similarly, the Poisson distribution is related to the exponential distribution. The exponential distribution represents the amount of time until the next occurrence of an event while the Poisson distribution represents the number of times an event occurs within a given period of time.

In some situations, as when the number of trials for the binomial distribution becomes very large, the normal and binomial distributions become very similar. For these two distributions, as the number of binomial trials approaches infinity, the probabilities become identical for any given interval. For this reason, you can use the normal distribution to approximate the binomial distribution when the number of trials becomes too large for Crystal Ball to handle (more than 1750 trials). You also can use the Poisson distribution to approximate the binomial distribution when the number of trials is large, but there is little advantage to this since Crystal Ball takes a comparable amount of time to compute both distributions.

The binomial and hypergeometric distributions are also closely related. As the number of trials and the population size increase, the hypergeometric trials tend to become independent like the binomial trials—the outcome of a single trial has a negligible effect on the probabilities of successive observations. The differences between these two types of distributions become important only when you are analyzing samples from relatively small populations. As with the Poisson and binomial distributions, Crystal Ball requires a similar amount of time to compute both the binomial and hypergeometric distributions.

The Weibull distribution is very flexible. Actually, it consists of a family of distributions that can assume the properties of several distributions. When the Weibull shape parameter is 1.0, the Weibull distribution is identical to the exponential distribution. The Weibull location parameter lets you set up an exponential distribution to start at a location other than 0.0. When the shape parameter is less than 1.0, the Weibull distribution becomes a steeply declining curve. A manufacturer might find this effect useful in describing part failures during a burn-in period. When the shape parameter is equal to 2.0, a special form of the Weibull distribution, called the Rayleigh distribution, results. A researcher might find the Rayleigh distribution useful for analyzing noise problems in communication systems or for use in reliability studies. When the shape parameter is set to 3.25, the Weibull distribution approximates the shape of the normal distribution; however, for applications when the normal distribution is appropriate, us it instead of the Weibull distribution.

The Erf distribution is also a special case of the normal distribution. It is a normal distribution with a mean of 0 and variance of $\frac{1}{2h^2}$.

The gamma distribution is also a very flexible family of distributions. When the shape parameter is 1.0, the gamma distribution is identical to the exponential distribution. When the shape parameter is an integer greater than one, a special form of the gamma distribution, called the Erlang distribution, results. The Erlang distribution is especially useful in the areas of inventory control and queueing theory, where events tend to follow Poisson processes. Finally, when the shape parameter is an integer plus one half (e.g. 1.5, 2.5, etc.), the result is a chi-square distribution, useful for modeling the effects between the observed and expected outcomes of a random sampling.

When no other distribution seems to fit your historical data or accurately describes an uncertain variable, you can use the custom distribution to simulate almost any distribution. The Data button on the Custom Distribution dialog lets you read a series of data points or ranges from value cells in your spreadsheet. If you like, you can use the mouse to individually alter the probabilities and shapes of the data points and ranges so that they more accurately reflect the uncertain variable.

Sampling methods

During each trial of a simulation, Crystal Ball selects a random value for each assumption in your model. Crystal Ball selects these values based on the sampling option in the Run > Run Preferences > Sampling dialog. The two sampling methods are:

Monte Carlo — Randomly selects any valid value from each assumption's defined distribution.

Latin hypercube — Randomly selects values, but spreads the random values evenly over each assumption's defined distribution.

Monte Carlo

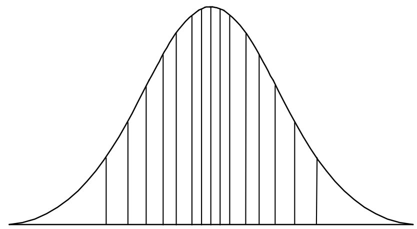
With Monte Carlo, Crystal Ball generates random values for each assumption's probability distribution, that are totally independent. In other words, the random value selected for one trial has no effect on the next random value generated.

Using Monte Carlo sampling to approximate the true shape of the distribution requires a larger number of trials than Latin hypercube.

Use Monte Carlo sampling when you want to simulate 'real world' what-if scenarios for your spreadsheet model.

Latin hypercube

With Latin hypercube sampling, Crystal Ball divides each assumption's probability distribution into non-overlapping segments, each having equal probability, as illustrated below.



Crystal Ball then selects a random assumption value for each segment according to the segment's probability distribution. The collection of values forms the Latin hypercube sample. After Crystal Ball uses all the values from the sample, it generates a new batch of values. This is why the simulation appears to stop at certain intervals — Crystal Ball is generating the new batch of values.

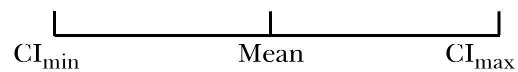
The Sample Size option (in the Run Preferences > Sampling dialog) controls the number of segments in the sample. See "Sampling preferences" on page 165 for more information on setting sampling method preferences.

Latin hypercube sampling is generally more precise when calculating simulation statistics than conventional Monte Carlo sampling, because the entire range of the distribution is sampled more evenly and consistently. Thus, with Latin hypercube sampling, you don't need as many trials to achieve the same statistical accuracy with Monte Carlo sampling. The added expense of this method is the extra memory required to hold the full sample for each assumption while the simulation runs.

Use Latin hypercube sampling when you are concerned primarily with the accuracy of the simulation statistics.

Confidence intervals

Since Monte Carlo simulation is a technique that uses random sampling to estimate model results, statistics computed on these results, such as mean, standard deviation and percentiles, will always contain some kind of error. A confidence interval (CI) is a bound calculated around a statistic that attempts to measure this error with a given level of probability. For example, a 95% confidence interval around the mean statistic is defined as a 95% chance that the mean's error will be contained within the specified interval. Conversely, there is a 5% chance that the mean's error will lie outside the interval. Shown graphically, a confidence interval around the mean looks like:



For most statistics, the confidence interval is symmetric around the statistic so that X = (CI_{max} - Mean) = (Mean - CI_{min}). This allows you to make statements of confidence such as "the mean's error will lie within the estimated mean plus or minus X with 95% probability." See Appendix B for a discussion on how CB calculates confidence intervals for different statistics.

Confidence intervals are important for determining the accuracy of statistics, and hence, the accuracy of the simulation. Generally speaking, as more trials are calculated, the confidence interval narrows and the statistics become more accurate. The precision control feature in Crystal Ball uses this characteristic of confidence intervals to determine when a specified accuracy of a statistic has been reached, and then stops the simulation accordingly. (See “Trials preferences” on page 164 and “Defining forecasts” on page 153 for more information on how to use precision control).

Understanding other statistical terms

Forecast Statistics window terms

These terms are listed in the Forecast Statistics window when you run a simulation in Crystal Ball and in the reports you can create for each forecast. The formulas for the statistics are contained in Appendix B, "Equations and Methods."

Statistic	*See:*
Mean	page 116
Median	page 116
Mode	page 117
Standard deviation	page 117
Variance	page 118
Skewness	page 119
Kurtosis	page 120
Coefficient of variability	page 120
Range (also range width)	page 121
Mean standard error	page 121

Mean

The mean of a set of values is found by adding the values and dividing their sum by the number of values. The term "average" usually refers to the mean. For example, 5.2 is the mean or average of 1, 3, 6, 7, and 9.

Median

The median is the middle value in a set of sorted values. For example, 6 is the median of 1, 3, 6, 7, and 9 (recall that the mean is 5.2).

If there is an odd number of values, the median is found by placing the values in order from smallest to largest and then selecting the middle value.

If there is an even number of values, the median is equal to the mean of the two middle values.

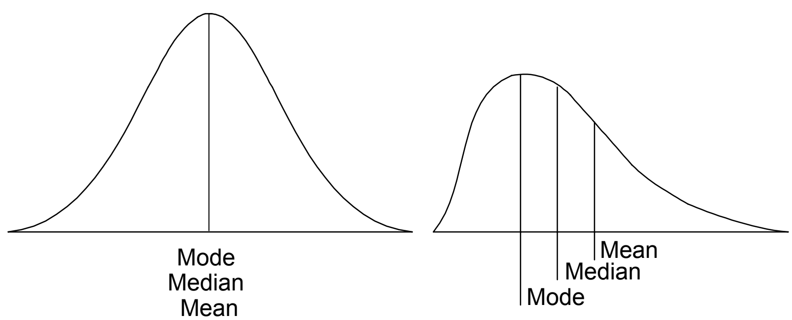
Mode

The mode is the value that occurs most frequently in a set of values. The greatest degree of clustering occurs at the mode.

The modal wage, for example, is the one received by the greatest number of workers. The modal color for a new product is the one preferred by the greatest number of consumers.

In a perfectly symmetrical distribution like the normal distribution (shown on the top left of the following page), the mean, median, and mode converge at one point.

In an asymmetrical or skewed distribution like the lognormal distribution, the mean, median, and mode tend to spread out, as shown in the following illustrations.



Crystal Ball Note: *When using continuous distributions, it is likely that your forecast will not have two values that are exactly the same. When this occurs, Crystal Ball sets the mode to '---' to indicate that the mode is undefined.*

Standard deviation

The standard deviation is the square root of the variance for a distribution. Like the variance, it is a measure of dispersion about the mean and is useful for describing the "average" deviation. See the description for the variance in the next section.

For example, you can calculate the standard deviation of the values 1, 3, 6, 7, and 9 by finding the square root of the variance that is calculated in the variance example on the following page.

The standard deviation, denoted as s, is calculated from the variance as follows:

$$s = \sqrt{10.2} = 3.19$$

Variance

Variance is a measure of the dispersion, or spread, of a set of values about the mean. When values are close to the mean, the variance is small. When values are widely scattered about the mean, the variance is larger.

To calculate the variance of a set of values:

1. **Find the mean or average.**
2. **For each value, calculate the difference between the value and the mean.**
3. **Square these differences.**
4. **Divide by n-1, where n is the number of differences.**

For example, suppose your values are 1, 3, 6, 7, and 9. The mean is 5.2. The variance, denoted by s^2, is calculated as follows:

$$s^2 = \frac{(1-5.2)^2 + (3-5.2)^2 + (6-5.2)^2 + (7-5.2)^2 + (9-5.2)^2}{5-1}$$

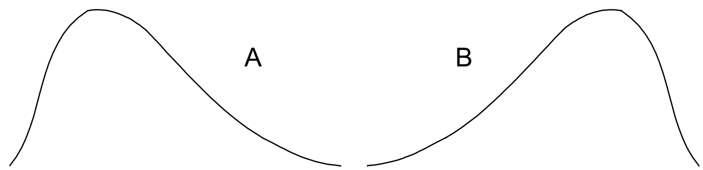
$$= \frac{40.8}{4} = 10.2$$

Crystal Ball Note: *The calculation uses n-1 instead of n to correct for the fact that the mean has been calculated from sample the data, thus losing one degree of freedom. This makes the sample variances slightly smaller than the variance of the entire population.*

Skewness

A distribution of values (a frequency distribution) is said to be "skewed" if it is not symmetrical.

For example, suppose the curves in the example below represent the distribution of wages within a large company.



Curve A illustrates positive skewness (skewed "to the right") where most of the wages are near the minimum rate, although some are much higher. Curve B illustrates negative skewness (skewed "to the left") where most of the wages are near the maximum, although some are much lower.

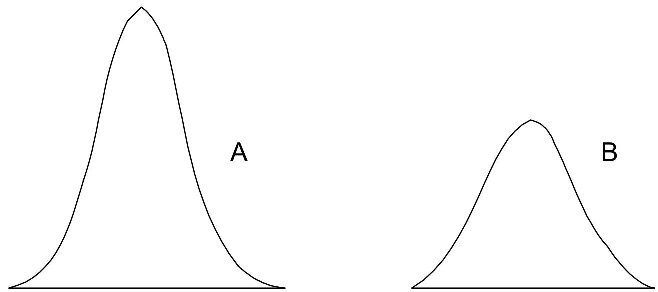
If you describe the curves statistically, curve A is positively skewed and might have a skewness coefficient of 0.5, while curve B is negatively skewed and might have a -0.5 skewness coefficient.

A skewness value greater than 1 or less than -1 indicates a highly skewed distribution. A value between 0.5 and 1 or -0.5 and -1 is moderately skewed. And a value between -0.5 and 0.5 indicates that the distribution is fairly symmetrical.

See Appendix B, "Equations and Methods," for an explanation of the formula used to calculate skewness.

Kurtosis

Kurtosis refers to the peakedness of a distribution. For example, a distribution of values might be perfectly symmetrical but look either very "peaked" or very "flat," as illustrated below:



Suppose the curves in the above examples represent the distribution of wages within a large company. Curve A is fairly peaked, since most of the employees receive about the same wage with few receiving very high or low wages. Curve B is flat topped, indicating that the wages cover a wider spread.

Describing the curves statistically, curve A is fairly peaked, with a kurtosis of about 4. Curve B, which is fairly flat, might have a kurtosis of 2.

A normal distribution is usually used as the standard of reference and has a kurtosis of 3. Distributions with a kurtosis value of less than 3 are described as platykurtic (meaning flat), and distributions with a kurtosis value of greater than 3 are leptokurtic (meaning peaked).

Crystal Ball Note: *Is some places, Crystal Ball uses a standard reference of 0.*

See Appendix B, "Equations and Methods," for an explanation of the formula used to calculate kurtosis.

Coefficient of variability

The coefficient of variability provides you with a measurement of how much your forecast values vary relative to the mean value. Since this statistic is independent of the forecast units, you can use it to compare the variability of two or more forecasts, even when the forecast scales differ.

For example, if you are comparing the forecast for a penny stock with the forecast for a stock on the New York Stock Exchange, you would expect the average variation (standard deviation) of the penny stock price to appear smaller than the variation of the NYSE stock. However, if you compare the coefficient of variability statistic for the two forecasts, you will notice that the penny stock shows significantly more variation on an absolute scale.

The coefficient of variability typically ranges from a value greater than 0 to 1. It might exceed 1 in a small number of cases in which the standard deviation of the forecast is unusually high.

The coefficient of variability is calculated by dividing the standard deviation by the mean, as follows:

$$\text{coefficient of variability} = \frac{s}{\bar{x}}$$

To present this in percentage form, simply multiply the result of the above calculation by 100.

Range (also range width)

The range minimum is the smallest number in a set, and the range maximum is the largest number.

The range is the difference between the largest and the smallest numbers in a set of values.

For example, if the range minimum is 10, and the range maximum is 70, then the range is 60.

Mean standard error

The mean standard error statistic lets you determine the accuracy of your simulation results and how many trials are necessary to ensure an acceptable level of error. This statistic tells you the probability of the estimated mean deviating from the true mean by more than a specified amount. The probability that the true mean of the forecast is within the estimated mean (plus or minus the mean standard error) is approximately 68%.

Statistical Note: *The mean standard error statistic only provides information on the accuracy of the mean and can be used as a general guide to the accuracy of the simulation. The standard error for other statistics, such as mode and median, will probably differ from the mean standard error.*

Additional statistical terms

This section explains other statistical terms used in Crystal Ball and to describe forecast results.

Correlation coefficient

Crystal Ball Note: *Crystal Ball uses rank correlation to determine the correlation coefficient of variables. For more information on rank correlation, see "Rank correlation" on page 123.*

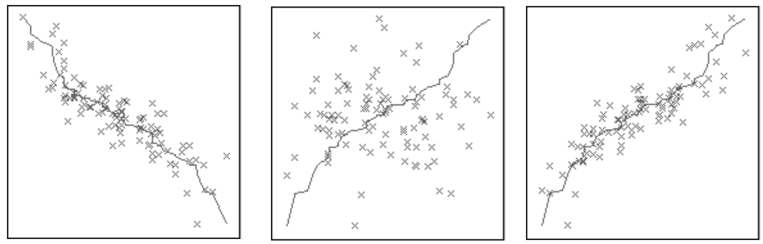
When the values of two variables depend upon one another in whole or in part, the variables are considered correlated. For example, an "energy cost" variable is likely to show a positive correlation with an "inflation" variable. When the "inflation" variable is high, the "energy cost" variable is also high; when the "inflation" variable is low, the "energy cost" variable is low.

In contrast, "product price" and "unit sale" variables might show a negative correlation. For example, when prices are low, sales are expected to be high; when prices are high, sales are expected to be low.

By correlating pairs of variables that are dependent on one another, you can increase the accuracy of your simulation forecast results.

The correlation coefficient is a number that describes the relationship between two dependent variables. Coefficient values range between -1 and 0 for a negative correlation and 0 and +1 for a positive correlation. The closer the absolute value of the correlation coefficient is to either +1 or -1, the more strongly the variables are related.

When an increase in one variable is associated with an increase in another variable, the correlation is called positive and indicated by a coefficient between 0 and 1. When an increase in one variable is associated with a decrease in another variable, the correlation is called negative and indicated by a coefficient between 0 and -1. A value of 0 indicates that the variables are unrelated to one another. The example below shows three correlation coefficients.



Negative Correlation **Zero Correlation** **Positive Correlation**

For example, assume that total hotel food sales might be correlated with hotel room rates. Total food sales are likely to be higher, for example, at hotels with higher room rates. If food sales and room rates correspond closely for various hotels, the correlation coefficient is close to 1. However, the correlation might not be perfect (correlation coefficient < 1). Some people might eat meals outside the hotel while others might skip some meals altogether.

When you select a correlation coefficient to describe the relationship between a pair of variables in your simulation, you need to consider how closely they depend upon one another.

You should never need to use an actual correlation coefficient of 1 or -1. Generally, you should represent these types of relationships as formulas on your spreadsheet.

Rank correlation

A correlation coefficient measures the strength of the linear relationship between two variables. However, if the two variables do not have the same probability distributions, they are unlikely to be related linearly. Under those circumstances, the correlation coefficient calculated on their raw values has little meaning.

If you calculate the correlation coefficient using rank values instead of actual values, the correlation coefficient is meaningful even for variables with different distributions.

You determine rank values by arranging the actual values in ascending order and replacing the values with their rankings. For example, the lowest actual value will have a rank of 1, the next lowest actual value will have a rank of 2, etc.

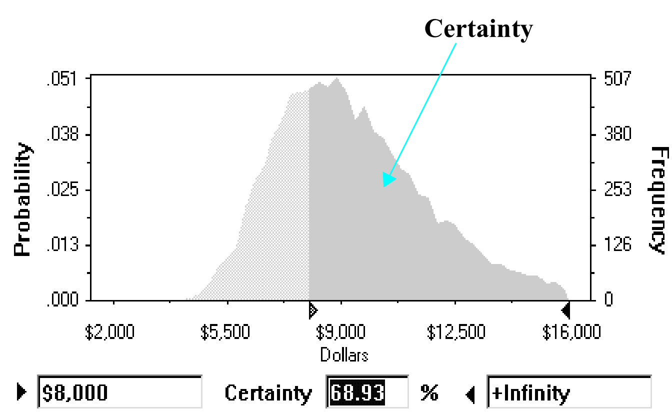
Crystal Ball uses rank correlation to correlate assumptions. The slight loss of information that occurs using rank correlation is offset by a couple of advantages.

- First, the correlated assumptions need not have similar distribution types. In effect, the correlation function in Crystal Ball is *distribution-independent*. The rank correlation method even works when a distribution has been truncated at one or both ends of its range.
- Second, the values generated for each assumption are not changed, they are merely *rearranged* to produce the desired correlation. In this way, the original distributions of the assumptions are preserved.

Certainty

Certainty describes the percentage of simulation results that fall within a range. For instance, in the Vision Research example from Chapter 1, if your objective was to make a minimum return of $2,000,000, you might choose a range of $2,000,000 to +Infinity.

.



By default, the range of certainty is from negative infinity to positive infinity. The certainty for this range is always 100%.

Percentile

A percentile is a number on a scale of zero to one hundred that indicates the percent of a distribution that is equal to or below a value. Standardized tests usually report results in percentiles. So if you were at the 95 percentile, that means that 95% of the people who took the test had the same score as you did or lower. This doesn't mean you answered 95% of the questions correctly. You might have answered only 20% correctly, but that was better than 95% of the people who took the test.

As another example, suppose you want to be 90% sure that you put enough money away for your retirement. You might create a model with all the uncertain variables, such as annual return on your investments, inflation, and expenses at retirement. The resulting distribution shows the most likely investment needed, but if you select the mean, you have a 50% chance of not having enough money. So you choose the value at the 90th percentile, which leaves only a 10% chance of not having enough money.

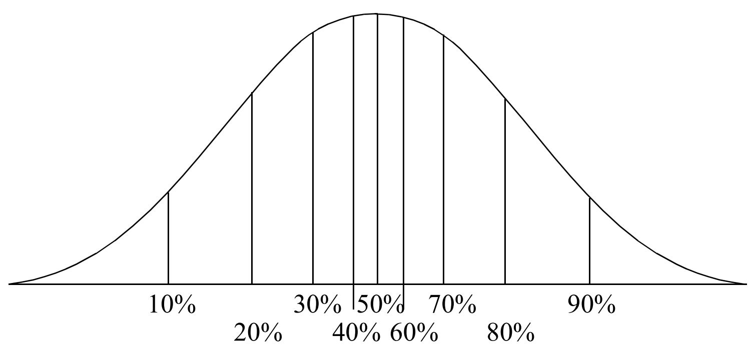


Figure 2.29 Percentiles for a normal distribution

Chapter 3

Setting Up and Running a Simulation

In this chapter

- Defining assumptions
- Using cell references
- Fitting distributions to data
- Specifying correlations between assumptions
- Editing assumptions
- Freezing assumptions
- Defining decision variables
- Defining forecasts
- Editing forecasts
- Selecting and reviewing your assumptions, forecasts, and decision variables
- Setting cell preferences
- Setting run preferences
- Running macros during the simulation
- Running/Stopping/Continuing/Resetting a simulation
- Saving and restoring a simulation
- Saving your models

This chapter provides step-by-step instructions for setting up and running a simulation in Crystal Ball. As you work through these instructions, you will learn to define assumption, forecast, and decision variable cells, and run a simulation. If you are a new user, you should read the sections in order. Later, when you become more familiar with Crystal Ball, you can refer to each section independently. After you complete this chapter, read Chapter 4 for information about interpreting simulation results.

If you need to use a command that is not described in this chapter, refer to Appendix D, "Index of Commands," in the back of this manual. The index of commands contains a brief description of each Crystal Ball command.

The Crystal Ball toolbar

To help you set up spreadsheet models and run simulation, Crystal Ball comes with a customized toolbar that provides instant access to the most commonly used menu commands.

The Crystal Ball toolbar looks like:

The tools in the first three groups are from the Cell menu. The tools from the next four groups are from the Run menu.

When you close Crystal Ball, Excel remembers the state of the toolbar. If the toolbar was showing when you closed Crystal Ball, it will be visible the next time you open Crystal Ball. If the toolbar was hidden when you closed Crystal Ball, it will be hidden the next time you start Crystal Ball.

To turn the Crystal Ball toolbar on or off, select View > Toolbars > Crystal Ball.

Starting Crystal Ball manually

When you installed Crystal Ball, you were prompted whether you wanted Crystal Ball to start automatically with Excel. If you answered No, you will need to start the program manually when you want to use Crystal Ball or set up Crystal Ball to start automatically in the future.

To set up Crystal Ball to automatically start with Excel:

1. **Start Excel.**
2. **Select Tools > Add-Ins.**

   The Add-Ins dialog appears.

3. **Check the Crystal Ball add-in option.**
4. **Click on OK.**

   Crystal Ball starts, and is set to automatically start with Excel from now on.

Other methods of starting manually

Select the icon

If you have not already started Excel, start Crystal Ball manually by:

1. **In Windows 95, Windows 98, Windows 2000, or Windows NT Version 4.0 or higher, select the Crystal Ball shortcut icon from the Windows Start >Programs > Crystal Ball menu.**

Using the File menu

If you have already started Excel and want to start Crystal Ball manually (without setting it up to start automatically in the future):

1. **Select File > Open.**

   The Open dialog appears.
2. **Find the Crystal Ball folder.**
3. **Select the CB.XLA file.**
4. **Click on OK.**

   The Crystal Ball toolbar and menus load into Excel.

Defining cells: an overview

Crystal Ball lets you define three types of cells: assumption cells decision variable cells, and forecast cells.

Assumption cells contain the values that you are unsure of—the uncertain independent variables in the problem you are trying to solve. The assumption cells must contain simple numeric values, not formulas or text.

Decision variable cells contain the values that are within your control to change. The decision variable cells must contain simple numeric values, not formulas or text.

Forecast cells (dependent variables) contain formulas that refer to one or more assumption and decision variable cells. The forecast cells combine the values in the assumption, decision variable, and other cells to calculate a result. A forecast cell, for example, might contain the formula C17*C20*C21.

Crystal Ball Note: *Crystal Ball automatically names the assumptions, decision variables, and forecasts in your spreadsheets. It also assigns default values to the distribution parameters (such as the probability, number of trials, and population size, among others) and variable bounds (the upper and lower bounds or decision variables). For a list of the defaults refer to Appendix E, "Default Names and Distribution Parameters."*

Defining assumptions

Crystal Ball uses probability distributions to describe the uncertainty in your assumption cells. From a gallery of seventeen distribution types, you choose the ones that best describe the uncertain variables in the problem you are trying to solve. When defining an assumption, there are two steps you must perform: 1) identify a distribution type, and 2) enter the assumption.

Identifying a distribution type

How do you know which distribution type to choose? If you are experienced in statistical methods, you probably will have a feel for which distribution to use. If your statistical background is limited or you need a review, consider using the following process for selecting the correct distribution:

1. **Look at the variable in question. List everything you know about the conditions surrounding this variable.**

   For example, look at the variable "patients cured" that was discussed in the Vision Research tutorial in Chapter 1. The company must test 100 patients. You know that the patients will either be cured or not cured. And you know that the drug has shown a cure rate of around 0.25 (25%). These facts are the conditions surrounding the variable.

2. **Review the descriptions of the probability distributions discussed in Chapter 2.**

   In Chapter 2, each distribution type is described in detail, with the conditions underlying the distribution outlined and with real-life examples. As you review the conditions and examples for each distribution, look for a distribution that characterizes the variable in question, as explained in the following step.

3. **Identify the distribution that characterizes the variable in question.**

   For a distribution to characterize a variable, the conditions of the distribution must closely approximate the conditions of the variable. Look at the conditions of the binomial distribution, as described in Chapter 2:

   - For each trial, only two outcomes are possible—success or failure.
   - The trials are independent. What happens on the first trial does not affect the second trial, and so on.
   - The probability of success remains the same from trial to trial.

   Now check the “patients cured” variable in question against the conditions of the binomial distribution:

   - There are two possible outcomes—the patient is either cured or not cured.
   - The trials (100) are independent of each other. What happens to the first patient does not affect the second patient.
   - The probability of curing a patient 0.25 (25%) remains the same each time a patient is tested.

   Since the conditions of the variable match the conditions of the binomial distribution, the binomial distribution would be the correct distribution type for the variable in question.

   Chapter 2, **“Understanding the Terminology”** has a more detailed discussion of this distribution selection process.

4. **If historical data is available, use distribution fitting to select the distribution that best describes your data.**

   Crystal Ball can automatically select the probability distribution that most closely approximates your data's distribution. The feature is described in detail in "Fitting distributions to data" beginning on page 138. You can also populate a custom distribution with your historical data.

After you select a distribution type, determine the parameter values for distribution. Each distribution type has its own set of parameters. For example, there are two parameters for the binomial distribution—trials and probability. The conditions of a variable contain the values for the parameters. In the example used, the conditions show 100 trials and 0.25 (25%) probability of success.

Entering an assumption

After you choose the correct distribution type, you are ready to enter the assumption.

To enter an assumption:

1. **Select a cell or a range of cells (Select value cells only. Assumptions cannot be defined for formula, non-numeric, or blank cells).**

Crystal Ball Note: *There is no absolute limit to the number of assumptions you can define per worksheet. As a rule of thumb, you should limit your combined Crystal Ball data elements (assumptions, forecasts, decision variables, and correlations) to 500 per worksheet (i.e., A+F+C+D< 500) for performance reasons.*

2. **Select Cell > Define Assumption.**

   For each selected cell or cells in the selected range, Crystal Ball displays the Distribution Gallery.

3. **Click the distribution you want from the gallery of choices.**

4. **Click on OK.**

   A dialog appears, showing the distribution type you chose for the selected cell (or for the first cell in a range of cells). Figure 3.1 provides an example of the binomial distribution.

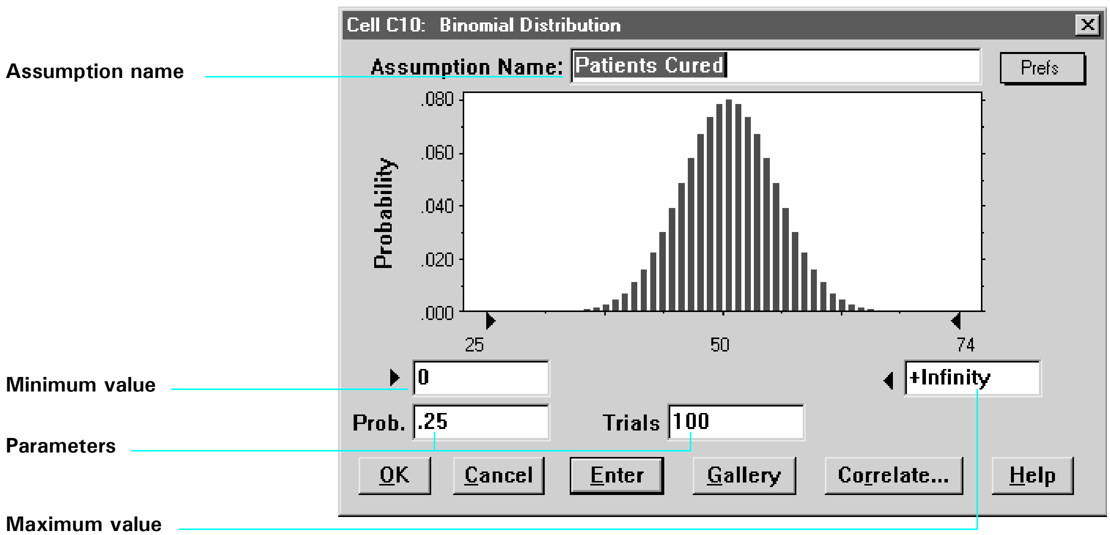


Figure 3.1 Binomial distribution

Crystal Ball Note: *If you want to change the distribution type, click on Gallery to return to the Distribution Gallery, and then select another distribution. If you want to cancel and return to the spreadsheet, click Cancel.*

5. **In the dialog, type a name for the assumption (optional).**

   If the assumption cell already has a name next to it or above it on the spreadsheet, the name appears in the dialog. You can use that name or type a new name.

6. **Type the parameters for the distribution.**

   Default values appear for the display range bounds and the distribution parameters. You can type new values or cell references (see "Entering cell references" on page 135) in any field. You can also change the display range bounds by moving the truncation grabbers.

Crystal Ball Note: *You can change the values entered by typing new values. However, if you are using the custom distribution, you must first clear the initial values with the following steps: click on Gallery, click the custom distribution, and then click on OK.*

7. Click on Enter.

The distribution changes to reflect the values you entered.

Crystal Ball Note: *If you click on OK instead of Enter, Crystal Ball accepts the parameters and closes the dialog.*

8. Click on OK.

If you selected a range of cells, repeat steps 3-8 to define the assumption for each cell.

Entering cell references

In addition to numeric values, you can enter a reference to a specific cell in a parameter field. Cell references must be preceded by an equals sign (=). Cell references can be either absolute or relative.

Relative referencing

Relative references remember the position of a cell relative to the cell containing the assumption. For example, an assumption in cell C6 refers to cell C5. If the assumption in C6 is copied to cell C9, the relative reference to C5 will then refer to the value in cell C8. This lets you easily set up a whole row or column of assumptions, each having similar distributions but slightly different parameters, by performing just a few steps. An absolute reference, on the other hand, always refers back to the originally referenced cell, in this case C5.

Absolute referencing

To indicate an absolute reference, you must use a dollar sign ($) before the row and the column. For example, to copy the exact contents of cell C5 into an assumption parameter field, you would enter the cell reference =C5. This causes the value in cell C5 to be used in the assumption cell parameter field. Later, if you decide to copy and paste this assumption in the spreadsheet, the cell references in the parameter field will refer to the contents of cell C5.

The Define Assumption dialog has two radio buttons labeled Static and Dynamic that become visible when you enter a cell reference (e.g., =A1) in any of the parameter fields.

When you select Static (the default selection), Crystal Ball resolves all cell references for the distribution at the moment the simulation starts, thus freezing the shape of the distribution for the entire simulation. This lets the simulation run faster since Crystal Ball does not have to check the parameters for errors.

When you select Dynamic, Crystal Ball resolves cell references each time the spreadsheet is recalculated. This feature is called Dynamic Cell Referencing. It gives you more flexibility in setting up models by letting you change an assumption's distribution during a simulation.

Alternate parameter sets

For almost all the continuous probability distributions, except uniform, you can define the distributions using percentiles for parameters. This option gives you added flexibility in setting up assumptions when only percentile information is available, or when specific attributes (e.g. mean, standard deviation) of the variable in your model are unknown.

For example, if you are defining a triangular distribution but are unsure of the absolute minimum and maximum values of the variable, you could instead define the distribution using the 10^{th} and 90^{th} percentiles along with the likeliest value. This gives you a distribution that has 80%, or four-fifths of the values, occurring between the two specified percentiles, as in Figure 3.2.

To change the parameter sets for the continuous distributions, use the popup "Parms" menu in the upper right corner of the Define Assumption dialog. The currently selected parameter set has a check mark next to it, as shown on the menu below.

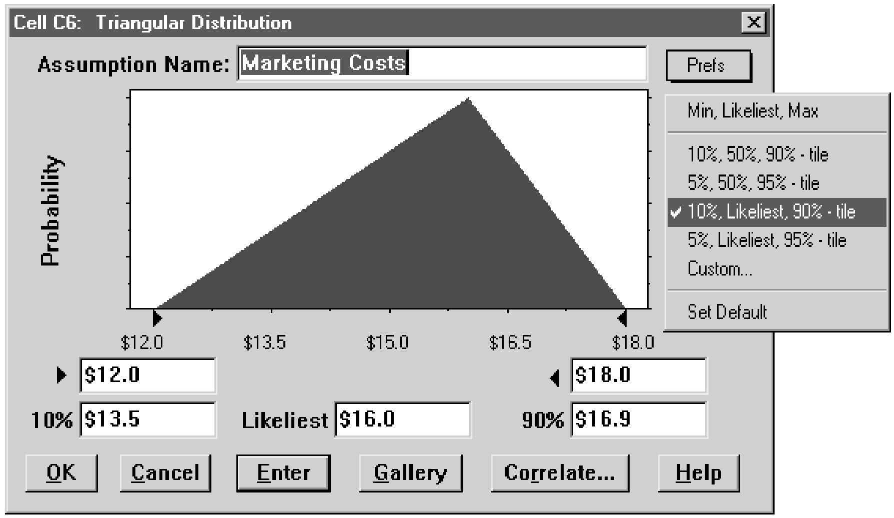


Figure 3.2 10/90 percentiles with likeliest parameter

In addition to the standard parameter set, each continuous distribution's Parms menu has four additional pre-defined parameter sets that include various combinations of the standard parameters and percentiles. There is also a custom option that lets you define your own parameter set.

For the Custom Parameters, you can select to replace any or all of the standard parameters with any percentile. You will always have the same number of parameters, either standard or alternate, for any given distribution. For example, even if you choose to use custom alternate parameters for a triangular distribution, you will always have three parameters, either minimum, likeliest, or maximum, or, for example, 10%-tile, likeliest, and 99%-tile.

To select a parameter set to use as the default when defining new assumptions of this type, select Set Default from the popup menu.

Several special parameter sets are available with the lognormal distribution, including geometric and logarithmic sets. For more information, see "Lognormal parameter sets" on page 83.

Distribution preferences

All the assumption distribution dialogs have a Prefs button in the top right corner. This button brings up a menu with the following options:

Glossary Term:
CDF—
A cumulative distribution function that represents the probability that a variable will fall at or below a given value.

Glossary Term:
PDF—
A probability density function that represents the probability that an infinitely small variable interval will fall at a given value.

View Cumulative Chart
: Changes the distribution to display either a CDF (when checked) or a PDF.

Show Coordinates
: Displays the mouse coordinates when the cursor is over the displayed distribution, when checked. For discrete distributions or continuous CDFs, both the *x* coordinate and the associated probability appear. For continuous PDFs, only the *x* coordinate appears.

Statistical Note: *Crystal Ball doesn't display the probability on the continuous PDFs because, for continuous random variables, the occurrence of any exact value of x has a probability of 0.*

Show Mean Line
: Displays a vertical line at the distribution mean, when checked.

Fitting distributions to data

If you have historical data available, Crystal Ball's distribution fitting feature can substantially simplify the process of selecting a probability distribution. Not only is the process simplified, but the resulting distribution more accurately reflects the nature of your data than if the shape and parameters of the distribution were estimated.

How distribution fitting works

Glossary Term:
Goodness-of-Fit—
A set of mathematical tests performed to find the best fit between a standard probability distribution and a data set's distribution.

In distribution fitting, Crystal Ball automatically matches your data against each continuous probability distribution. A mathematical fit is performed to determine the set of parameters for each distribution that best describe the characteristics of your data. The quality or *goodness* of each fit is judged using one of several standard **goodness-of-fit** tests. The distribution with the highest ranking fit is chosen to represent your data.

At your discretion, you can review the distributions sorted in order of their fit tests using a comparison chart. This chart shows the fitted distributions superimposed over your data, allowing you to visually check the quality of the fits. Several chart preferences, such as a Difference option, make it easier to pinpoint discrepancies in the fits. A summary field shows the goodness-of-fit statistics. If desired, you can override the highest-ranking probability distribution with another one of your choice.

Crystal Ball Note: *Distribution fitting can also be used to check the characteristics of a forecast chart. See "Using distribution fitting with the overlay chart" on page 212 for more information.*

Crystal Ball Note: *Only continuous distributions are considered for distribution fitting. Continuous and discrete distributions are defined on page 63.*

In Crystal Ball, you use two Fit Distribution dialogs to specify the source of your data, the distributions to be fitted, and a goodness-of-fit test. Each goodness-of-fit test is calculated for every distribution, but only the selected test determines how the distributions are ranked.

Distribution fitting example

To use distribution fitting:

1. **Create a new spreadsheet and enter a value in any cell.**

2. **Select Cell > Define Assumption.**

   The Distribution Gallery appears.

3. **Click on Fit to select the source of the fitted data.**

   The first of two Fit Distribution dialogs appears, as in Figure 3.3.

4. Select one of the following two options.

- If the historical data is in the front-most spreadsheet, select the Active Worksheet option, and then enter the data's cell range.
- If the historical data is in a separate text file, click on File, and then either enter the path and name of the file or click on Browse to search for the file. For this example, click on Browse and locate the TEST DATA.txt file in the Examples folder.

Crystal Ball Note: *When you use a file as your source of data, each data value in the file must be separated with either a comma, a tab character, or an end-of-line character.*

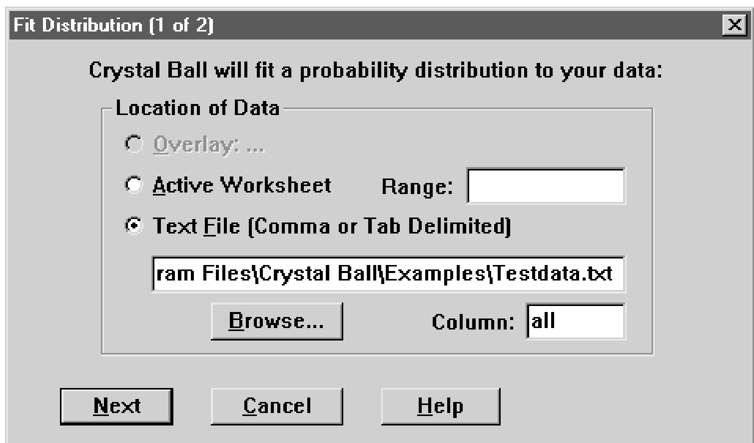


Figure 3.3 Fit Distribution (1 of 2) dialog

Crystal Ball Note: *The overlay option is only visible on this dialog when fitting a forecast histogram with a distribution on the overlay chart.*

5. Click on Next to select the distribution fitting characteristics.

The second Fit Distribution dialog is used to specify which distributions are to be fitted and how they should be ranked, as in Figure 3.4. The number of data values also appears at the top of the dialog. The amount of data has a direct effect on the duration of the fitting process. For our example, use these settings:

- Under "Fit to Which Distributions," select All Continuous Distributions.

  All Continuous Distributions fits the data to all those distributions in which every value in the distribution's range is possible (these distributions appear as solid shapes on the Distribution Gallery). Chosen Distributions displays another dialog, from which you can select a subset of the continuous distributions to include in the fitting. The third option selects the distribution that was highlighted on the Distribution Gallery.

- Under "Ranking Method," select the Chi-Square Test.

  In ranking the distributions, you can use any one of three standard goodness-of-fit tests. Each is described below:

  Chi-square. This test is the oldest and most common of the goodness-of-fit tests. It gauges the general accuracy of the fit. The test breaks down the distribution into areas of equal probability and compares the data points within each area to the number of expected data points. Generally a *p-value* greater than 0.5 indicates a close fit.

  Kolmogorov-Smirnov. The result of this test is essentially the largest vertical distance between the two cumulative distributions. Generally a value less than 0.03 indicates a close fit.

  Anderson-Darling. This method closely resembles the Kolmogorov-Smirnov method, except that it weights the differences between the two distributions at their tails greater than at their mid-ranges. Use this method when you need a better fit at the extreme tails of the distribution. Generally a value less than 1.5 indicates a close fit.

- Check the Show Comparison Chart And Goodness-of-Fit Statistics option. The Comparison Chart and Fit Statistics appear after the fits are made.

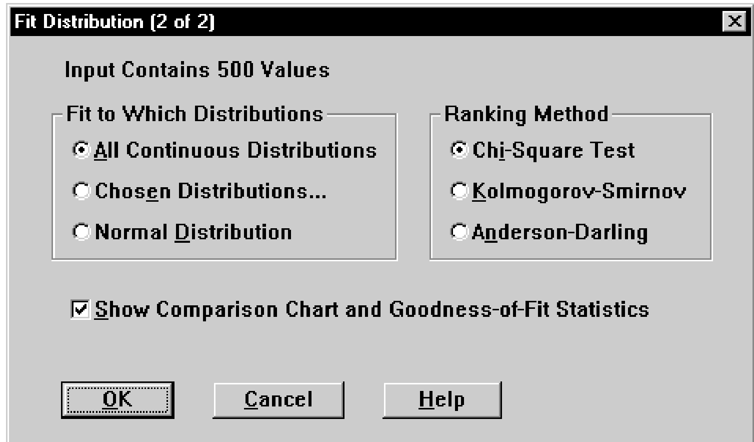


Figure 3.4 Fit Distribution (2 of 2) dialog

6. Click on OK to fit the distributions to your data.

Crystal Ball successively fits the selected distributions to your data. If you left the Show Comparison Chart and Goodness-of-Fit Statistics option *unchecked*, the Define Assumption dialog immediately appears with the highest-ranked distribution. If you *checked* the Show Comparison Chart and Goodness-of-Fit Statistics option, the fitted distributions appear, starting with the highest-ranked distribution down through to the lowest.

You can scroll through the fitted probability distributions by using the Next Distribution button. Each probability distribution is shown superimposed over the data, as shown in Figure 3.5.

Comparison chart

You can use the Comparison Chart dialog to visually compare the quality of the fits or to view the goodness-of-fit statistics. Use the Comparison Chart features as described below:

- Click on the Next Distribution/Prev. Distribution buttons to scroll through the fitted distributions. You can view the quality of each fit graphically and statistically in decreasing order.

- Click on Prefs to change chart features so that similarities or differences are more clearly accentuated. You can find a relevant explanation of these features in “Customizing the overlay chart” beginning on page 209.
- Click on Accept to select the currently displayed distribution for use in your assumption.
- Click on Cancel to return to the second Fit Distribution dialog.

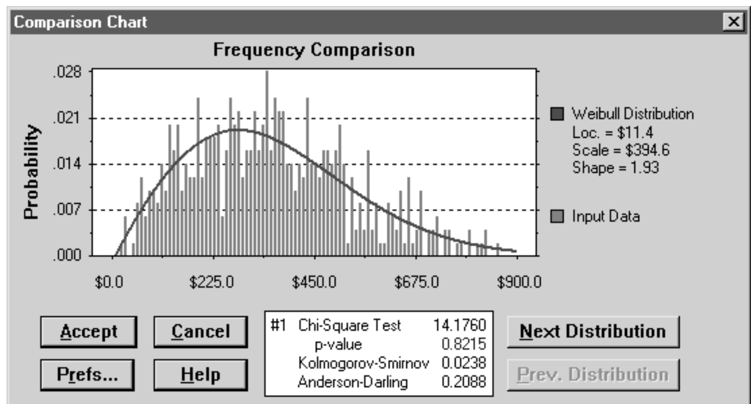


Figure 3.5 Comparison Chart dialog

In our Test Data example, the Weibull distribution had the best fit of any distribution using the chi-square fit test. As you can see in Figure 3.5, the chi-square p-value is 0.8215, which indicates a good fit. The parameters that were calculated for the Weibull distribution are displayed in the Comparison Chart legend.

Crystal Ball Note: *The goodness-of-fit statistic that Crystal Ball used to rank the distributions appears in blue.*

Comparison Chart - Difference

You can use the Difference option to more clearly highlight the difference between the fitted distribution and the data's distribution. Figure 3.6 shows the difference between the fitted Weibull distribution and the test data.

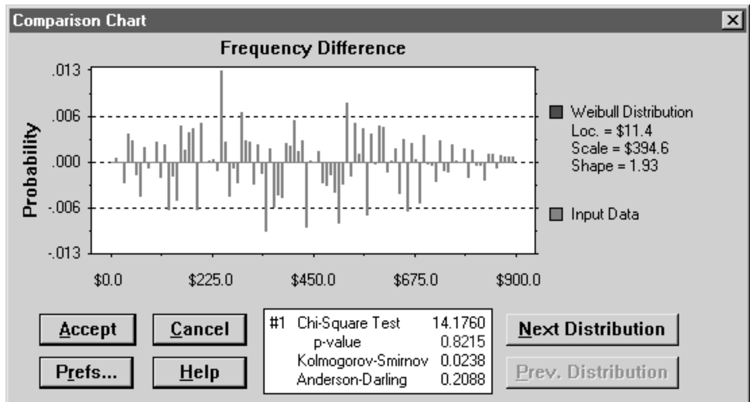
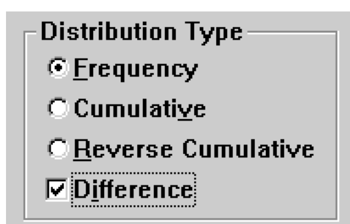


Figure 3.6 Comparison Chart - Difference display option

Select the Difference option by pressing the *i* key while you are in the chart dialog or checking the option in the Distribution Type area of the Chart Preferences dialog, as shown below. You can use this option with all three distribution types.



7. **Click on Accept to choose the currently displayed distribution.**

   The Assumption dialog appears with the parameter entries taken from the chosen distribution. You can change the distribution parameters before you click on OK.

Specifying correlations between assumptions

In Crystal Ball, assumptions are usually calculated independently of each other. Crystal Ball generates random numbers for each assumption without regard to how random numbers are generated for other assumptions.

However, dependencies often do exist between variables in a system being modeled. The Correlated Assumptions feature in Crystal Ball lets you use correlation coefficients to define the dependencies between assumptions. Crystal Ball then uses the correlation coefficients to rearrange the generated random numbers to produce the desired correlations. Correlation coefficients and correlation in general are explained in Chapter 2.

Crystal Ball Note: *Crystal Ball uses rank correlation to calculate correlation coefficients. For more information on how Crystal Ball calculates rank correlation coefficients, see "Rank correlation" on page 123.*

Crystal Ball Note: *When using dynamic cell referencing, you cannot correlate the assumptions since the shape of the distribution is unpredictable.*

When defining correlations, you are usually limited by an approximate, practical limit of 50 to 70 correlations. Also, the more correlations you define, the greater the possibility that some correlations might be in conflict with each other, preventing Crystal Ball from running a simulation. Conflicts can arise when a group of assumptions are improperly related to each other by large positive and/or large negative correlation coefficients. When this condition occurs, the correlations are said to be "inconsistent." For more information, see "Correlated assumptions" on page 251.

With Crystal Ball, you use the Correlation dialog to specify a correlation coefficient for any pair of assumptions.

To define a correlation coefficient:

1. **Select the cell of one of the assumptions you want to correlate; for example, a cell describing the inflation rate.**

Statistical Note: *Which of the pair you select first is not important, since the correlation coefficient is bidirectional.*

2. Select Cell > Define Assumption.

The distribution previously defined for this assumption appears, as in Figure 3.7.

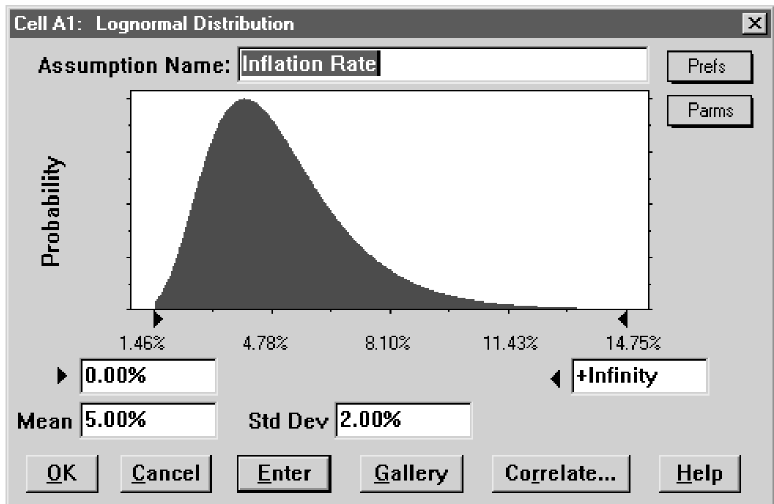


Figure 3.7 Lognormal distribution

3. Click on Correlate.

The Correlate dialog appears, as in Figure 3.8.

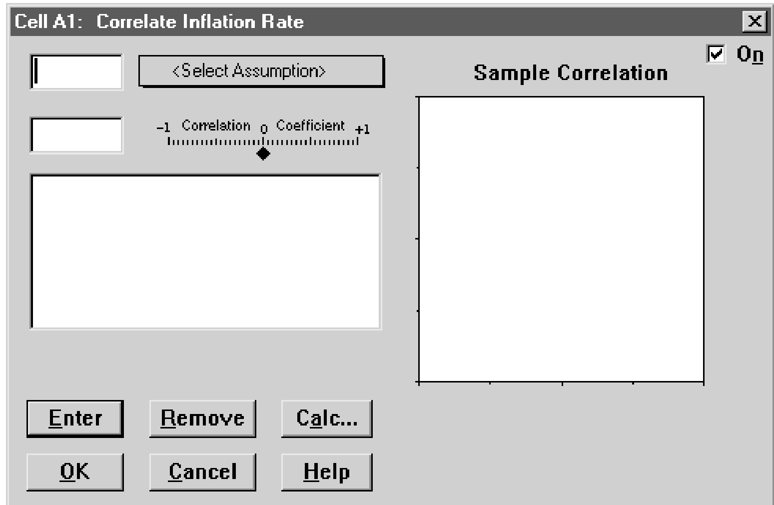


Figure 3.8 Correlate dialog

4. **Select the second assumption from the Select Assumption menu on the Correlate dialog.**

   The Select Assumption menu provides a list of the names of all the assumptions defined on your spreadsheet. Select one of the assumption names on the list. The cell reference of the assumption appears in the list to the left of the menu.

   Alternatively, enter the cell reference of the assumption directly into the first parameter field and press <Enter>.

   After you have selected the second assumption, the cursor moves to the Correlation Coefficient field.

5. **Enter a correlation coefficient. Either:**

   - Enter a value between -1 and 1(inclusive) in the Coefficient Value field.

     Type the number that you want to use in the field to the left of the slider control. After you type the number, the slider control on the Correlation Coefficient scale moves to the selected value.

- Drag the slider control along the Correlation Coefficient scale.

  The value you select appears in the field to the left of the scale.

- Click on Calc...

  A small dialog appears at the bottom of the first dialog. Enter the range of cells on your spreadsheet that contain the empirical values that Crystal Ball should use to calculate a correlation coefficient.

  Enter the range of cells in the standard A1:A2 format, where A designates the column and 1 and 2 designate the first and last cell rows, respectively. For example, if one set of values is in column Q, rows 10 through 15 and the second set of values is in column R, rows 10 through 15, enter the range in the left field as Q10:Q15 and the range in the right field as R10:R15.

  Crystal Ball calculates the correlation coefficient, enters it in the field to the left of the Correlation Coefficient scale, and moves the slider control to the correct position.

Crystal Ball Note: *The two cell ranges do not necessarily have to have the same dimensions, but they must contain the same number of value cells. The cell ranges are read in a row-by-row fashion.*

Each time you select a new assumption cell or correlation coefficient, Crystal Ball displays a sample correlation chart.

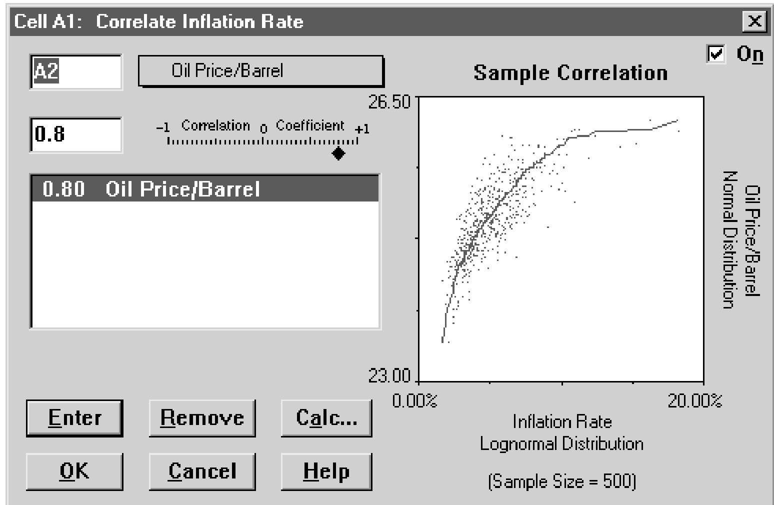


Figure 3.9 Correlate chart

The points on the chart represent the pairing of assumption values as they would actually occur when running a simulation. The solid line running through the middle of the chart indicates the location where values of a perfect correlation (+1.0 or -1.0) would fall. The closer the points are to the solid line, the stronger the correlation.

In the example above, an Inflation assumption and an Oil Price/ Barrel assumption have been correlated using a coefficient of 0.8, a strong positive correlation. As the points on the chart show, higher inflation values tend to be associated with higher oil prices and vice versa. This chart can help you begin to understand how the two assumptions are related.

You can specify as many of these paired correlations as you want for each assumption, up to the total number of assumptions defined.

You can ignore correlation if one or both variables do not impact the output or are not highly correlated.

Crystal Ball Note: *The correlation sample size is determined by the Sample Size option in the Run Preferences dialog. See "Setting run* *preferences" on page 163 or "Simulation accuracy" on page 243 for more information*

Crystal Ball Note: *For Crystal Ball to display the sample correlation chart, random values are actually produced for both of the assumptions and then correlated to the specified coefficient. This process takes time, depending on the size of the random samples and the types of distributions. You can turn the process off by unchecking the On option in the upper right corner of the dialog.*

Freezing assumptions

The Freeze Assumptions command lets you exclude or "freeze" certain assumptions from a simulation. The Freeze Assumptions command gives you the ability to:

- Temporarily disable assumptions that have been defined but are not wanted for the current simulation.
- See the effect of certain assumptions on the forecast cells while holding other assumptions to their spreadsheet value.

Note the following features of the Freeze Assumptions dialog (see Figure 3.10).

- Clicking on Frozen excludes those assumptions listed in the right list from the current simulation.
- Clicking on Enabled includes only those assumptions listed in the right list in the current simulation.
- Clicking on Choose moves the assumptions selected in the left list to the right list.

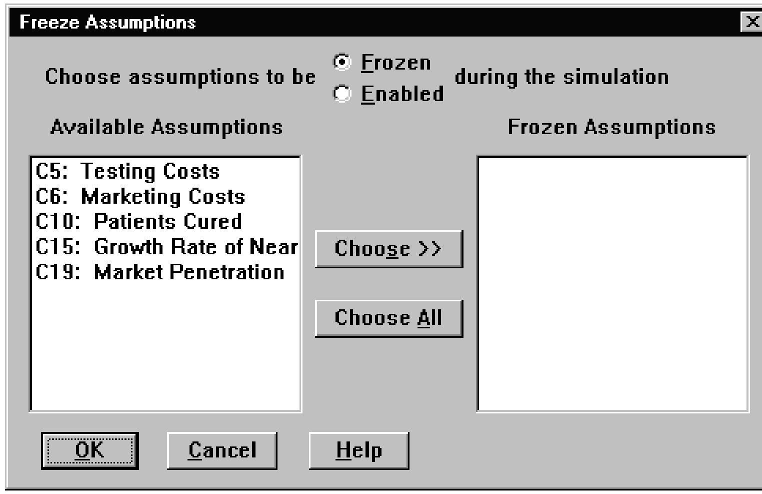


Figure 3.10 Freeze Assumptions dialog

Defining decision variables

Decision variables are not required for simulation models, but they can be very useful when comparing and selecting alternate situations. Several of the Crystal Ball tools discussed in Chapter 6, **“Crystal Ball Tools”**, use and benefit from decision variables. Decision variables are the variables that you can control, such as rent to charge or the amount of money to invest in a fund.

OptQuest Note: *You also use decision variables with OptQuest, available with Crystal Ball 2000, Professional Edition.*

To define a decision variable:

1. **Select a cell or range of cells.**

   Select value cells only. You cannot define a decision on a formula, non-numeric, or blank cell.

2. **Select Cell > Define Decision.**

   The Define Decision Variable dialog appears. For details about the fields and options in this dialog, see “Define Decision Variable dialog” below.

3. **Complete the Define Decision Variable dialog.**
4. **Click on OK.**
5. **Repeat steps 1–4 for each decision variable in your model.**

Crystal Ball Note: *There is no absolute limit to the number of decision variables you can define per worksheet. As a rule of thumb, you should limit your combined Crystal Ball data elements (assumptions, forecasts, decision variables, and correlations) to 500 per worksheet (i.e., A+F+C +D< 500) for performance reasons.*

Define Decision Variable dialog

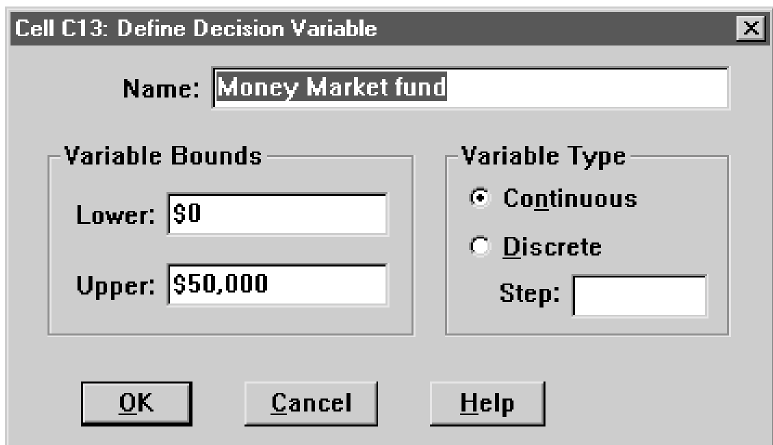


Figure 3.11 Define Decision Variable dialog

The Define Decision Variable dialog has the following fields and options:

Name	Is the name of the decision variable. You can replace the default name with a name up to 40 characters. You cannot use special characters (+, -, *, /, ^, <, =, >). By default, Crystal Ball uses the range name for a cell or, if the cell isn't part of a range, it uses text from an adjacent text cell to the left of or above the variable cell. If Crystal Ball can't find a range name or text cell to use, it uses the cell coordinates.

Lower	Is the lower limit for the decision variable range. The default is 90% of the cell value.
Upper	Is the upper limit for the decision variable range. The default is 110% of the cell value.
Variable Type	Defines whether the variable is continuous or discrete. You must define a step for a discrete variable. Selecting Continuous automatically clears the Step field. The default is Continuous.
Step	Sets the difference between successive discrete values within the defined range. Discrete variables require a step. If you enter a value in this field, the Variable Type automatically changes to Discrete. The default step is 1.

Defining forecasts

After you define the assumption cells and decision variables, you are ready to select and define forecast cells. Forecast cells usually contain formulas that refer to one or more assumption and decision variable cells. The forecast cells combine cells in your model to produce the results you need.

When you define a forecast cell, you:

- Name the forecast
- Specify the units for the forecast
- Select the size of the forecast window
- Indicate whether you want to automatically display the forecast window during the simulation
- Set precision control settings for the forecast

Crystal Ball Note: *Forecast cells can be defined directly on top of assumption cells so that the random values generated during a forecast can be captured and reported.*

To define forecast cells:

1. **Select a cell or a range of contiguous value or formula cells.**

Crystal Ball Note: *There is no absolute limit to the number of forecasts you can define per worksheet. As a rule of thumb, you should limit your combined Crystal Ball data elements (assumptions, forecasts, decision variables, and correlations) to 500 per worksheet (i.e., A+F+C +D< 500) for performance reasons.*

2. **Select Cell > Define Forecast.**

   The Define Forecast dialog appears, as in Figure 3.12.

3. **Complete the Define Forecast dialog.**
4. **Click on OK.**
5. **Repeat steps 1–4 for each forecast in your model.**

Define Forecast dialog

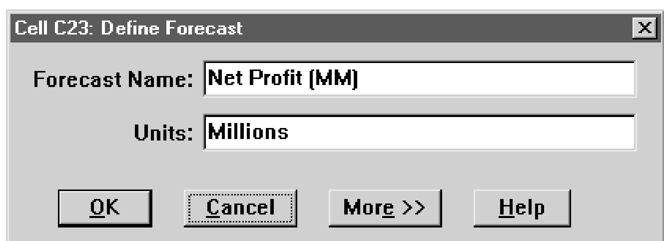


Figure 3.12 Define Forecast dialog

The Define Forecast dialog has the following fields and options:

Name	Is the name of the forecast. You can replace the default name with a name up to 40 characters.
	By default, Crystal Ball uses the range name for a cell or, if the cell isn't part of a range, it uses text from an adjacent text cell to the left of or above the forecast cell. If Crystal Ball can't find a range name or text cell to use, it uses the cell coordinates.
Units	Is the name of the units that will appear in at the bottom of the forecast chart.

You can expand the Define Forecast dialog by clicking on the More >> button.

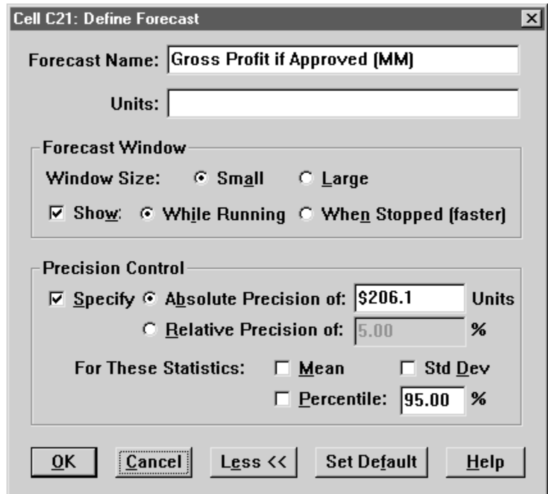


Figure 3.13 Expanded Define Forecast dialog

The additional fields and options are:

Window Size — Sets whether forecast windows appear Small or Large. Size is defined for each chart separately. After selecting the Large chart size, more groups become available for this forecast on the Chart Preferences dialog.

Show — Sets whether to display the forecast window automatically while the simulation is running or when the simulation is stopped.

You can display one or more forecasts while the simulation is running. If you decide not to display the forecast, the simulation continues to run. You can override this option and close all the forecast windows using the Suppress Forecast Windows option in the Run Preferences dialog. For more information, see “Speed preferences” on page 167.

Crystal Ball Note: *Running a simulation with the forecast windows closed decreases the time required to run the simulation.*

While Running Displays the forecast window automatically during simulations. This slows down the simulation.

When Stopped Displays the forecast window automatically after the simulation stops. This is faster than displaying the window during simulations.

Specify Turns on the precision control settings for the forecast. Crystal Ball only uses these settings if the simulation is set to stop when it reaches the specified precision from the Run Preferences dialog. For more information, see "Trials preferences" on page 164.

Absolute Precision
Determines the size of the confidence interval, in actual forecast units, used to test the precision of the forecast statistics.

Relative Precision
Determines the size of the confidence interval, in percentage terms, used to test the precision of the forecast statistics.

Crystal Ball Note: *See "Confidence intervals" on page 115 for more information about how the absolute and relative precision relate to the confidence interval.*

For These Statistics
Selects which statistics to check for the precision indicated above. The statistics available for precision control are: mean, standard deviation, and an indicated percentile.

Percentiles If you select Percentile, you can enter any number between percentile value between 1.00% and 99.00% to use as a precision control statistic.

Crystal Ball Note: *For more information on setting the Precision Control option and settings, see "Precision Control" on page 243.*

Set Default	Applies all the forecast options, such as window size, window display, and precision control settings, to all future forecast definitions.

Editing Crystal Ball data

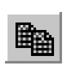
Crystal Ball provides edit commands to let you copy, paste, or clear assumptions, decision variables, or forecasts from cells. This feature lets you set up a whole row or column of assumptions, decision variables, or forecasts with just a few steps.

Excel Note: *Using the Excel cut, copy, and paste commands only copies the cell value and attributes, including cell color or pattern. The Excel commands do not copy the Crystal Ball data, even though they appear to since the cell colors copy. To copy Excel data and Crystal Ball data, you must use the Excel edit commands and then use the Crystal Ball edit commands to copy the Crystal Ball data.*

To copy Crystal Ball data:

1. **Select a cell or a range of cells that contains the Crystal Ball data you want to copy.**

Crystal Ball Note: *Make sure the range only contains one type of Crystal Ball data, such as assumptions, decision variables, or forecasts. If your range contains mixed Crystal Ball data, Crystal Ball only copies the first data type in the range.*



2. **Select Cell > Copy Data.**

   Crystal Ball copies the selected Crystal Ball data to the clipboard.

To paste Crystal Ball data:

1. **Select a cell or a range of cells that you want to paste into.**

2. **Select Cell > Paste Data.**

   Crystal Ball pastes each copied item from the clipboard area (mapping one for one, row by row) into the cells you selected. Any existing Crystal Ball data is replaced with the new data.

Crystal Ball Note: *If there are more copied Crystal Ball items in the clipboard area than cells in the selection, the remaining Crystal Ball items are ignored. If there are more cells in the selection than Crystal Ball items in the clipboard area, the Crystal Ball items in the clipboard area are reused, starting with the first one in the list. If copied Crystal Ball data remain in the clipboard, the clipboard empties when you run a simulation.*

To clear Crystal Ball data:

1. **Select a cell or a range of cells that contains the Crystal Ball data you want to clear.**

Crystal Ball Note: *Make sure the range only contains one type of Crystal Ball data, such as assumptions, decision variables, or forecasts. If your range contains mixed Crystal Ball data, Crystal Ball only clears the first data type in the range.*

2. **Select Cell > Clear Data.**

   Crystal Ball clears the Crystal Ball data from each cell.

To clear all of one type of Crystal Ball data from all cells in the active spreadsheets:

1. **Select either:**
   - Cell > Select All Assumptions
   - Cell > Select All Decisions
   - Cell > Select All Forecasts

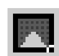
2. **Select Cell > Clear Data.**

   Crystal Ball clears the Crystal Ball data from all the selected cells in the active spreadsheets.

Selecting and reviewing your data

After you define the assumption, decision variable, and forecast cells and return to the spreadsheet, you might want to check the assumptions, decision variables, and forecasts to make sure you have defined the cells as you intended. With Crystal Ball, you can select all the assumption, forecast, and decision variable cells and check their corresponding dialogs in sequence. If you find a cell that is not defined correctly, you can redefine it.

To review all assumption cells:



1. **Select Cell > Select All Assumptions.**

   Crystal Ball selects each cell defined as an assumption.

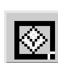
2. **Select Cell > Define Assumption.**

   The distribution dialog for the first assumption cell appears.

3. **Change the distribution (optional).**
4. **Click on OK.**

   If you have more than one assumption cell, each assumption appears in turn. Repeat steps 3 and 4 to check the assumptions for each cell.

To review all decision variable cells:



1. **Select Cell > Select All Decisions.**

   Crystal Ball selects each cell defined as a decision variable.

2. **Select Cell > Define Decision.**

   The Define Decision Variable dialog for the first decision variable cell appears.

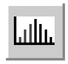
3. **Change the decision variable parameters (optional).**
4. **Click on OK.**

   If you have more than one decision variable cell, each decision variable appears in turn. Repeat steps 3 and 4 to check the decision variables for each cell.

To review all forecast cells:

1. **Select Cell > Select All Forecasts.**

   Crystal Ball selects each cell defined as a forecast.



2. **Select Cell > Define Forecast.**

   The forecast dialog for the first forecast cell appears.

3. **Change the forecast definition (optional).**

4. **Click on OK.**

   If you have selected more than one forecast cell, each forecast appears in turn. Repeat steps 3 and 4 to check the forecast for each cell.

The Cell > Select Some command lets you view, select, or define assumptions, decision variables, and forecasts from a list. You can use the Select Some command to:

- Quickly locate assumptions, decision variables, and forecasts on your spreadsheet.
- Modify assumptions, decision variables, and decision variables chosen from a list.

The Select Some dialog lets you choose from either a list of assumptions, a list of decision variables, or a list of forecasts.

Clicking on Select highlights the chosen assumptions, decision variables, or forecasts on the spreadsheet.

Clicking on Define brings up the Define Assumption, Define Decision Variable, or Define Forecast dialog for each chosen item.

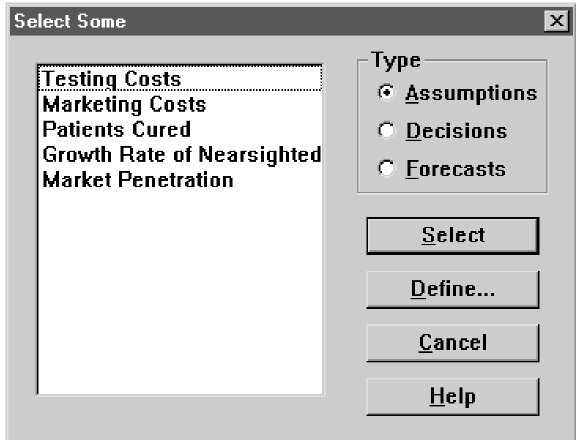


Figure 3.14 Select Some dialog

To select and define assumptions, decision variables, or forecasts from a list:

1. **Select Cell > Select Some.**

   The Select Some dialog appears listing all of the assumptions in your spreadsheet, as in Figure 3.14.

2. **Choose the type of data you want to select from by clicking either Assumptions, Decisions, or Forecasts.**
3. **Select the items you want to work with from the list.**
4. **Click on Select or Define to select or redefine the selected items.**

Setting cell preferences

Crystal Ball lets you change the appearance of assumption, forecast, and decision variable cells so you can quickly identify them on your spreadsheets. You can set Crystal Ball to change the appearance of these cells as you define them or you can change the appearance of predefined cells.

Pattern — Changes the pattern of each Crystal Ball data cell. You can change the pattern for each cell type separately. The new pattern applies to any existing and new Crystal Ball cells.

Color — Changes the color of each Crystal Ball data cell. You can change the color for each type separately. The new color applies to any existing and new Crystal Ball cells.

Add Excel Note — Adds an Excel note that describes the assumption, forecast, or decision variable.

For assumption cells, the note includes the assumption's name and its distribution parameters. For forecast cells, the note just includes the forecast's name. For decision variables, the note includes the decision variable's name, the minimum and maximum values, and step size, if applicable.

Crystal Ball Note: *Crystal Ball only updates cell notes when you define or redefine an assumption, decision variable, or forecast.*

Set To Distribution Mean — Changes the assumption cell value to the distribution mean, when no simulation is running.

Set To Range Midpoint — Changes the decision variable cell value to the range midpoint when no optimization is running.

Setting run preferences

With the Run Preferences dialog you can change a number of factors that determine how Crystal Ball runs a simulation. The options are grouped on different Run Preferences dialog pages, and you can open any of these by clicking on the corresponding button:

- Trials
- Sampling
- Speed
- Macros
- Options
- Turbo

To change any preferences:

1. **Select Run > Run Preferences.**
2. **Click on the button with the preferences you want to change.**
3. **Set, select, check, uncheck, or change any preferences you want to change.**
4. **Repeat steps 2 and 3 for any other preferences you want to change on a different dialog.**
5. **Click on OK.**

Trials preferences

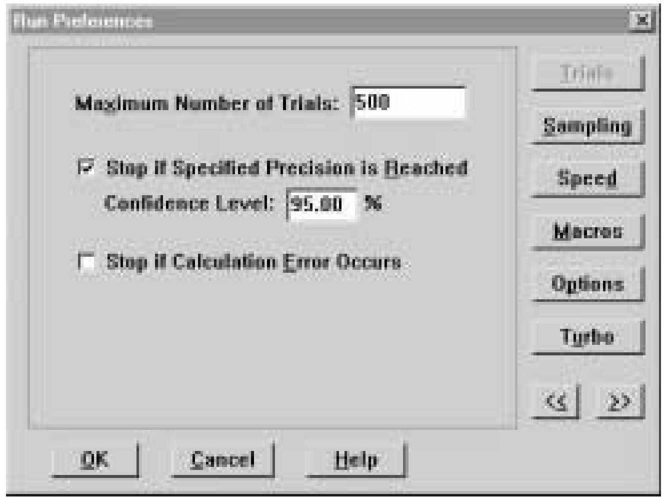


Figure 3.15 Run Preferences Trials dialog

The preferences on the Trials dialog are:

Maximum Number Of Trials
: Defines the total number of trials that Crystal Ball runs before it stops the simulation. If you check either of the options on this dialog, Crystal Ball only uses the maximum number of trials if forecast results do not meet the other Stop criteria first.

Stop If Specified Precision Is Reached
: Stops the simulation when the statistics reach a specified precision. You define the precisions that trigger this option in each Define Forecast dialog. Any forecasts set to use precision control must all reach their specified precision (within the confidence level) to stop the simulation.

: This is On by default. If all the forecasts set to use precision control don't meet the specified precision, the simulation stops when it reaches the Maximum Number Of Trials.

Confidence Level

Is used in the calculation of the confidence intervals for the Precision Control statistics. If you select the Stop If Precision Is Reached option, Crystal Ball stops the simulation when the specified statistics for all the forecasts set to use precision control (indicated in the Define Forecast dialogs) fall within the calculated confidence interval.

Crystal Ball Note: *For more information on how Precision Control uses the Confidence Level, see Chapter 5,* **"Maximizing Your Use of Crystal Ball"** *and Appendix B, "Equations and Methods".*

Stop On Calculation Error

Stops the simulation when a mathematical error (such as division by zero) occurs in any forecast cell.

If no calculation errors occur, the simulation continues until it reaches the Maximum Number Of Trials or (if set) when the specified precision is reached.

Sampling preferences

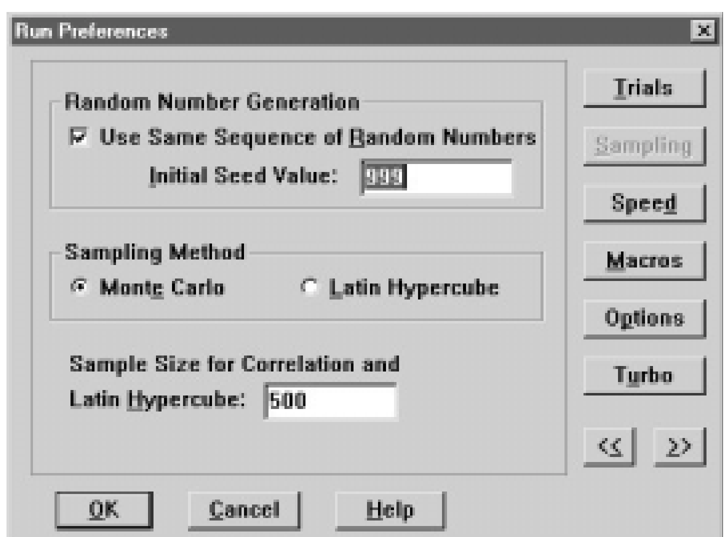


Figure 3.16 Run Preferences Sampling dialog

The Sampling preferences are:

Use Same Sequence Of Random Numbers
: Sets the random number generator to generate the same set of random number for assumptions, letting you repeat simulation results.

  When you select this option, enter a seed value in the Initial Seed Value field.

Crystal Ball Note: *To reproduce the sample results shown in this manual, check this option and use a seed value of 999.*

Initial Seed Value
: Determines the first number in the sequence of random numbers generated for the assumption cells.

Sampling Method
: Sets whether to use Monte Carlo or Latin hypercube simulation. Latin hypercube sampling generates values more evenly and consistently across the distribution, but requires more memory. For more information, see “Sampling methods” on page 113.

Sample Size For Correlation And Latin Hypercube
: Sets the number of assumption values in a group generated ahead of time for use with correlated assumptions or Latin hypercube sampling. The default is 500. The maximum is 5000.

  A higher value increases the number of values generated ahead of time, increasing accuracy but requiring extra memory. The accuracy of both features increases at roughly a linear rate with increases in sample size. For more information, see “Simulation accuracy” on page 243.

  Crystal Ball might automatically increase this number to make the sample size at least as large as the number of correlated assumptions.

CB Turbo Note: *CB Turbo also uses the Sample Size option. If you use CB Turbo, see your CB Turbo User Manual for information on using this option with that product.*

Speed preferences

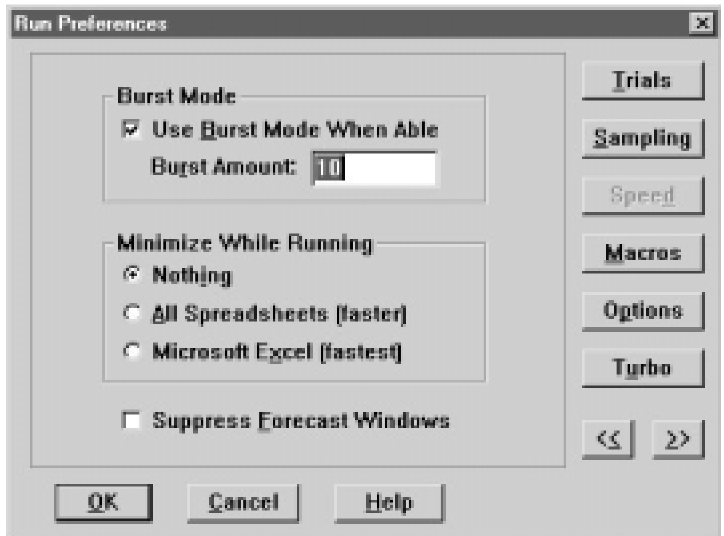


Figure 3.17 Run Preferences Speed dialog

The Speed preferences are:

Use Burst Mode When Able
: Makes Crystal Ball batch messages when it communicates with Excel. This decreases the amount of time required to run a simulation, but also reduces the responsiveness of the system.

: Disable this option for large models or if you plan to work on other parts of the system during simulations.

Burst Amount
: Sets the size of the message batches Crystal Ball uses to communicate with Excel. This must be a whole number greater than 0. The default is 10.

Minimize While Running
: Sets whether to minimize spreadsheets, Excel, or nothing during simulations. Minimizing spreadsheets or Excel speeds up simulations. Minimizing Excel speeds up simulations the most.

Suppress Forecast Windows
: Sets whether forecast windows appear during a simulation. If you don't check this option, during a simulation, forecast windows that were originally defined to appear during the simulation appear normally. If you check this option, you can only view forecast windows after a simulation using the Open Forecast Windows dialog.

Macros preferences

The Run Preferences Macros dialog lets you select command macros to run before the simulation begins, during a simulation (either before or after each spreadsheet recalculation), or after the simulation ends.

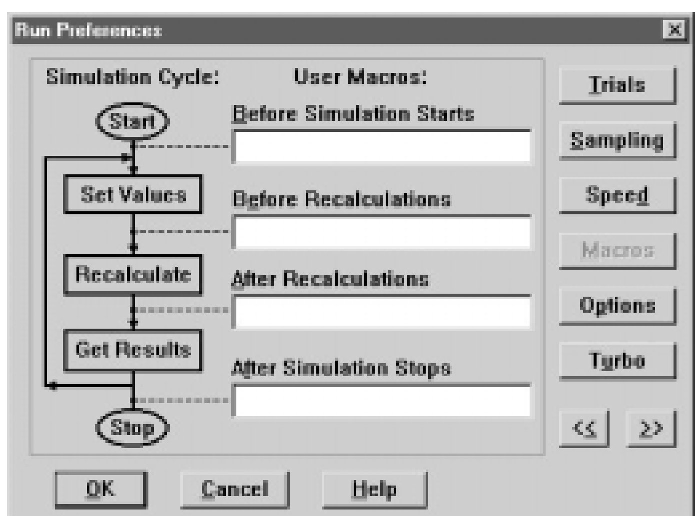


Figure 3.18 Run Preferences Macros dialog

The diagram on the left side of the dialog shows the parts of the Crystal Ball simulation cycle. In the middle, under the heading User Macros, are four text fields where you can enter the macro names to execute during the simulation.

Crystal Ball Note: *You cannot call Crystal Ball macros (or macros containing Crystal Ball macros) from this dialog.*

The text fields are:

Before Simulation Starts
: Runs the macro only once before the simulation begins, immediately after you select the Run command. This macro doesn't run again if you continue the simulation.

Before Recalculation
: Runs the macro once before every simulation trial, before Crystal Ball recalculates the spreadsheet.

After Recalculation
: Runs the macro once after every simulation trial, after Crystal Ball recalculates the spreadsheet.

After Simulation Stops
: Runs the macro only once after the simulation, immediately after you select the Stop command or the simulation stops itself.

For more information on writing macros, see your Excel user documentation.

Excel Note: *Use the following macro name format:*

VBA macro: book1.xls!sheet1.macro1

Options preferences

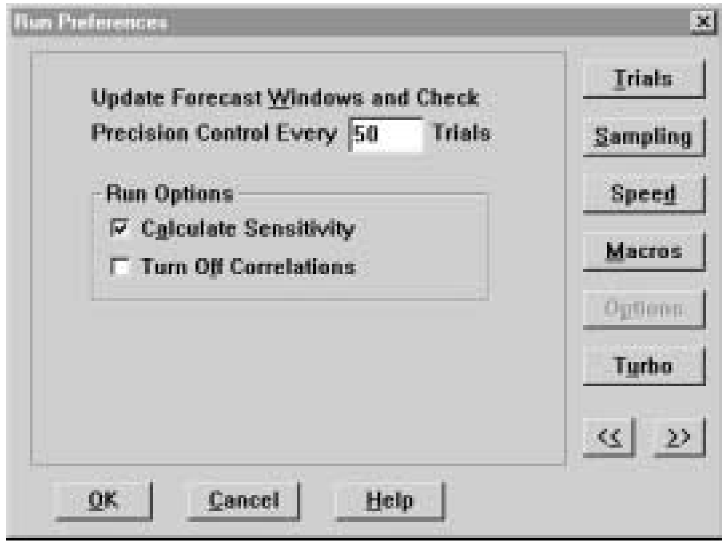


Figure 3.19 Run Preferences Options dialog

The Options preferences are:

Update Forecast Windows And Check Precision Control
: Controls how often Crystal Ball redraws the forecast charts and checks precision control during a simulation, e.g., every 50 trials or every 500 trials. The default is every 50 trials.

Calculate Sensitivity
: Sets Crystal Ball to generate sensitivity information during simulations. This information appears in the sensitivity chart to display the influence each assumption has on a particular forecast. For more information on the sensitivity chart, see "Understanding sensitivity and using the sensitivity chart" on page 223.

: The sensitivity chart is not available unless you selected this option before you run your simulation.

Turn Off Correlations
: Disables any defined correlations.

Turbo preferences

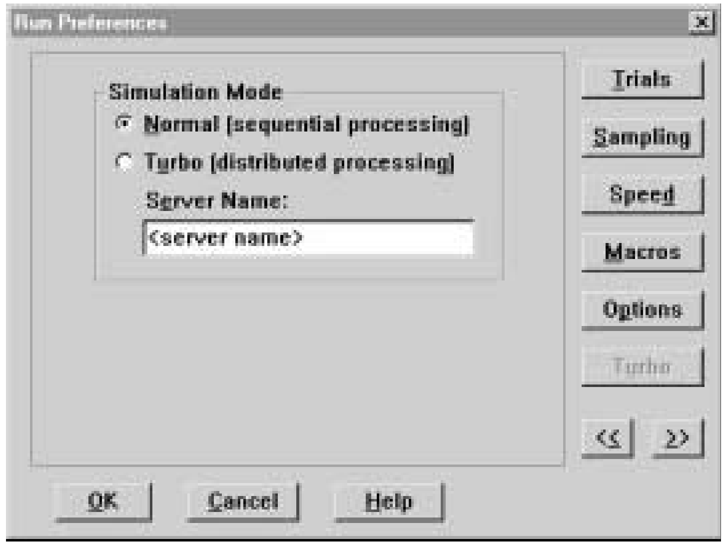


Figure 3.20 Run Preferences Turbo dialog

The Turbo preferences are:

Normal — Sets to run simulations in normal mode (without CB Turbo).

This is the default, and the only option you should use unless you have CB Turbo. For more information about CB Turbo, see the ABOUTME.DOC file in the Crystal Ball\CBTurbo folder.

Turbo — Sets to run simulations in turbo mode, which requires CB Turbo. For more information about CB Turbo, see the ABOUTME.DOC file in the Crystal Ball\CBTurbo folder.

Server Name — Indicates which server and port number to use for turbo simulations. The syntax is

<host_name>:<port_number>

where <host_name> is the CB Turbo server host name and <port_number> is the port number on the server that CB Turbo uses, usually 24158. See your CB Turbo network administrator to change your server settings.

Running the simulation

After you define assumption, forecast, and decision variable cells in the spread-sheet model, you are ready to run the simulation. Crystal Ball uses a technique called Monte Carlo simulation to forecast the entire range of results most likely to occur in the situation you have defined in your spreadsheet model. This technique involves generating random numbers for the assumption variables. While the simulation is running, Crystal Ball displays these assumption results in a forecast chart that shows the entire range of possible outcomes. Crystal Ball provides descriptive statistics for any forecast, which summarize the results numerically. Descriptive statistics are discussed in Chapter 4, **"Interpreting the Results"**.

How Crystal Ball uses Monte Carlo simulation

Crystal Ball implements Monte Carlo simulation in a repetitive three-step process. For each trial of a simulation, Crystal Ball repeats the following three steps:

1. **For every assumption cell, Crystal Ball generates a number according to the probability distribution you defined and places it into the spreadsheet.**
2. **Crystal Ball recalculates the spreadsheet.**
3. **Crystal Ball then retrieves a value from every forecast cell and adds it to the graph in the forecast windows.**

Although Monte Carlo simulation is meant to portray a real-world situation, keep in mind that the spreadsheet model can only approximate the elements of the real world. As you examine the problem, you can eliminate certain variables and add others to the spreadsheet model. Continue to refine the model until you feel it describes the real-world situation as accurately as possible.

Simulation overview

When running a simulation in Crystal Ball, you can stop, continue, or reset the simulation at any time. You can display the forecast chart for each forecast cell or run the simulation with the forecast windows closed. You can select windows, cascade windows, open all of the forecast charts at the same time, or bring the spreadsheet or a forecast window to the front of the other windows.

You can even continue to work in Excel as you normally would while the simulation is running, with the following exceptions:

- Do not change the assumption or forecast cells.
- Do not change the calculation options.
- Do not close spreadsheets that contain assumptions or forecasts.
- Do not exit Excel.

While the simulation is running, Crystal Ball creates a forecast chart for each forecast cell using frequency distributions. A frequency distribution shows the number or frequency of values occurring in a given group interval.

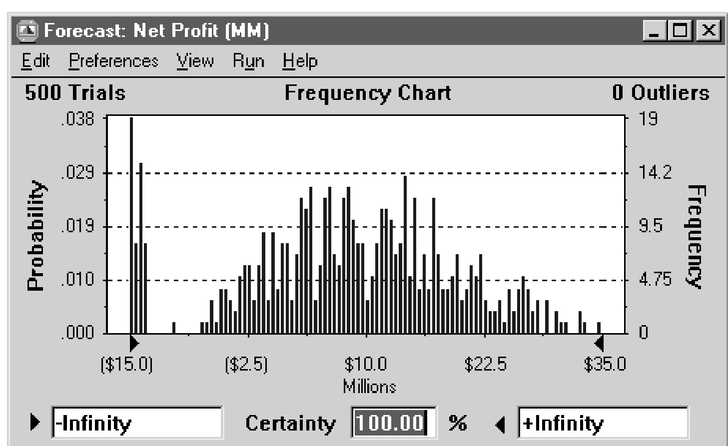


Figure 3.21 Net Profit Forecast dialog

In Chapter 4, you will find a more detailed explanation on how to interpret the forecast chart. For now, remember that the chart displays the forecast results, showing you how the forecast values are distributed and the likelihood of achieving a given result. As you run the simulation, observe the shape of the distribution to get a feel for the problem you are trying to solve. In Figure 3.21 (frequency distribution for Net Profit), for example, you see that most of the values are clustered around central group intervals. You also see the variation around these group intervals and a clump of values at the right end.

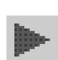
Excel Note: *Each forecast window has its own set of Edit, Preferences, View, Run, and Help menus with the Run menu duplicating the Run menu on Excel's menu bar. You can use whichever menu bar is the most convenient. Clicking on the Excel menu bar brings Excel forward and makes the forecast windows disappear. If this happens, you can quickly bring the forecast windows back to the front by pressing <Alt>-<Tab>.*

During the simulation, Crystal Ball saves the forecast values in a list that grows as the simulation progresses. Later in this chapter, you will learn how to export these forecast values to other programs, such as a statistical analysis program. Chapter 4 also explains how to focus on a particular range of forecast values.

Running/Stopping/Continuing/Resetting

As you run a simulation, you can stop, continue, or reset at any time. Crystal Ball runs trials automatically, but you also have the option of observing each individual trial to see how the values change.

To run a simulation:



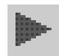
1. **Select Run > Run.**

   If you set the forecast to appear in the Define Forecast dialog, a forecast window appears. As the simulation proceeds, the frequency distribution is updated on the forecast chart to reflect the changing values in the forecast cell.

Crystal Ball Note: *The Total Trials counter in the lower left corner of Excel's status bar shows the simulation's progress.*

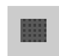
2. **To stop a simulation, Select Run > Stop (use the Run menu on the front-most forecast window).**

To continue a simulation:



3. **Select Run > Continue.**

To reset the simulation and start over again:



4. **Select Run > Stop.**

5. **Select Run > Reset.**

   A dialog appears with a message asking you to confirm your request to reset the simulation.

6. **Click on OK.**

   Crystal Ball resets the number of trials to 0 and clears the list of forecast values and statistics for each forecast. However, the assumption and forecast definitions remain.

7. **Change any assumptions or forecasts as needed.**

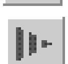
8. **Select Run > Run.**

   The simulation starts over again.

Before you run a simulation or after you have stopped it, you can use the Single Step command to watch the Monte Carlo process generate one set of values at a time for the assumption cells and recalculate the spreadsheet. This feature is useful if you are trying to track down a calculation error or to verify that the values being produced for your assumption cells are valid.

To observe an individual trial (you can only do this when the simulation is not running):

1. **Select Run > Reset.**



2. **Select Run > Single Step to execute one iteration of the simulation.**

Crystal Ball Note: *Single Step does not create or update forecast windows; it only produces values for the assumption cells and recalculates the spreadsheet.*

Closing forecasts

You can close the forecast windows at any time and the simulation will continue. Running a simulation with the forecast windows closed decreases the time required to run the simulation.

To close a forecast window, click on the Close icon in the top right corner of the forecast window.

Managing windows

While the simulation is running, you can use the Run > Forecast Windows command to open or close all the forecast windows, open or close selected forecast windows, or cascade the forecast windows.

To manage forecast windows:

1. Select Run > Forecast Windows.

A dialog appears that gives you the following options:

Open All Forecasts
: Click this to open all forecast windows.

Close All Forecasts
: Click this to close all forecast windows.

Open Selected Forecasts
: Select the forecasts you want to open and click Open Selected Forecasts. You can select as many forecasts as you want before clicking Open Selected Forecasts.

Close Selected Forecasts
: Select the forecasts you want to close and click Close Selected Forecasts. You can select as many forecasts as you want before clicking Close Selected Forecasts.

Cascade Forecasts
: Click this to neatly stack all windows in front of the spreadsheet.

Responding to problems with correlated assumptions

If Crystal Ball detects inconsistently correlated assumptions, it first determines whether small adjustments to the correlation coefficients are possible. This process might take a long time, depending on the number of correlated assumptions. Crystal Ball displays the message "Examining the Correlation Coefficients". If you get this message, you should probably stop and redefine your correlations.

If small adjustments to the correlation coefficients are possible, a dialog appears, letting you decide whether to cancel the simulation or continue with the adjusted coefficients.

Select one of the following responses:

- Click on Adjust Coefficients This Time Only to continue the simulation with the adjusted coefficients.

  Since adjusted coefficients are not saved permanently, a dialog appears again if you stop and restart the simulation.

- To continue the simulation, replacing your original correlation coefficients with the adjusted ones, click on Adjust Coefficients Permanently.

Crystal Ball Note: *Correlation coefficients that were specified using a cell reference in place of an actual value are replaced with a permanent value.*

- To terminate the simulation, click on Cancel.

If small adjustments to the correlation coefficients are not possible, the simulation stops with an error message prompting you to reexamine your coefficients. To solve the problem, you can usually make large coefficients smaller or change your spreadsheet model to use formulas to calculate assumptions with large correlation coefficients. Creating a report containing just your assumptions might make it easier to spot problems.

Saving and restoring a simulation

You can save your Crystal Ball simulation to disk and later restore it. This is useful in several ways:

- You can run a simulation, set the chart settings the way you want them, save your simulation, and go on to other work. You can then, at a moment's notice, call up the forecast windows and show them to someone else without running the whole simulation again.
- You can open a saved run and graph its results in a different way, or use the Extract Data and Create Report features.
- If you have a simulation that takes a particularly long time to run, you can stop the simulation, save it to disk, and continue it. That way, should the simulation be interrupted (by a power failure, for instance) you can restore the simulation from the last time you saved it and continue processing as if nothing had happened.
- You can copy a saved simulation and its spreadsheets to a floppy disk and open them on another computer where Crystal Ball is installed.

Saving a Crystal Ball simulation

To save a Crystal Ball simulation, select Run > Save Run. You can only save a run after a simulation stops. Crystal Ball uses the name of the topmost spreadsheet for the default name of the save run file.

The Save Run dialog appears, as in Figure 3.22. It lets you determine the name and location of the saved run file.

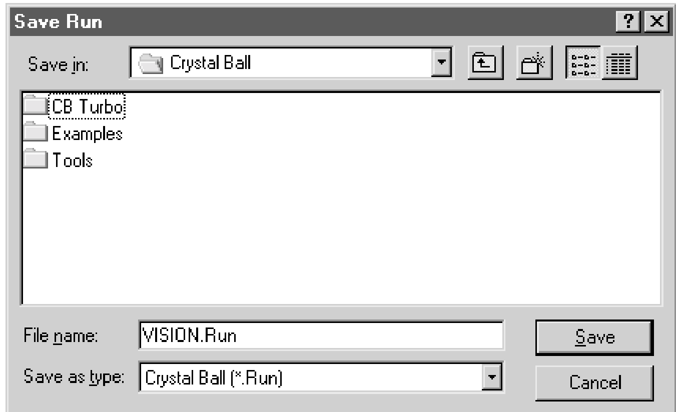


Figure 3.22 Save Run dialog

Crystal Ball saves a lot of information about your simulation to the file including a list of all the spreadsheets that were open that contain one or more assumptions or forecasts. If there is a need to save other spreadsheets that must automatically reopen when you restore a simulation, simply create a dummy assumption on these other spreadsheets: define a cell in the sheet as a uniform distribution.

Crystal Ball checks to make sure that all the sheets in the saved run have been saved to disk. If you have created a new sheet and have added assumptions or forecasts to it, Crystal Ball asks you to save it first.

Crystal Ball also records which spreadsheet was active when you saved the run. Before you save your run, you can click on the most important spreadsheet to bring it forward. Then when you restore the run, Crystal Ball will bring that sheet to the front.

Crystal Ball Note: *Saving a run does not save sensitivity data. You can only save part of the sensitivity information by creating a report with the sensitivity chart included and then saving the report.*

Once you have saved the simulation you can continue working with Crystal Ball.

Restoring a Crystal Ball simulation

To restore a Crystal Ball simulation that you saved earlier, select Run > Restore Run. The Restore Run dialog appears, as in Figure 3.23.

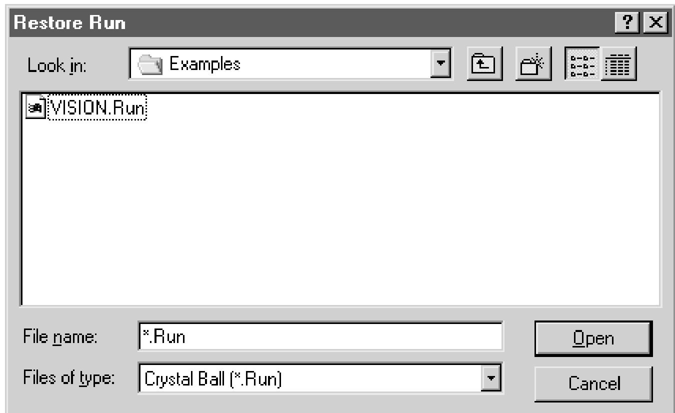


Figure 3.23 Restore Run dialog

Crystal Ball Note: *You cannot restore a simulation if the current simulation is not reset first.*

The file list contains only files that have the RUN file type. If you try to open a file that is not a Crystal Ball saved run, or if Crystal Ball recognizes that a saved run file has somehow been damaged, you will not be able to open the file, and a dialog appears containing the message, “Sorry, that file is not a Crystal Ball saved run.”

This simply means that Crystal Ball did not recognize the contents of the file. Click on OK to continue.

Crystal Ball opens all the spreadsheets associated with the saved run file if they are not already open.

If you are restoring a large run file with many spreadsheets and forecasts, you might get an “out of memory error.” Check your computer to determine the amount of free memory available, then release it by quitting other applications. Check free memory again before retrying your run file.

Saving your models

The distributions you specified for each assumption cell, the settings you specified for each forecast cell, and the range information for each decision variable cell are saved with their spreadsheet through the Excel save process.

When you later reopen the spreadsheet, Crystal Ball "remembers" your assumption, forecast, and decision variable cells, so you can pick up where you left off.

This chapter provided the step-by-step instructions you need to run a simulation in Crystal Ball. You are now ready to read Chapter 4 to learn how to interpret the results of your simulation.

Closing Crystal Ball

To close Crystal Ball without exiting Excel, select Run > Close Crystal Ball.

A dialog appears asking you to confirm your decision. If you click on Yes, the Crystal Ball menus are removed from the menu bar and Crystal Ball closes.

Crystal Ball also closes automatically when you exit Excel.

Chapter 4

Interpreting the Results

In this chapter

- Understanding the forecast chart
- Determining the certainty level
- Focusing on the display range
- Customizing the forecast chart
- Interpreting the statistics
- Understanding and using the overlay chart
- Understanding and using the trend chart
- Understanding sensitivity and using the sensitivity chart
- Changing chart patterns and colors
- Creating reports
- Extracting data

This chapter helps you interpret the forecast results, either during or after the simulation. As you read through the chapter, you will learn to focus on a particular range of the forecast results, customize the forecast chart, interpret the descriptive statistics, and print comprehensive reports.

You will also learn how to develop an overlay chart to view forecasts superimposed over each other, to develop and customize a trend chart to view the certainty ranges of all your forecasts on a single chart, and to develop a sensitivity chart to view the impact your assumptions have on the forecast results.

Understanding the forecast chart

While a simulation is running, Crystal Ball creates a forecast chart for each forecast cell. The forecast chart displays the results of the simulation graphically in an easily understood format.

As discussed in the section “Simulation overview” on page 173, the forecast chart uses frequency distributions to show the number (frequency) of values occurring in a given interval. The highest value on the frequency scale to the right of the chart is the frequency for the group interval that contains the greatest number of forecast values—the mode of the frequency distribution. You can estimate the frequency for other points on the forecast chart using this frequency scale.

The highest value on the probability scale to the left of the chart is the probability for the mode. You can estimate the probability for other points on the forecast chart using this probability scale.

In the example below, the mode (on the extreme left of the distribution) has a frequency of 19, meaning that there are 19 values in the interval expressed by that column. The mode has a probability of 0.038 (or 3.8%), meaning that there is a 3.8% chance of a value falling within this interval.

The elements of the forecast chart appear in Figure 4.1.

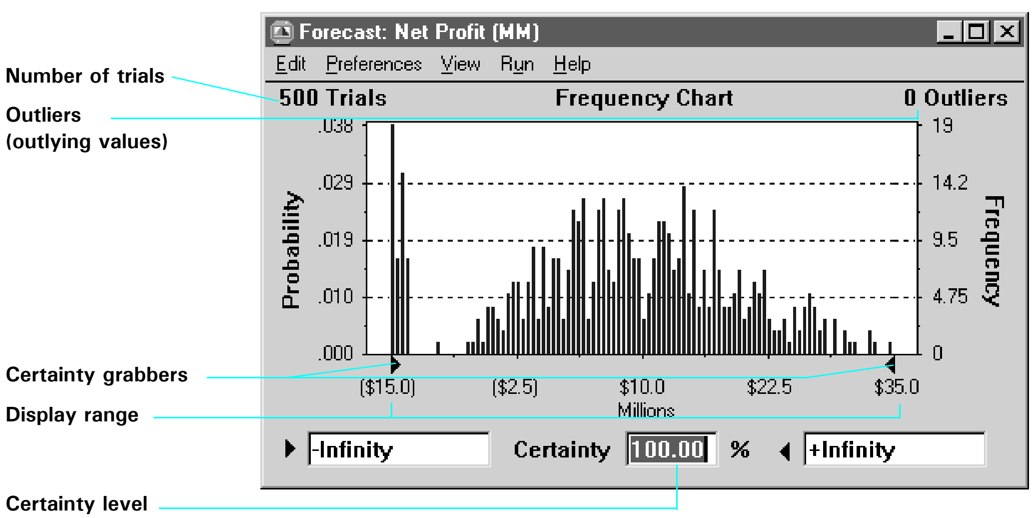


Figure 4.1 Forecast chart

Crystal Ball forecasts the entire range of results for a given situation. However, the forecast chart shows only the display range, which is a subrange of the whole. By default, the display range includes all values within 2.6 standard deviations of the mean (approximately 99% of the forecast values). Crystal Ball then rounds the display range to the next even number of units. For this reason, outlying values might be excluded from the display range.

Crystal Ball Note: *To display outliers, change the display range preferences (from the Forecast chart, select Preferences > Display Range) to display fixed endpoints between -Infinity and +Infinity.*

The number of outlying values is shown in the upper right corner. The number of trials run for this forecast appears in the upper left corner of the forecast chart. The display range is the linear distance from ($15.00) to $35.00, as shown in Figure 4.1.

The forecast chart also shows the certainty range for the forecast. In Crystal Ball, the certainty range includes all values between the certainty grabbers. In the next section, you will learn how to change the certainty range.

Crystal Ball Note: *When the certainty grabbers are at -infinity and +infinity, every forecast value is included in the certainty range regardless of the size of the display range.*

Crystal Ball compares the number of values in the certainty range with the number of values in the entire range to calculate the certainty level for the forecast. The previous example shows a certainty level of 100% since the default certainty range includes all possible values.

By default, Crystal Ball calculates the certainty level based on the entire range of forecast values. To calculate the certainty level based just on the display range, in the forecast window select Preferences > Statistics to display the Statistics Preferences dialog. In this dialog, select the Base Statistics On Display Range option.

On either side of the Certainty field, the range minimum and the range maximum for the certainty range appear.

Determining the certainty level

The certainty level is one of Crystal Ball's key statistics because it shows the certainty of achieving the values within a specific range. With Crystal Ball, you can determine the certainty level for specific value ranges either by moving the certainty grabbers on the forecast chart or typing the range minimum and range maximum in the fields. You can also type any certainty level in the certainty field to get a certainty range centered around the mean (or median).

When you move the certainty grabbers, the certainty range changes and Crystal Ball recalculates the certainty level. When you type minimum and maximum values, Crystal Ball moves the certainty grabbers for you and recalculates the certainty level. When you type the certainty level in the certainty field, Crystal Ball moves the certainty grabbers to show you the value range for the certainty level you specified. The certainty level shows you the certainty of achieving the values between the certainty grabbers.

To determine the certainty level for a specific value range using certainty grabbers:

1. **Make sure the forecast chart you want to use is the front window.**
2. **Move the certainty grabbers on the forecast chart.**

   The range minimum field shows the value that corresponds to the position of the left certainty grabber; the range maximum field shows the value that corresponds to the position of the right certainty grabber. The Certainty field shows the certainty level for the area between the certainty grabbers. Crystal Ball shades the columns outside the certainty grabbers a different color to show that those values have been excluded, as illustrated in Figure 4.2.

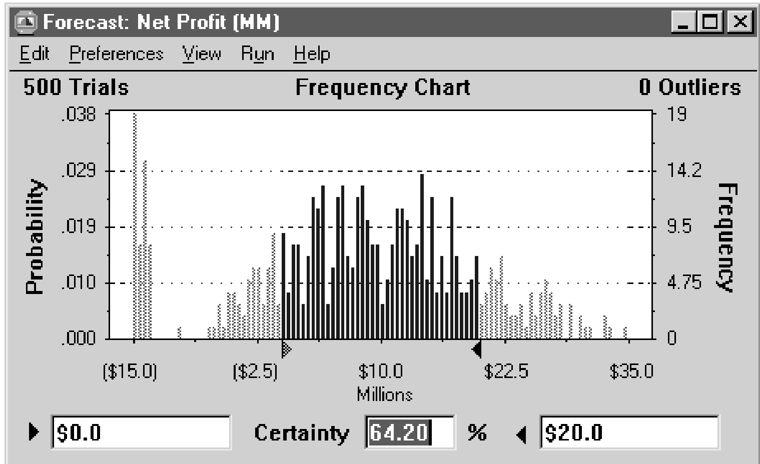


Figure 4.2 Certainty level: values $0 to $20

The Net Profit forecast chart in Figure 4.2 is the same as the example preceding it, except that the certainty grabbers have been moved. The range minimum shows $0.0 and the range maximum shows $20.0. The key statistic is the certainty level of 64.20% in the Certainty field. By moving the certainty grabbers, you have changed the certainty range. Crystal Ball compares the number of values lying within the certainty range to the number of values in the entire range—from negative $15 million to positive $35 million—to recalculate the certainty level. With a certainty level of 64.20%, you can be 64.2% confident of making a net profit between $0 and $20 million.

When you type values in the range minimum and range maximum fields on the forecast chart, Crystal Ball moves the certainty grabbers for you and recalculates the certainty level.

To determine the certainty level for a specific value range using the range minimum and range maximum fields:

1. **In the forecast chart, Type a value in the range minimum field.**
2. **Press <Tab> twice and type a value in the range maximum field.**

3. **Press <Enter>.**

   The certainty grabbers move to correspond to the values you entered. The Certainty field shows the certainty level for the area between the range minimum and the range maximum. Crystal Ball shades the columns outside the certainty grabbers a lighter color to show that those values have been excluded.

Alternately, you can type the certainty level on the forecast chart and let Crystal Ball calculate the minimum and maximum for you. Make sure that the certainty grabbers are free (whenever you click on an certainty grabber, it turns to a lighter color and is considered anchored). To free the lighter color certainty grabber, click in the chart area. Then when you type in a certainty, it will be centered on the mean (or median, depending on the setting in the Statistics Preferences dialog).

In the following procedure you will enter a certainty level of 10% and watch Crystal Ball move the certainty grabbers:

4. **Type 10 in the Certainty field.**
5. **Press <Enter>.**

   Crystal Ball moves the certainty grabbers to include 10% of the values centered around the mean of the entire range.

You can also anchor an certainty grabber, type the certainty level, and Crystal Ball moves the free grabber to correspond to the value range for the level.

6. **Move the left certainty grabber until the minimum range value is $0.0.**

   The certainty grabber turns to a lighter color and is considered anchored.

7. **Press <Tab> and type 75 in the Certainty field.**
8. **Press <Enter>.**

   Crystal Ball moves the free certainty grabber toward the anchored certainty grabber to include 75% of the values as in Figure 4.3.

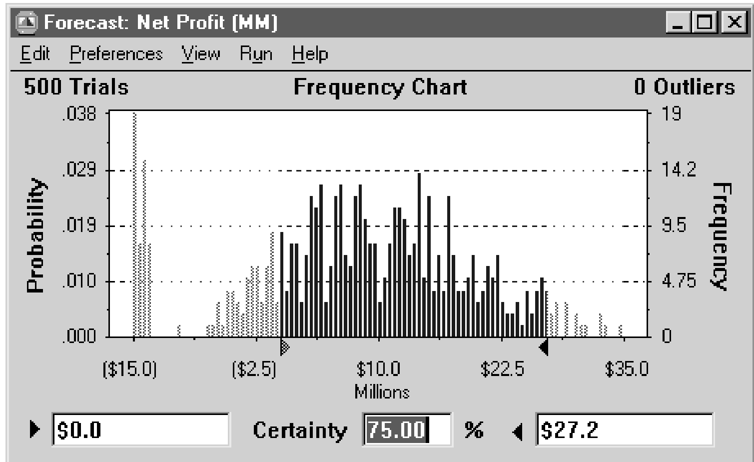


Figure 4.3 Certainty level: 75% from left certainty grabber

You also can cross over the certainty grabbers to determine the certainty level for the two ends.

You can determine the certainty level for specific value ranges at any time, either during or after the simulation.

To reset the original certainty range before beginning the next section:

1. **In the forecast chart, either:**
   - Move the right certainty grabber until the range maximum field shows positive infinity.
   - Type +Infinity in the right field.
2. **Either:**
   - Move the left certainty grabber until the range minimum field shows negative infinity.
   - Type -Infinity in the left field.

   The forecast chart displays a certainty level of 100%, as displayed in the original forecast.

Focusing on the display range

With Crystal Ball, you can focus on a particular range of the forecast results by changing one of the options on the Display Range Preferences dialog. When you select an option, the display range of the forecast changes to reflect the option you selected. As you change the display range, you can compare the probabilities from one range to another to determine the probability of attaining a particular value. The Apply To All button in the Display Range Preferences dialog lets you apply an option to all other forecasts.

To select one of the options on the Display Range Preferences dialog, make sure the forecast chart you want to change is the front window and either:

- Select Preferences > Display Range
- Double-click the axis values on the forecast chart

The Display Range Preferences dialog appears, as in Figure 4.4.

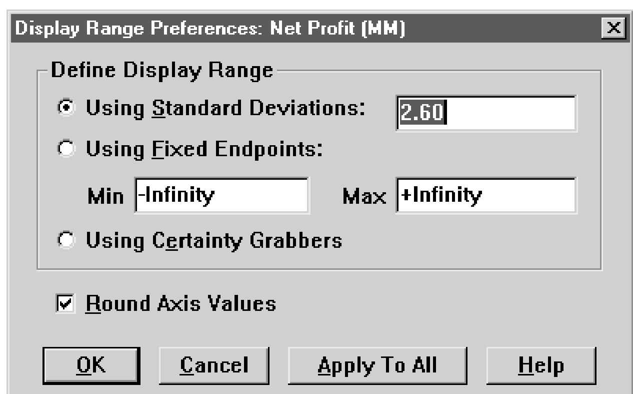


Figure 4.4 Display Range Preferences dialog

You can define the display range in three ways:

Using Standard Deviations
: Defines the number of standard deviations worth of value you want to display.

Using Fixed Endpoints
: Sets the display range endpoints manually.

Using Certainty Grabbers
: Sets the display range to the current certainty grabber positions on the Forecast Window.

There is also a Round Axis Values option that constrains the display range to round numbers. Each item is described in the following sections.

Defining the display range using standard deviation

Crystal Ball uses a default display range of 2.6 standard deviations from the mean, which includes about 99% of the forecast values. Click the Using Standard Deviations option if you want to change the standard deviation to focus on values centered around the mean. For example, you could change the display range to 1 standard deviation from the mean to look at approximately 68% of the forecast values.

The forecast chart redraws with the new display range defined by the standard deviations you entered. If your statistics preferences are set to base the statistics on the display range, the number of trials changes to correspond to the number of values within the new display range. The number of total trials for the forecast remains the same.

Defining the display range using fixed endpoints

The Using Fixed Endpoints option lets you focus on particular value ranges. For example, you can focus on positive values only to look at the profit for a profit/loss forecast.

To define the display range using fixed endpoints:

1. **In the forecast window, select Preferences > Display Range.**

   The Display Range Preferences dialog appears.

2. **Select the Using Fixed Endpoints option.**

3. **Type the value for the minimum endpoint in the Min field.**

4. **Press <Tab>.**

5. **Type the value for the maximum endpoint in the Max field.**

Crystal Ball Note: *To see the entire range of forecast values (100%), use the default values in the Min and Max fields—negative infinity and positive infinity. If these values do not appear in the fields:*

- *For negative infinity, type -inf.*
- *For positive infinity, type +inf.*

Show the entire range only when the simulation is completed. Otherwise, Crystal Ball attempts to rebuild the forecast chart every time a new minimum or maximum value occurs, which can slow down the simulation significantly.

6. **Click on OK.**

   The forecast chart appears again, reflecting the values you entered. Figure 4.5 shows the Net Profit forecast chart with display range endpoints of $20.0 and $35.0.

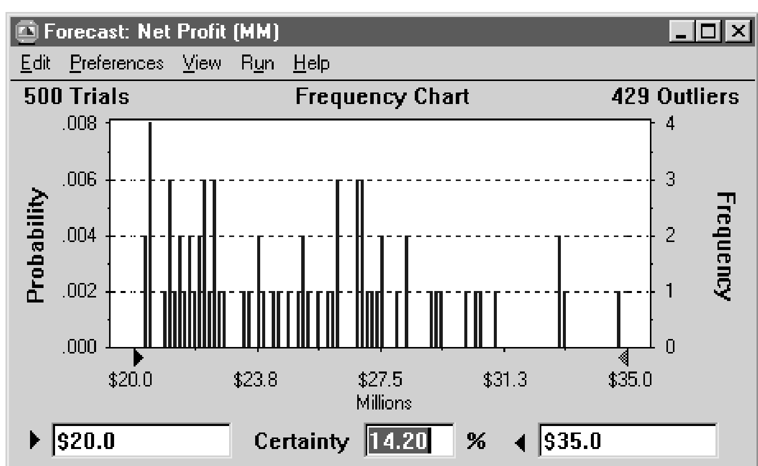


Figure 4.5 Display range: $20.0 to $35.0

Defining the display range using certainty grabbers

To define the display range using certainty grabbers:

1. **In the forecast chart, move the certainty grabbers to focus on a particular part of the display area.**
2. **In the forecast window, select Preferences > Display Range.**

   The Display Range Preferences dialog appears.
3. **Select Using Certainty Grabbers.**
4. **Click on OK.**

   The display range changes to reflect the location of the certainty grabbers.

Crystal Ball Note: *After Crystal Ball generates the new display range, the Certainty Grabber option automatically reverts back to the Fixed Endpoints option, the default setting.*

Changing the axis values to actual values

By default, the *x* axis values numbers are automatically adjusted to round numbers to make the forecasts easier to read.

To turn off automatic rounding of the axis value numbers:

1. **In the forecast window, select Preferences > Display Range.**
2. **Uncheck the Round Axis Values option.**
3. **Click on OK.**

Customizing the forecast chart

You can change the appearance of a forecast chart using the Chart Preferences and Format Preferences dialogs. Changing the appearance of the forecast helps you interpret the results of the simulation by viewing the data in different ways.

To use the Chart Preferences dialog:

1. **Make sure the forecast chart you want to change is the front window**
2. **In the forecast window, either:**
   - Select Preferences > Chart
   - Double-click on the chart

   The Chart Preferences dialog appears, as in Figure 4.6.

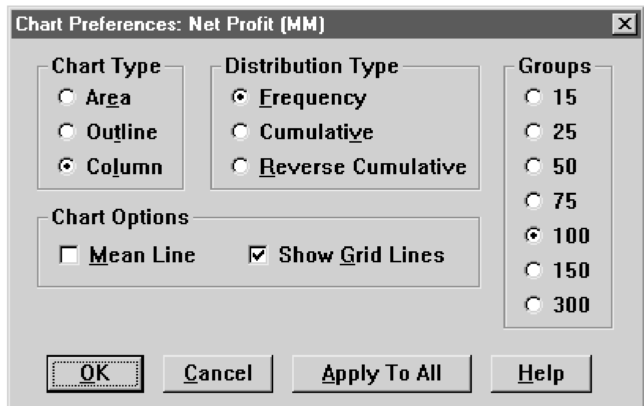


Figure 4.6 Chart Preferences dialog

The Chart Preferences dialog provides a variety of options for changing the appearance of the forecast chart. Change the appearance of the forecast chart by selecting one or more options at a time. Use the Apply to All button in the Chart Preferences dialog to apply an option to all other forecasts.

For some of the options, you can use hot keys to bypass the Chart Preferences dialog. The following sections describe these hot keys. They are also listed in Appendix C, “Keyboard Commands.”

Changing the chart type

Use the Chart Type options to select the type of forecast chart you want to display:

area chart	Shows the forecast results as darkened peaks and valleys.
outline chart	Shows an outline of the peaks and valleys.
column chart	Displays the forecast results as vertical columns that correspond to the group intervals of the distribution. The column chart is the default chart type.

To change the forecast chart type:

1. **In the forecast window, either:**
   - Select Preferences > Chart
   - Double-click on the chart

   The Chart Preferences dialog appears.
2. **Select a chart type.**
3. **Click on OK.**

   The forecast chart appears in the new chart type, such as outline, as in Figure 4.7.

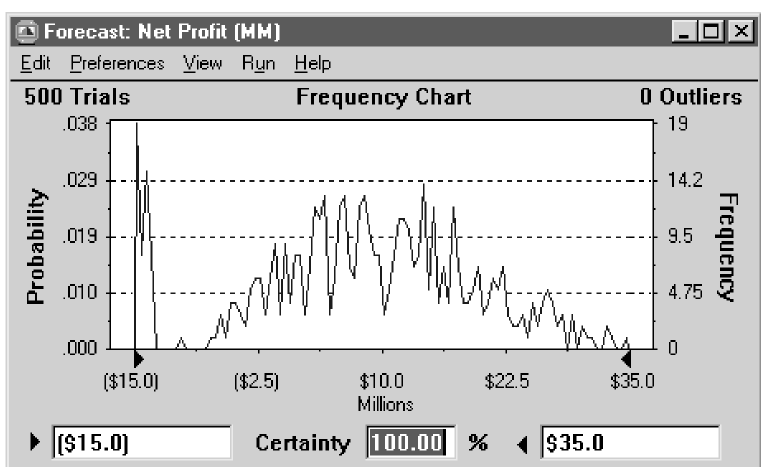


Figure 4.7 Forecast chart—outline type

__Crystal Ball Note:__ Use a Crystal Ball hot key to bypass the Chart Preferences dialog. Each time you press __t__, the chart type changes from an area chart, to an outline chart, to a column chart, and back again to an area chart.

Changing the distribution type

Use the Distribution Type options to select the type of distribution to display on the forecast chart:

frequency distribution
: Shows the number or frequency of values occurring in a given interval. This is the default distribution type.

cumulative distribution
: Shows the number or proportion (percentage) of values less than or equal to a given amount.

reverse cumulative distribution
: Shows the number or proportion (percentage) of values greater than or equal to a given amount.

To change the forecast distribution type:

1. **In the forecast window, either:**
   - Select Preferences > Chart
   - Double-click on the chart

   The Chart Preferences dialog appears.
2. **Select a Distribution Type.**
3. **Click on OK.**

   If you chose the Cumulative option, the forecast chart appears as a cumulative distribution, as shown in Figure 4.8.

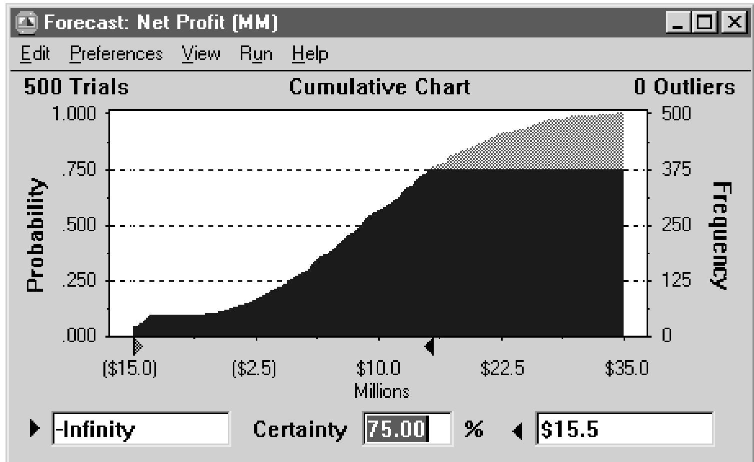


Figure 4.8 Forecast chart—cumulative distribution

Crystal Ball Note: *You can use a Crystal Ball hot key to bypass the Chart Preferences dialog. Each time you press* ***d,*** *the frequency distribution changes from a frequency distribution to a cumulative distribution to a reverse cumulative distribution and back to a frequency distribution.*

Figure 4.8 shows the Net Profit forecast chart as a cumulative distribution. To create this chart, the frequencies are added cumulatively, from one end of the range to the other, and then plotted as a cumulative frequency curve. To understand the cumulative distribution, look at a particular value, $15.5 (in the example above). The chart shows that approximately 75% of the values are less than $15.5, while approximately 25% are greater. That means there is a 75% chance for a net profit/loss below $15.5 million and a 25% chance for a net profit above $15.5 million.

Changing groups

The forecast values are always grouped internally into 300 group intervals for small charts and 450 groups for large charts. To view the data with less detail, use the Groups option to select fewer group intervals. Working with fewer group intervals might give you a better overall feel for the results.

To change the number of groups for the forecast results:

1. **In the forecast window, either:**
   - Select Preferences > Chart
   - Double-click on the chart

   The Chart Preferences dialog appears.

2. **Select a number of groups.**
3. **Click on OK.**

Using fewer details than before, Crystal Ball displays the forecast results for a frequency type distribution and a column type chart.

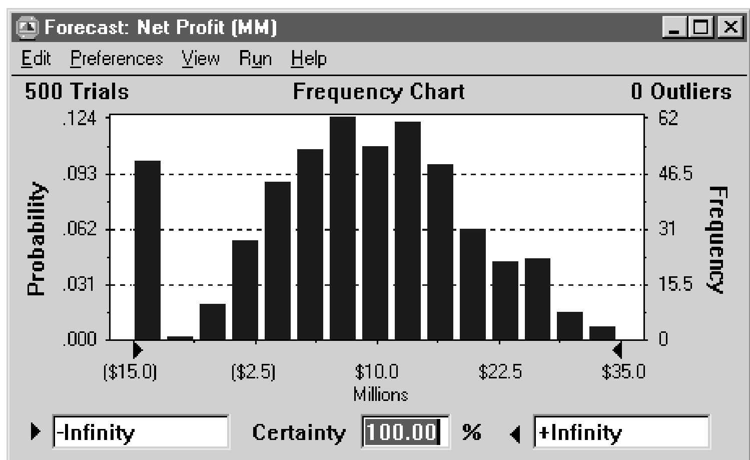


Figure 4.9 Frequency and column types

Crystal Ball Note: *Use a Crystal Ball hot key to bypass the Chart Preferences dialog. Press* ***g****. The number of groups decreases from 300 to 15, and then jumps back up to 300 (450 to 25 for large charts).*

Showing grid lines

Use the Show Grid Lines option to turn on or off horizontal grid lines on the forecast chart. The grid lines help you determine the frequency and probability for any point on the distribution.

To select horizontal grid lines on the forecast chart:

1. **In the forecast window, either:**
   - Select Preferences > Chart
   - Double-click on the chart

   The Chart Preferences dialog appears.
2. **Click the Show Grid Lines check box.**
3. **Click on OK.**

Crystal Ball Note: *Use a Crystal Ball hot key to bypass the Chart Preferences dialog. Press* ***l*** *to toggle the horizontal grid lines on and off.*

Using the mean line

Use the Mean Line check box to enable or disable a mean line on the forecast chart. The mean line helps you locate the mean value of the forecast. To turn on the mean line on the forecast chart:

1. **In the Forecast Window, select Preferences > Chart.**
2. **Click the Mean Line check box.**
3. **Click on OK.**

Crystal Ball Note: *Use a Crystal Ball hot key to bypass the Chart Preferences dialog. Press* ***m*** *to turn the Mean Line option on and off.*

Formatting forecast charts

The number format displayed on the forecast chart comes from the underlying format of the forecast cell. You can select another cell format using the Format Preferences dialog.

To change the format of a forecast:

1. **In the forecast window, select Preferences > Format.**

   The Format Preferences dialog appears, as in Figure 4.10.

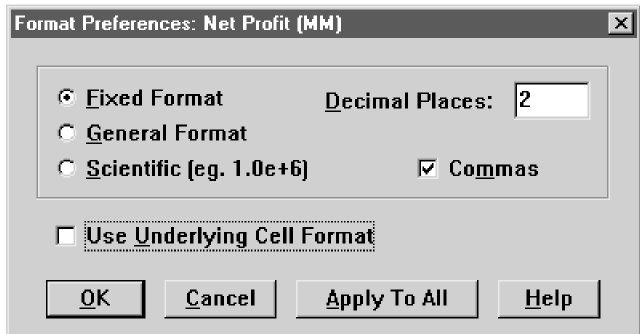


Figure 4.10 Format Preferences dialog

The Forecast Preferences are:

Use Underlying Cell Format
: Selects or deselects the underlying cell format.

Crystal Ball Note: *You must disable the Use Underlying Cell Format check box to select one of the other options.*

Fixed Format
: Displays the forecast chart values with a specified number of decimal places in a fixed decimal format.

General Format
: Displays the forecast chart values with as much precision as possible and without any formatting.

Scientific
: Displays the forecast chart values in scientific notation, for example 1.0e+6 for 1,000,000.

Decimal Places
: Changes the number of decimal places displayed for fixed and scientific formats.

Commas
: Displays the forecast chart values with commas separating thousands for fixed formats.

When you click on OK, Crystal Ball reformats the numbers on the forecast chart. This number format also appears in the statistics window and on the reports you print.

Use the Apply To All button in the Format Preferences dialog to apply an option to all forecasts.

Interpreting the statistics

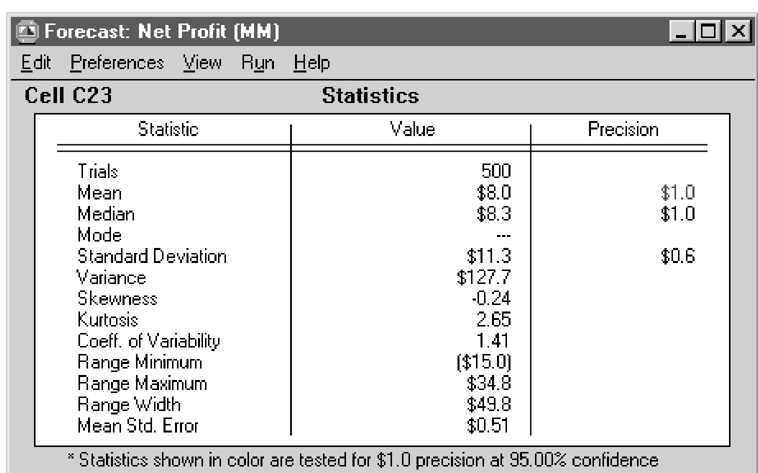
Crystal Ball provides summary descriptive statistics in a forecast window by using statistical and percentiles views.

Statistics view

You can display a full set of descriptive statistics for a simulation in the forecast window by selecting View > Statistics.

In the example at the top of the following page, statistics are shown for the entire range of values (100% of the forecast values, including the extreme values excluded from the default display range).

Statistical terms listed on this table are discussed in Chapter 2, "Understanding the Terminology," and the glossary.



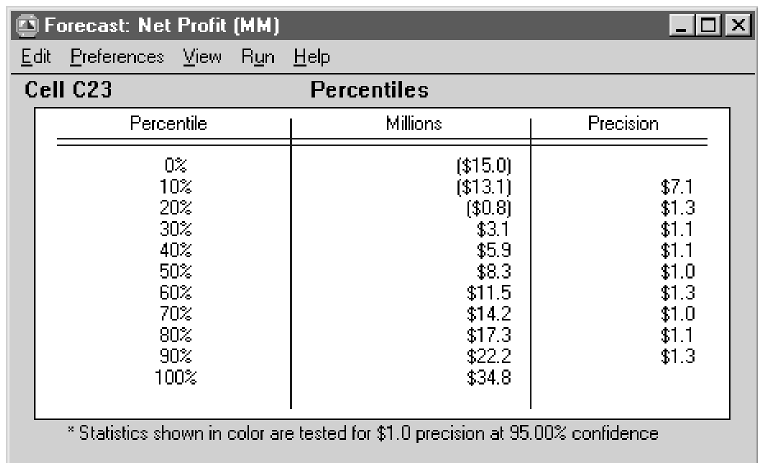
Statistic	Value	Precision
Trials	500	
Mean	$8.0	$1.0
Median	$8.3	$1.0
Mode	---	
Standard Deviation	$11.3	$0.6
Variance	$127.7	
Skewness	-0.24	
Kurtosis	2.65	
Coeff. of Variability	1.41	
Range Minimum	($15.0)	
Range Maximum	$34.8	
Range Width	$49.8	
Mean Std. Error	$0.51	

Figure 4.11 Forecast window—Statistical view

__Crystal Ball Note:__ When the Precision Control feature is not on, or if the forecast does not have any Precision Control options set, the Precision column doesn't appear in the statistics view.

Percentiles view

You can display percentile information in 10% increments in the forecast chart by selecting the View > Percentiles.



Percentile	Millions	Precision
0%	($15.0)	
10%	($13.1)	$7.1
20%	($0.8)	$1.3
30%	$3.1	$1.1
40%	$5.9	$1.1
50%	$8.3	$1.0
60%	$11.5	$1.3
70%	$14.2	$1.0
80%	$17.3	$1.1
90%	$22.2	$1.3
100%	$34.8	

* Statistics shown in color are tested for $1.0 precision at 95.00% confidence

Figure 4.12 Forecast—Percentile view

__Crystal Ball Note:__ When the Precision Control feature is not on, or if the forecast does not have any Precision Control options set, the Precision column doesn't appear in the percentile view.

Changing how statistics are calculated

Crystal Ball lets you change the way the statistics and certainty levels are computed. By default, this information is calculated based on the entire range of forecast values. However, the statistics preference dialog lets you base the calculations on the display range instead, allowing you to focus on a specific range of the frequency distribution. For example, you could determine the certainty of making at least $1,000 per month *if* the project is profitable.

To use these options:

1. **In the forecast window, select Preferences > Statistics.**

   The Statistics Preferences dialog appears, as in Figure 4.13.

2. **Change the preferences.**

3. **Click on Apply To All.**

   The changes apply to all forecast windows.

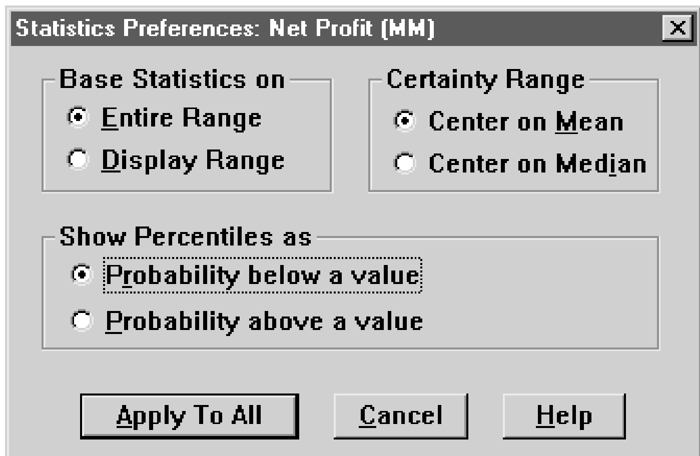


Figure 4.13 Statistics Preferences dialog

The statistics preferences are:

Entire Range	Bases the calculated statistics on all the generated forecast values, including outliers that don't appear on the forecast chart.
Display Range	Bases the calculated statistics on only the forecast values that fall in the display range. For more information on the display range, see "Focusing on the display range" on page 191.

Center On Mean

Centers the certainty range around the mean of the frequency distribution, unless one of the certainty grabbers is locked. See "Determining the certainty level" on page 187. This is the default.

This option also affects how the trend chart draws certainty bands.

Center On Median

Centers the frequency distribution on the median, unless one of the certainty grabbers is locked. See "Determining the certainty level" on page 187.

This option also affects how the certainty bands are centered on the trend chart.

Probability Below A Value

Defines percentiles as the percent chance (probability) that the associated variable value is below a particular value.

Selecting this option also affects the percentiles used for the assumption alternate parameters.

This is the default percentile setting.

Probability Above A Value

Defines percentiles as the percent chance (probability) that the associated variable value is above a particular value.

Selecting this option also affects the percentiles used for the assumption alternate parameters.

Understanding and using the overlay chart

After completing a simulation with multiple related forecasts, you can use Crystal Ball's overlay chart feature to view the relative characteristics of those forecasts on one chart. The overlay chart superimposes the frequency data from selected forecasts in one location so that you can compare differences or similarities that otherwise might not be apparent. There is no limit to the number of forecasts you can view at one time on the overlay chart.

For example, if a model had several forecasts based on slightly different interest rate assumptions, the overlay chart can be used to show how these slight differences manifest themselves in the variability of the forecasts. You can customize the overlay chart to accentuate these differences or similarities.

After the simulation is stopped, you can also use the overlay chart to fit a standard distribution to any single forecast. This process is similar to the distribution fitting feature described in Chapter 3, except that the fit is applied to forecasts, not historical data.

Crystal Ball Note: *You can only use the overlay chart after a simulation is stopped.*

Creating the overlay chart

In the Crystal Ball Examples folder there is a Reliability Engineering spreadsheet you can use to experiment with the overlay chart function. The overlay chart that you create shows the reliability of a design component when it is made from different materials and is subjected to varying stresses.

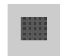
To create an overlay chart:

1. **Close any spreadsheet windows that are currently open.**
2. **Select File > Open.**
3. **Open the Reliability spreadsheet (Reliability.xls) from the Crystal Ball Examples folder.**

   The Reliability Engineering spreadsheet loads. You can scroll the spreadsheet down one page to see a detailed description of how the model works.

4. **Select Run > Run.**

   The three forecast charts for this example appear, one for each of the materials used to make the design component.



5. **After at least 500 trials, select Run > Stop.**

   Each forecast chart displays the design component's reliability as a ratio of two distributions, strength to stress. A reliable component has values greater than one, because the component's strength exceeds the stress subjected to it. An unreliable component has values less than one, because its strength is less than the stress.

   You can slide the left certainty grabber to "1.00" or type "1" in the range minimum field to view the predicted reliability of the component in the certainty statistic.

6. **Select Run > Open Overlay Chart.**

   An empty Overlay Chart dialog appears, as in Figure 4.14.

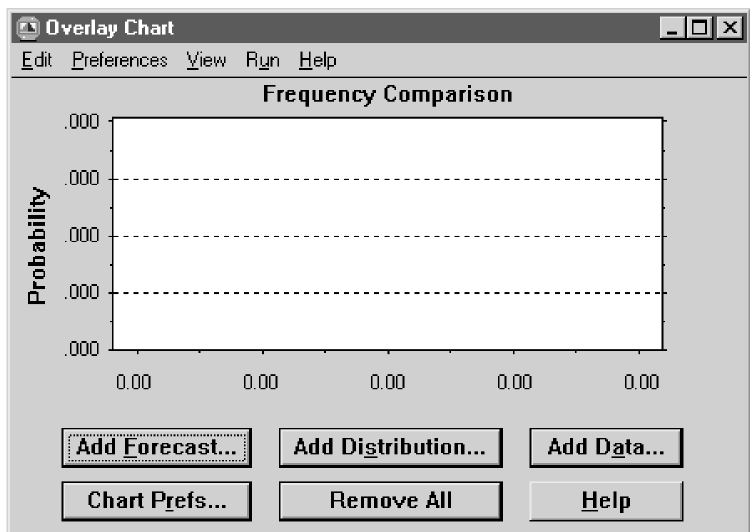


Figure 4.14 Empty overlay chart

7. **Click on Add Forecast... to select the forecasts to include in the overlay chart.**

   The Choose Forecasts dialog appears, listing all of the forecasts in the simulation.

8. **Click on Choose All to include all three forecasts in the overlay chart display.**

   The order in which you chose the forecasts determines the order in which they are displayed on the overlay chart. You can order the forecasts differently by highlighting each, then clicking on Choose Forecast.

9. **Click on OK to place the forecasts on the overlay chart.**

   The overlay chart is built and displayed with the frequency distributions for the three forecasts superimposed over each other. In this view, it is easy to see each forecast's variability and location with respect to one another.

10. **Press the "d" key until the Reverse Cumulative Comparison type appears.**

Crystal Ball Note: *Similar to steps 10 and 11, keyboard equivalents for quickly changing each of the overlay chart preferences can be found under "Overlay/Comparison chart hot keys" on page 348.*

11. **Press the "t" key until the outline chart type appears.**

    The chart type changes to show all three distributions completely, as in Figure 4.15. This chart type most clearly demonstrates that the third distribution, Material 3 Reliability, **dominates** the other distributions. We can conclude that Material 3 has the highest reliability and is superior to the other two since a greater proportion of its distribution is to the right of 1.00.

Glossary Term:
Dominant—
A relationship between distributions in which one distribution's values for all percentile levels are higher than another's.

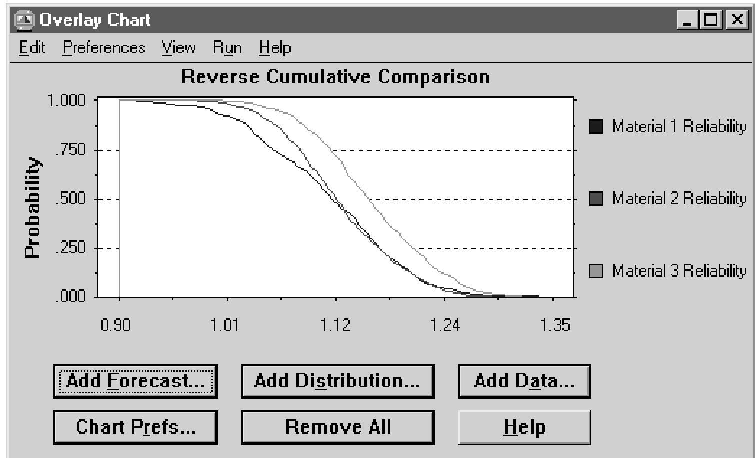


Figure 4.15 Overlay chart with three distributions

Customizing the overlay chart

You can customize the overlay chart by making selections on the Overlay Chart Preferences dialog. Customizing the overlay chart helps you compare the forecasts by viewing their differences in several ways. For example, the area and column chart types might obscure parts of some distributions behind other distributions, but the outline and line chart types show virtually all of each distribution.

To use the Overlay Chart Preferences dialog:

1. In the Overlay Chart dialog, either:

- Click on Chart Prefs
- Select Preferences > Chart

The Chart Preferences dialog appears as shown in Figure 4.16. Most of the chart preferences features are the same as those used for the forecast charts, except for those described below.

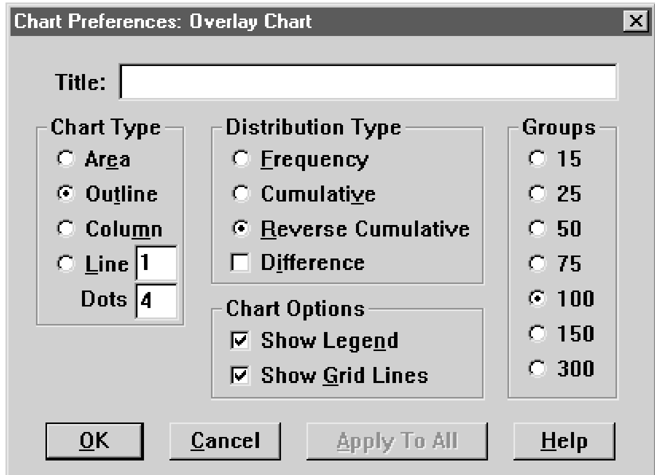


Figure 4.16 Overlay chart preferences

2. **Set your options.**
3. **To return to the overlay chart, click on OK.**

   The chart display updates with your selections.

The options for the overlay chart are:

Title	Assigns a title to the overlay chart. This title appears in the overlay chart window and on any reports that include the chart.
Chart Type	Sets how the chart is displayed.

Type	*Displays the chart as:*
Area	A filled area.
Outline	An outline of the top of the area with the forecast values.
Column	A series of bars, like a histogram.

Lines/Dots	A line connecting dots that represent a point in each interval. The line and dot values control thickness of the line and the size of the dots (in pixels). The values must be less than 16.

Distribution Type

Controls whether the distribution is a:

Frequency	Height determines the number of times the value fell within that interval. This is the default.
Cumulative	Height determines the probability that the variable falls at or below the associated value.
Reverse Cumulative	Height determines the probability that the variable falls at or above the associated variable.
Difference	Displays the difference between different items on the overlay chart. All selected forecasts are subtracted from the first forecast.

Show Legend — Displays the name of each forecast or probability distribution along the right side of the chart.

Show Grid Lines

Displays horizontal grid lines on the overlay chart.

Groups — Defines the number of intervals that Crystal Ball groups the forecast values into for display. The choices range from 15 to 300 for regular charts and 25 to 450 for maximized charts.

You can also change each distribution's chart preferences individually to view them in different formats. Press the <Ctrl> button, and then click the legend name to display the Chart Preferences dialog for only one distribution.

Using distribution fitting with the overlay chart

Similar to the distribution fitting with historical data described in Chapter 3, the overlay chart can be used to fit standard probability distributions to forecasts, in addition to comparing forecasts to each other. Much of the process of fitting distributions to forecasts is the same as that of fitting to historical data and is described beginning on page 138. The steps are included here; those that are different from "How distribution fitting works" in Chapter 3 have a detailed description.

To fit a probability distribution to one of your forecasts:

1. **Run the example ending on page 208, if you haven't already done so.**
2. **In the Overlay Chart dialog, click on Add Distribution.**

   The Distribution Gallery appears. If you just want to fit a particular probability distribution to your forecast, select the distribution from the gallery and click on OK. However, in this example you will choose all of the distributions.
3. **Click on Fit to begin the distribution fitting process.**

   The first of two Fit Distribution dialogs appears, which is used to select the data to which the distribution is fitted. You can use one of the overlay forecasts or historical data from a file; this example uses an overlay forecast.

   These dialogs are discussed in detail in "Fitting distributions to data" beginning on page 138.
4. **Click the radio button for Overlay: Material 1 Reliability to display the forecasts for selection.**

   The Choose Overlay dialog appears, as in Figure 4.17.

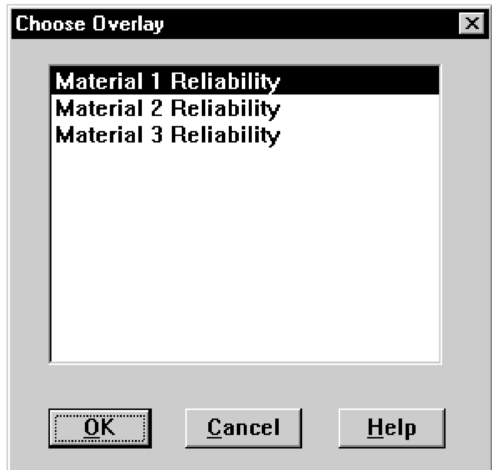


Figure 4.17 Choose Overlay dialog

5. **Click on Material 3 Reliability to select the forecast for fitting.**

   Crystal Ball fits probability distributions to the Material 3 forecast.

6. **Click on OK to return to the first Fit Distribution dialog.**

7. **Click on Next to select the distributions to be fitted and the ranking method.**

   The Fit Distribution (2 of 2) dialog appears. For our example, choose All Continuous Distributions and the Chi-square Test ranking method.

8. **Click the Show Comparison Chart And Goodness-of-Fit Statistics option, if it is not already checked.**

9. **Click on OK to fit the distributions and to display the comparison chart.**

   Crystal Ball fits the distributions, and then displays the selected forecast with each fitted probability distribution.

10. **Optionally, click on Next/Prev. Distribution buttons to scroll through the fitted distributions.**

11. **Optionally, click on Prefs to customize the display.**

12. Click on Accept.

The Distribution dialog appears.

13. Click on OK.

Crystal Ball places the currently displayed distribution on the overlay chart.

At this point, the overlay chart displays with a fitted distribution superimposed over the Material 3 Reliability forecast on the overlay chart, similar to Figure 4.18. Use this display to view the forecasts' shapes and characteristics, as in the example below.

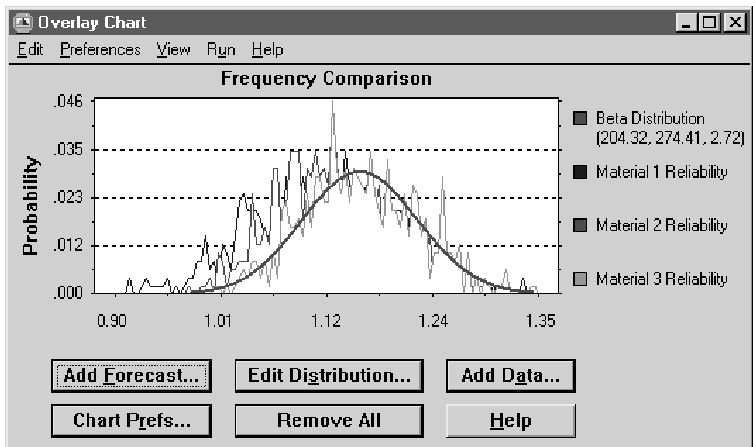


Figure 4.18 Overlay chart with fitted distribution

Other overlay chart features

In addition to the features described in this section, several more are available.

Edit Distribution
Appears after a you add a distribution to the overlay chart. Use this button to change or remove the overlaid distribution.

Add Data
Adds data from a text file to the overlay chart. The text file must contain only numbers separated by tabs and/or line breaks. Any numbers proceeded by a special character (such as $) are ignored.

Remove All	Clears the overlay chart. You can then re-order the forecasts, rebuild the overlay chart, and fit the probability distributions again.

Understanding and using the trend chart

After completing a simulation with multiple related forecasts, you can create a trend chart to view the certainty ranges of all the forecasts on a single chart. A trend chart summarizes and displays information from multiple forecasts, making it easy to discover and analyze trends that might exist between related forecasts. You can customize your trend chart to display the probability that given forecasts will fall in a particular part of a value range.

For example, if a model contains forecasts related through time, the trend chart could be used to view the certainty ranges for each forecast side by side. The certainty ranges for an early time period and a later time period could be compared with each other at a glance.

Crystal Ball Note: *The trend chart is only available after a simulation is completed and is only meaningful when you have multiple forecasts that are related to each other.*

Understanding the trend chart

The trend chart displays certainty ranges for multiple forecasts in a series of patterned bands. Each band represents the certainty ranges into which the actual values of your forecasts fall. For example, the band which represents the 90% certainty range shows the range of values into which your forecast has a 90% chance of falling. By default, the bands are centered around the mean of each forecast.

Use the Trend Chart Preferences dialog to change the location of the bands so that they are anchored at either the high end or the low end of the projected forecast ranges. Changing the location of the bands is useful when analyzing cumulative or reverse cumulative probabilities (unless your certainty range is set to center on the median).

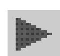
Creating the trend chart

In the Crystal Ball Examples folder there is a Sales Projection spreadsheet you can use to experiment with the trend chart function. The trend chart that you create displays the probability that a manufacturer's gross sales will fall within a particular range of values over a three-year period.

To create a trend chart:

1. **Close any spreadsheet windows that are currently open.**
2. **Select File > Open.**
3. **Open the Sales Projection.xls spreadsheet from the examples folder.**

   The Quarterly Sales Projection spreadsheet appears.



4. **Select Run > Run.**

   One of the forecast charts for the manufacturer's spreadsheet appears (there are 12 forecasts for this spreadsheet). As the simulation proceeds, this chart changes to reflect the changing values in the forecast cells. Run about 200 trials before stopping the simulation.

5. **Select Run > Stop.**

6. **Select Run > Open Trend Chart.**

   The trend chart appears, as in Figure 4.19.

The trend chart displays the certainty ranges on a quarterly basis over a three year period. Because the model contains quarterly forecast formulas dependent on the previous quarter's results, the bands widen in the future. This occurs because the standard deviation of the forecasts increases or widens for each quarter. Trend charts like the one illustrated in Figure 4.19 demonstrate the compounding of uncertainty that occurs as predictions are made farther and farther into the future.

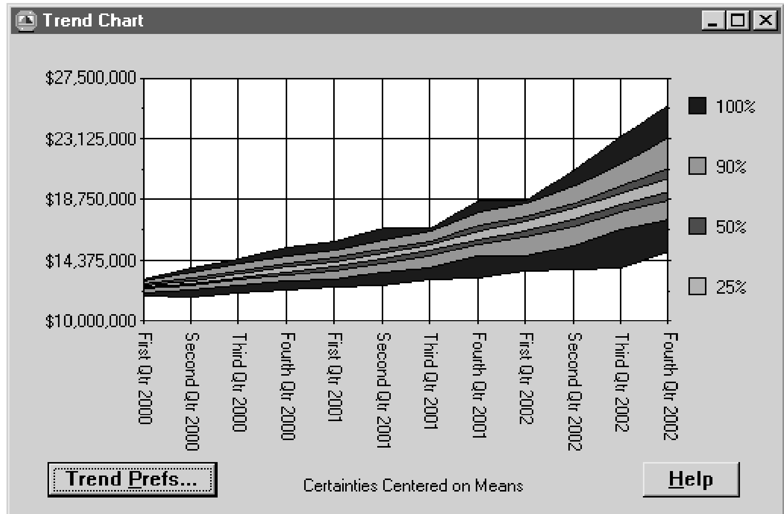


Figure 4.19 Trend Chart dialog

Customizing the trend chart

Use the Trend Preferences dialog to customize the trend chart. Changing the appearance of the trend chart helps you interpret the results of the simulation by allowing the data to be viewed in different ways.

To use the Trend Preferences dialog:

1. **At the bottom of the Trend Chart dialog, click on Trend Prefs.**

   The Trend Preferences dialog appears, as in Figure 4.20.

   The Trend Preferences dialog provides a variety of options for changing the appearance of the trend chart and selecting specific forecasts for display. The appearance of the trend chart can be changed by selecting one or more options at a time.

   For some of the options, you can use hot keys to bypass the Trend Preferences dialog. These hot keys are described in the following sections and are listed in Appendix C, "Keyboard Commands."

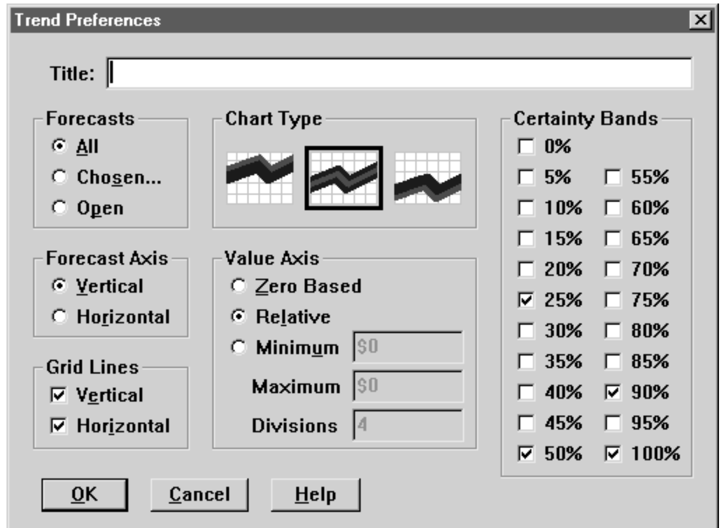


Figure 4.20 Trend Preferences dialog

Choosing forecasts

Use the Forecasts option to select the forecasts to include on the trend chart. You can select all forecasts, selected forecasts, or just those forecast charts that are visible.

To select the forecasts to display:

1. **In the Trend Chart, click on Trend Prefs.**
2. **Either:**
   - To include all the forecasts, click on All.
   - To include only the forecasts that have open forecast windows, click on Open.
   - To include only selected forecasts, click on Chosen.

     A selection dialog appears, letting you choose from a list of available forecasts.
3. **Click on OK.**

   The new trend chart appears.

Changing the certainty bands

Use the Certainty Bands option to define the certainty levels for the trend chart in 5-percent increments from 0% to 100%.

To change the certainty bands:

1. **In the Trend Chart, click on Trend Prefs.**
2. **Select the Certainty Bands check boxes for the certainty levels to display on the trend chart.**

   The chart can display up to seven bands at one time.
3. **Click on OK.**

   The trend chart appears with the selected certainty bands. A legend appears on the right side of the trend chart indicating which percentages are represented by which bands.

Changing the placement of the certainty bands

Use the Chart Type option to change the placement of the percentile bands within the trend chart. The default setting centers the bands around the mean or median of each forecast (depending on the Forecast Window statistics preference setting—see “Changing how statistics are calculated” on page 203). You can change the location of the bands so that they are anchored at either the high end or the low end of the projected forecast ranges.

Crystal Ball Note: *Smaller bands always appear on top of larger bands. This obscures the larger bands. Do not confuse the actual width of a band with the portion that is visible. You can display the true size of a band using the Certainty Bands option in the Trend Preferences dialog. This option lets you display the bands one at a time on the trend chart.*

To change the placement of the certainty bands:

1. **In the Trend Chart, click on Trend Prefs.**
2. **Either:**

   - To display the forecasts anchored at the higher end of the forecast range, click the left Chart Type icon.

     This plot shows the certainty that the forecast values will be above a given value (reverse cumulative probability).

- To display the forecasts centered around the mean of each forecast value, click the middle Chart Type option icon.

  This plot shows the certainty that the forecast values will fall on either side of the 50th percentile. Use the Statistics Preference to center the bands around the means of each forecast instead.

- To display the forecasts anchored at the lower end of the forecast range, click the right Chart Type option icon.

  This plot shows the certainty that the forecast values will be below a given value (cumulative probability).

3. Click on OK.

The trend chart appears showing the new placement of the percentile chart types. Figure 4.21 shows the trend chart with bands organized around the lower end.

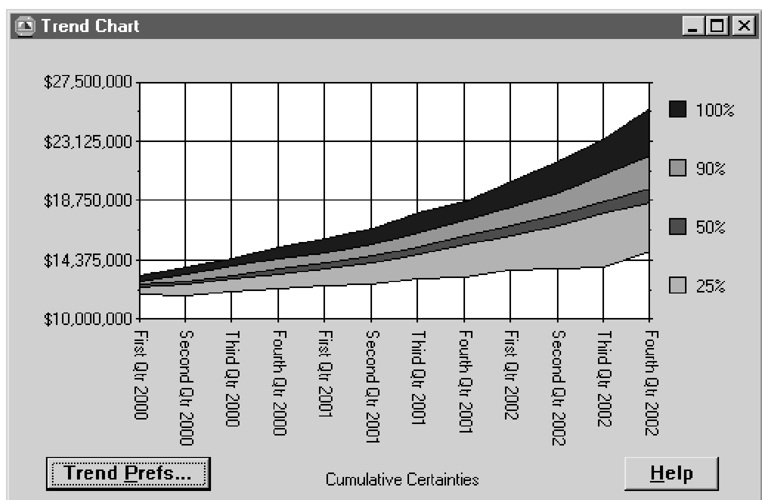


Figure 4.21 Trend chart bands

Crystal Ball Note: *Use a Crystal Ball hot key to bypass the Trend Preferences dialog. Press* ***d.*** *The placement of the bands changes from centered to lower to higher and back to centered.*

Changing the forecast axis

Use the Forecast Axis option to display the forecast names vertically or horizontally at the bottom of the trend chart. You can display or hide individual names by clicking on the names in the forecast window. To change the forecast axis:

1. **In the Trend Chart, click on Trend Prefs.**
2. **Either:**
   - To display the forecast names vertically, click on Vertical.
   - To display the forecast names horizontally, click on Horizontal.
3. **Click on OK.**

   The trend chart appears with the forecast axis changed.

Crystal Ball Note: *Use a Crystal Ball hot key to bypass the Trend Preferences dialog. Press* ***f*** *to toggle the forecast names between horizontal and vertical formats.*

Changing the value axis

Use the Value Axis option to change the minimum or maximum values used for the value axis endpoints. You can let Crystal Ball define one or both endpoints by using the minimum and maximum values from all selected forecasts or you can enter your own values. For the latter option, you also can change the number of divisions displayed on the axis.

Crystal Ball Note: *The number format for the axis values is taken from the first forecast that appears on the trend chart.*

To change the value axis:

1. **In the Trend Chart, click on Trend Prefs.**

2. **Either:**

   - To let Crystal Ball determine the minimum and maximum values to use for the axis endpoints, select Relative.
   - To use zero for the bottom endpoint instead of the minimum value, select Zero Based.
   - To use actual values for the minimum or maximum endpoints or specify the number of divisions on the axis chart, select Minimum and then enter the values in the Minimum, Maximum, and Divisions fields.

   By changing the minimum or maximum endpoint values, you can zoom in or out on selected ranges of the trend chart.

3. **Click on OK.**

   The trend chart appears with the new endpoints and the new axis divisions.

Crystal Ball Note: *Use a Crystal Ball hot key to bypass the Trend Preferences dialog. Press* ***v****. The value axis endpoints change from relative to minimum to maximum to zero based and back to relative.*

Changing the grid lines

Use the Trend Preferences Grid Lines option to display vertical and/or horizontal grid lines on the trend chart.

To change the grid lines:

1. **In the Trend Chart, click on Trend Prefs.**
2. **To display vertical grid lines, select the Vertical check box.**
3. **To display horizontal grid lines, select the Horizontal check box.**
4. **Click on OK.**

   The new grid lines appear on the trend chart.

Crystal Ball Note: *Use a Crystal Ball hot key to bypass the Trend Preferences dialog. Press* ***l****. The grid lines change from horizontal to vertical to horizontal and vertical to no grid lines and back to horizontal.*

Adding a title

Use this option to add or change a title on the chart.

To add or change a title:

1. **In the Trend Chart, click on Trend Prefs.**
2. **Type the title in the Title field.**
3. **Click on OK.**

   The trend chart appears with the new title.

Understanding sensitivity and using the sensitivity chart

Glossary Term:
sensitivity—
The amount of uncertainty in a forecast cell that is caused by both the uncertainty and model sensitivity of an assumption cell.

As you become more proficient at building spreadsheet models, you will want to know how much a given assumption affects your result. In other words, you want to determine the **sensitivity** of the forecast to each assumption. The overall sensitivity of a forecast to an assumption is a combination of two factors:

- The **model sensitivity** of the forecast to the assumption
- The assumption's uncertainty

To determine the model sensitivity yourself, you must algebraically analyze the various relationships between your forecast cell and your assumption cells. These relationships include all of the formulas in the spreadsheet that link the assumption cells to the forecast cells.

For example, you have two assumptions: Assum1 and Assum2. You also have one forecast, Forecast1, defined as:

Forecast1 = Assum1*0.6 + Assum2*0.3

The model sensitivity in this case indicates that Forecast1 is twice as sensitive to changes in the value of Assum1 than Assum2. In other words, if Assum1 increased by 5, Forecast1 would increase by twice as much as if you increased Assum2 by 5 instead.

Glossary Term:
model sensitivity—
The overall effect that a change in an assumption cell produces in a forecast cell. This effect is solely determined by the formulas in the spreadsheet model.

You then have to calculate how the uncertainty of your assumption cells affect the forecast cell to compute a final answer. The uncertainty of an assumption depends on the type of distribution and the parameters of that distribution.

If both assumptions change by the same amount, such as increasing by 5, as in the above example, then the uncertainty of the assumption would not change the overall sensitivity of the forecast to the two assumptions. However, if they change by different amounts, you must take into account both the model sensitivity and the uncertainty. At some point, the uncertainty for Assum2 might overcome the model sensitivity and result in an overall sensitivity greater than that of Assum1.

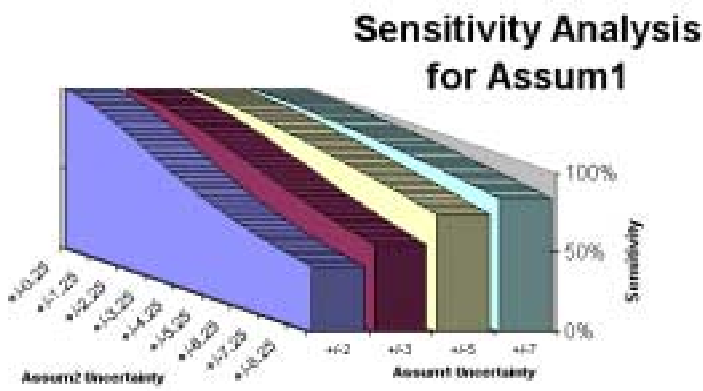


Figure 4.22 Changing sensitivity with uncertainty only

Figure 4.22 shows the overall sensitivity of Forecast1 to Assum1 and Assum2. When the uncertainty of Assum1 is high, the sensitivity is above 50%. This means that Assum1 has more influence over Forecast1 than Assum2. However, when the uncertainty of Assum1 is low, a high uncertainty in Assum2 can overcome the higher model sensitivity of Assum1 and have a greater influence over Forecast1 than Assum1.

Calculating this sensitivity could be a difficult and time-consuming task without a program like Crystal Ball.

Understanding the sensitivity chart

The sensitivity chart feature provides you with the ability to quickly and easily judge the influence each assumption cell has on a particular forecast cell. During a simulation, Crystal Ball ranks the assumptions according to their importance to each forecast cell. The sensitivity chart displays these rankings as a bar chart, indicating which assumptions are the most important or least important in the model. You can output (print) the sensitivity chart on the report or copy it to the clipboard.

The sensitivity chart feature provides three key benefits:

- You can find out which assumptions are influencing your forecasts the most, reducing the amount of time needed to refine estimates.
- You can find out which assumptions are influencing your forecasts the least, so that they can be ignored or discarded altogether.
- As a result, you can construct more realistic spreadsheet models and greatly increase the accuracy of your results because you will know how all of your assumptions affect your model.

Creating the sensitivity chart

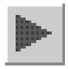
In the Crystal Ball Examples folder, there is a Toxic Waste Site spreadsheet you can use to experiment with the sensitivity chart feature. The sensitivity chart created in the steps below displays, in descending order, the assumptions in a risk assessment of a toxic waste site. The assumption with the highest level of sensitivity can be considered as the most influential assumption in the model.

To create a sensitivity chart:

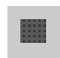
1. **Close any spreadsheet windows that are currently open.**
2. **Select File > Open.**
3. **Open the Toxic Waste Site spreadsheet (Toxic Waste Site.xls).**

4. **Select Run > Run Preferences> Options.**
5. **Make sure the Calculate Sensitivity option is checked.**

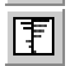
6. **Click on OK.**



7. **Run a simulation.**



8. **Stop the simulation.**



9. **Select Run > Open Sensitivity Chart.**

   The Sensitivity Chart dialog appears, displaying the sensitivity rankings of the assumptions in your simulation, as in Figure 4.23.

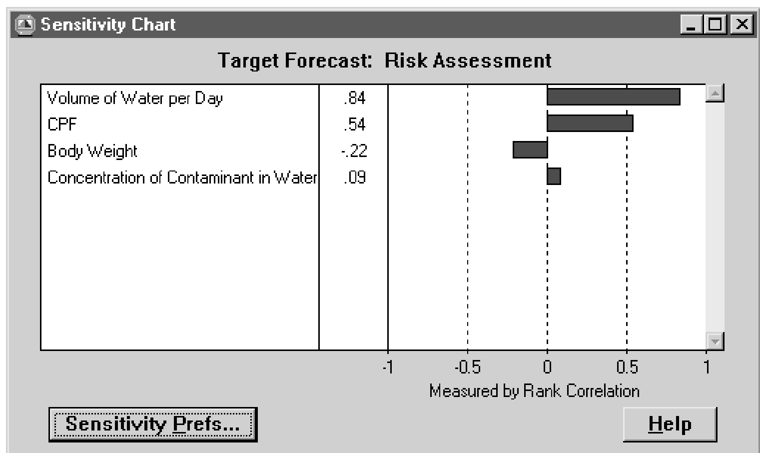


Figure 4.23 Sensitivity Chart dialog

If you select Open Sensitivity Chart but the Calculate Sensitivity option was not checked in the Run Preferences dialog, reset the simulation and run it again.

The assumptions (and possibly other forecasts) are listed on the left side, starting with the assumption with the highest sensitivity. The default color values are that assumptions appear as green bars and forecasts appear as blue bars, unless you change the color settings. Use the scroll bar to view the entire bar chart.

In this example, there are four assumptions listed in the sensitivity chart. The first assumption, Volume Of Water Per Day, has the highest sensitivity ranking and can be considered the most important assumption in the model. A researcher running this model would want to investigate this assumption further in the hopes of reducing its uncertainty, and therefore its effect on the target forecast.

The last assumption, Concentration Of Contaminant In Water, has the lowest sensitivity ranking and is the least important assumption in the model. The effect of this assumption on the target forecast is not as great as the others and, in this particular case, could be ignored or altogether eliminated by clearing it from the spreadsheet. Sensitivity charts like this one illustrate that one or two assumptions typically have a dominant effect on the uncertainty of a forecast.

How Crystal Ball calculates sensitivity

Glossary Term:
rank correlation—
A method whereby Crystal Ball replaces assumption values with their ranking from lowest value to highest value using the integers 1 to N prior to computing the correlation coefficient.

Crystal Ball calculates sensitivity by computing **rank correlation** coefficients between every assumption and every forecast while the simulation is running. Correlation coefficients provide a meaningful measure of the degree to which assumptions and forecasts *change together*. If an assumption and a forecast have a high correlation coefficient, it means that the assumption has a significant impact on the forecast (both through its uncertainty and its model sensitivity). Positive coefficients indicate that an increase in the assumption is associated with an increase in the forecast. Negative coefficients imply the reverse situation. The larger the absolute value of the correlation coefficient, the stronger the relationship.

Crystal Ball also computes the correlation coefficients between all pairs of forecasts while the simulation is running. You might find this sensitivity information useful if your model contains several intermediate forecasts that feed into a final forecast.

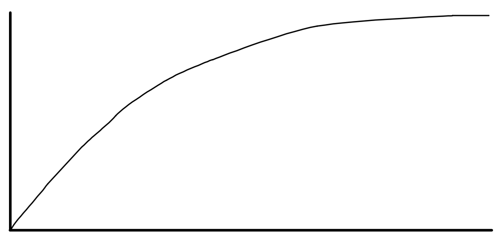
An option in the Sensitivity Preference dialog lets you display the sensitivities as a percentage of the contribution to the variance of the target forecast. This option, called Contribution To Variance, doesn't change the order of the items listed in the sensitivity chart and makes it easier to answer questions such as "what percentage of the variance or uncertainty in the target forecast is due to assumption X?". However, it is important to note that this method is only an approximation and is *not precisely* a variance decomposition. Crystal Ball calculates Contribution To Variance by squaring the rank correlation coefficients and normalizing them to 100%.

Caveats

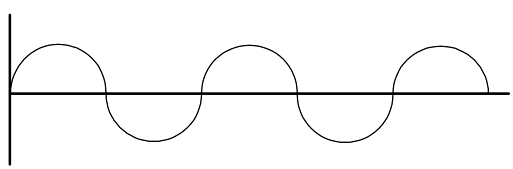
The sensitivity chart feature has several limitations you should be aware of:

- **Correlated assumptions.** The sensitivity calculation might be inaccurate for correlated assumptions. For example, if an important assumption were highly correlated with an unimportant one, the unimportant assumption would likely have a high sensitivity with respect to the target forecast. Assumptions that are correlated are flagged as such on the sensitivity chart. In some circumstances, turning off correlations in the Run Preference dialog might help you to gain more accurate sensitivity information.
- **Non-monotonic relationships.** The sensitivity calculation might be inaccurate for assumptions whose relationships with the target forecast are not monotonic. A monotonic relationship means that an increase in the assumption tends to be accompanied by a strict increase in the forecast; or an increase in the assumption tends to be accompanied by a strict decrease in the forecast.

  For example, the relationship y = Log(x) is monotonic:

  While y = Sin(x) is not:

  The Tornado Chart tool can help you discover if any of your assumptions have non-monotonic relationships with the target forecast. For more information, see “Tornado Chart tool” on page 289.

Customizing the sensitivity chart

Use the Sensitivity Prefs dialog to customize the sensitivity chart. As you become more familiar with the sensitivity chart, practice selecting preferences that help you get the answers you seek and that are appropriate for your data.

1. Click on Sensitivity Prefs to open the Sensitivity Preferences dialog, as in Figure 4.24.

The sensitivity options are:

Target Forecast
: Lets you select which forecast cell is the target of the sensitivity analysis.

Measure By
: Defines whether Crystal Ball shows the sensitivities as rank correlations or contributions to variance. Rank correlations range from -1 to +1 and indicate both magnitude and direction. Contributions to variance range from 0% to 100% and indicate relative importance.

Include
: Selects which types of Crystal Ball data to include in the rank against the target forecast: assumptions, forecasts, or both (by selecting both Assumptions and Other Forecasts).

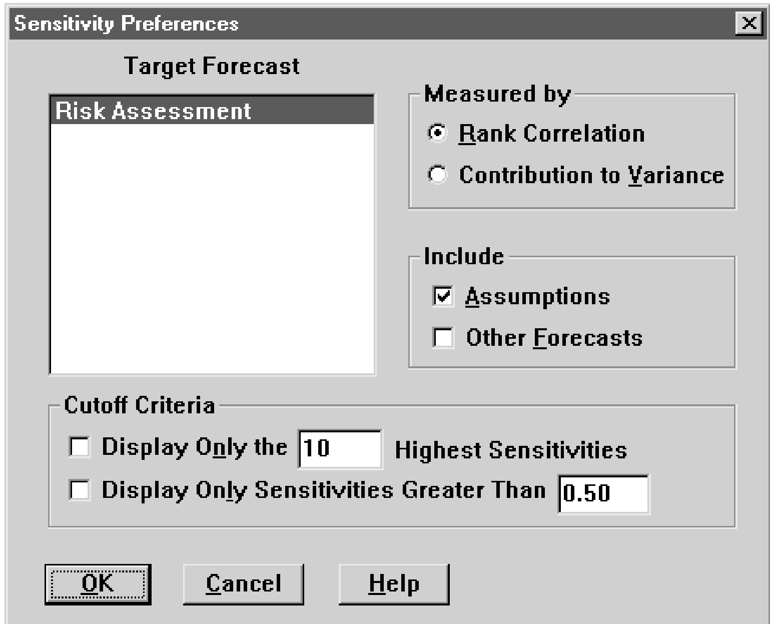


Figure 4.24 Sensitivity Preferences dialog

Cutoff Criteria
: Limits which sensitivities the sensitivity chart displays. You can limit them in two ways:

Display Only The Highest Sensitivities
: Only displays a certain number of sensitivities, starting with the highest, depending on the number (integer) you enter in the field.

Display Only Sensitivities Greater Than
: Only displays the sensitivities over the indicated number in the field. The number should be between 0 and 1.

Crystal Ball Note: *Crystal Ball always includes all the assumptions and forecasts in your model even though they might be unrelated. Generally, this is not an issue since unrelated assumptions and forecasts have sensitivity rankings close to zero. However, correlation might affect sensitivity analysis if there is a strong correlation between the variables and at least one of the variables is highly sensitive.*

Changing chart patterns and colors

Use the pattern and color dialog to select different patterns and, if you have a color monitor, different colors to enhance the appearance of the charts used in Crystal Ball. To change a chart's patterns and colors:

1. Display the Pattern And Color dialog:

- For the overlay, trend, or sensitivity charts, click inside the chart area or band, or click one of the names in the legend to the right of the chart.
- For an assumption or forecast distribution, press and hold <Ctrl>, and then click in the distribution area.

The Pattern And Color options appear, as in Figure 4.25.

Crystal Ball Note: *The pattern of the correlation chart is fixed. However, you can change the color of the correlation chart by pressing and holding <Shift>, then clicking in the correlation chart.*

2. Click on the desired pattern.

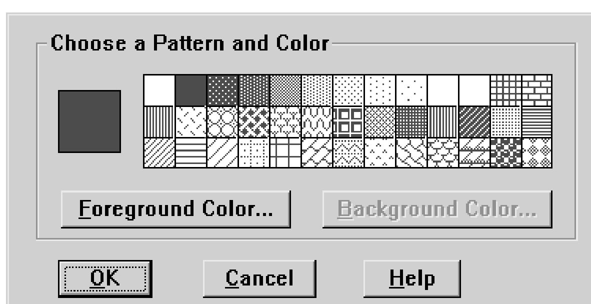


Figure 4.25 Pattern And Color dialog

3. **To display the color choices for the foreground or background patterns, click on Foreground Color or Background Color.**

    a. **Click on the color of your choice.**

    b. **Click on OK.**

4. **Click on OK.**

Creating reports

You can create a report for your simulation using the Report command. Any or all of the following items can be included in the report using the Report dialog:

- Overlay charts
- Trend charts
- Sensitivity charts
- Forecast summaries
- Forecast statistics
- Forecast charts
- Forecast percentiles
- Forecast frequency counts
- Assumption parameters
- Assumption charts
- Decision variables

Crystal Ball Note: *Reports are only available after you run a simulation.*

To create a report:

1. **Select Run > Create Report.**

The Create Report dialog appears, as in Figure 4.26.

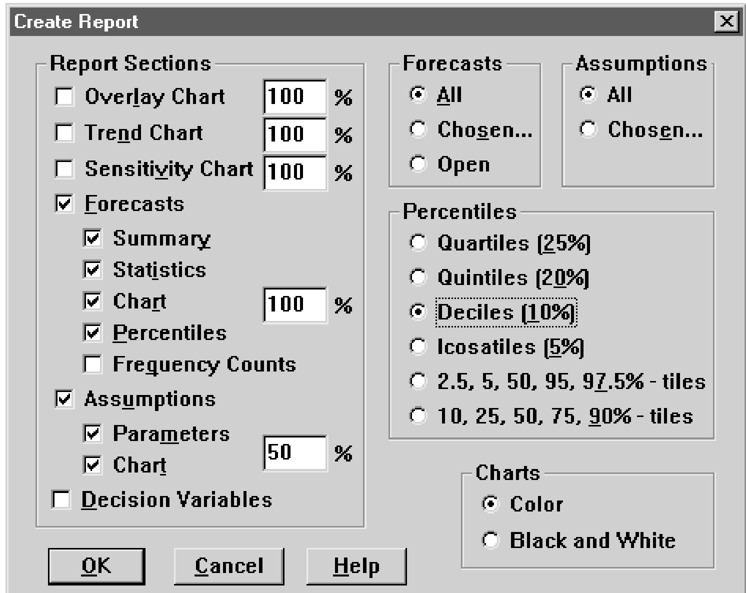


Figure 4.26 Create Report dialog

2. **Select your options and what to include in the report.**

   The descriptions of the selections are in the following sections.

3. **Click on OK.**

   Crystal Ball creates the report as an Excel spreadsheet. You can modify, print, or save the report in the same way as any other spreadsheet. For example, you can select the File > Print option for your spreadsheet model as you would for a normal spreadsheet.

Crystal Ball Note: *To suppress the initial report header and any cell references that occur in the report, hold down the <Shift> key when selecting Run > Report.*

Report sections

Select the different report sections you want to include in the report. The Report Sections options are:

Overlay Chart
: Includes the overlay chart on the report. You can also set the scale of the chart on the report by entering a percentage in the field.

  This is off by default.

Trend Chart
: Includes the trend chart on the report. You can also set the scale of the chart on the report by entering a percentage in the field.

  This is off by default.

Sensitivity Chart
: Includes the sensitivity chart on the report. You can also set the scale of the chart on the report by entering a percentage in the field.

  This is off by default.

Forecasts
: Includes various types of forecast information, including the name, cell, and other options. All these options are on by default, except for Frequency Counts.

  For more information on which forecasts to include, "Selecting forecasts" on page 235.

  Summary
  : Includes the endpoints of the entire range and the display range, as well as the mean standard error of the final result.

  Statistics
  : Lists all the statistics and their values for the forecast.

  Chart
  : Includes the forecast chart in the report. You can also set the scale by entering a percentage in the field.

  Percentiles
  : Lists percentiles for the forecast, as set by the Percentiles option in the right half of the dialog.

Frequency Counts
: Includes the start point and endpoint for each frequency range, plus the number of trials that fell in that range and the probability of falling in that range.

Assumptions
: Includes the name, cell, and two other types of information (listed below) for each assumption, as selected by the Assumptions option on the right side of the dialog. Both these options are on by default.

For more information on which assumptions to include, see "Selecting assumptions" on page 236.

Parameters	Lists the parameters used to define the distribution.
Chart	Includes the distribution chart in the report. You can also set the scale by entering a percentage in the field.

Decision Variables
: Lists the name, cell, lower and upper bound, and step (if discrete) for each decision variable.

Selecting forecasts

If you checked the Forecasts option in Report Sections, select the forecasts to include in the report by choosing one of three options:

All	Includes all forecasts in the report.
Chosen	Opens a Choose Forecasts dialog that lets you select which forecasts to include in the report and in what order. Only the selected forecasts then appear in the report, in the order that you added them to the Chosen Forecast list.

Open	Includes only the currently open forecasts in the report.

Selecting assumptions

If you checked the Assumptions option in Reports Sections, select the Assumptions to include in the report by choosing one of two options:

All	Includes all assumptions in the report.
Chosen	Opens a Choose Assumptions dialog that lets you select which assumptions to include in the report and in what order. Only the selected assumptions then appear in the report, in the order that you added them to the Chosen Assumptions list.

Percentiles

The Percentiles options for the forecast report show the certainty of achieving a value below a particular threshold. For example, if a report was printed for the Net Profit forecast discussed in the Vision Research tutorial in Chapter 1, the Percentiles section might show that you could be 80% certain of achieving a net profit or loss below the threshold value of $17.3 million.

Crystal Ball Note: *You can reverse the meaning of the percentiles by changing the setting in the Statistics Preferences dialog. For more information, see "Changing how statistics are calculated" on page 203.*

If you checked the Percentiles option in the Report Sections, select the appropriate button to choose which percentiles you want to display. The six pre-defined Percentiles options divide the distribution into value levels of different sections:

- Quartiles. Every 25th percentile for four sections: 25%, 50%, 75%, and 100%.
- Quintiles. Every 20th percentile for five sections.
- Deciles. Every 10th percentile for 10 sections.
- Icosatiles. Every fifth percentile for 20 sections.

- 2.5%, 5%, 50%, 95%, and 97.5%
- 10%, 25%, 50%, 75%, and 90%

Charts

When including either forecast or assumption charts in your report, you can display the charts as color or black and white. This can improve the results on certain printers and monitors.

The two charts options are:

Color — Copies the color version of the forecast and assumption charts to the report.

Black And White — Copies black and white versions of the forecast and assumption charts to the report.

Extracting data

Crystal Ball lets you extract forecast information generated by the simulation. Crystal Ball places the extracted data in a new workbook.

Crystal Ball Note: *To extract data for an assumption, define the assumption as a forecast.*

Crystal Ball Note: *You can only extract data after you run a simulation.*

To extract data:

1. **Select Run > Extract Data.**

   An Extract Data dialog appears, as in Figure 4.27.

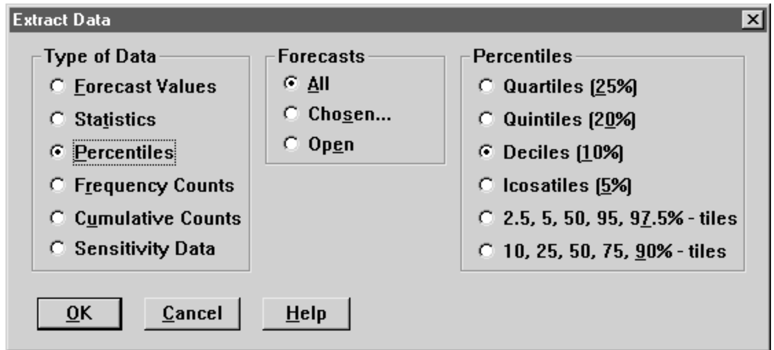


Figure 4.27 Extract Data dialog

2. **To select the type of data to extract, click the appropriate option in the Type Of Data options.**

These options are:

Option	***Extracts:***
Forecast Values	The forecast results for each trial. For more information on which forecasts to include, see "Selecting forecasts" on page 235.
Statistics	Descriptive statistics summarizing the forecast values.
Percentiles	The certainty of achieving values below a particular threshold in the selected increments for the forecast.
Frequency Counts	For each group interval: the interval range as well as the probability and frequency of occurrences within the interval for the forecast.
Cumulative Counts	Same as Frequency Counts, except each group interval is shown with cumulative probability and frequency information.
Sensitivity Data	Sensitivities for all pairs of assumptions and forecasts.

3. **In the Forecasts area, either:**

    - To include the selected data for all forecasts, select All.
    - To include the selected data for only selected forecasts, and to specify the order in which they will appear in the data sheet, select Chosen.

      The Choose Forecasts dialog appears letting you choose from list of available forecasts.
    - To include the selected data for only currently open forecasts, select Open.

4. **If you selected to extract percentiles, select which percentiles to include in the report.**

    The Percentiles options for the forecast report show the certainty of achieving a value below a particular threshold.

Crystal Ball Note: *You can reverse the meaning of the percentiles by changing the setting in the Statistics Preferences dialog. For more information, see* *"Changing how statistics are calculated" on page 203*.

    The six Percentiles options that divide the distribution into value levels of different sections are:

    - Quartiles. Every 25th percentile for four sections: 0%, 25%, 50%, 75%, and 100%.
    - Quintiles. Every 20th percentile for five sections.
    - Deciles. Every 10th percentile for 10 sections.
    - Icosatiles. Every fifth percentile for 20 sections.
    - 2.5%, 5%, 50%, 95%, and 97.5%
    - 10%, 25%, 50%, 75%, and 90%

5. **Click on OK.**

    Crystal Ball extracts the simulation data as an Excel spreadsheet. The spreadsheet is arranged as columns of forecasts and rows of data. You can sort, modify, print, or save the data in the same way as any other spreadsheet.

Chapter 5

Maximizing Your Use of Crystal Ball

In this chapter

All Crystal Ball users

- Simulation accuracy
- Memory use
- Simulation speed
- Correlated assumptions
- Sensitivity analysis

Excel users

- Customizing reports and data layouts

This chapter contains information that helps you get the most out of Crystal Ball, whether by using different Crystal Ball features or by the way you build your models.

Getting the most out of Crystal Ball

This chapter contains information that you can use to improve the overall performance of Crystal Ball. These improvements occur in terms of the accuracy of your model or speed of the results.

This chapter contains two primary sections, including:

- Crystal Ball tips
- Excel tips

Crystal Ball tips

This section contains ways to improve Crystal Ball's performance for all users.

Simulation accuracy

The accuracy of your simulation may be controlled in three ways:

- Using Precision Control
- Changing the sample size
- Changing the sampling method

Precision Control

The Precision Control feature lets you set how precise you want one of three forecast statistics to be. Crystal Ball runs the simulation until those statistics reach the required precision as determined by calculating confidence intervals.

Crystal Ball Note: *See "Confidence intervals" on page 115 for more information about how Crystal Ball calculates confidence intervals.*

Generally speaking, as more trials are calculated, the confidence interval narrows and the statistics become more accurate. The precision control feature in Crystal Ball uses this characteristic of confidence intervals to determine when a specified accuracy of a statistic has been reached. It compares the specified precision to the confidence interval. When the calculated confidence interval drops to less than the specified precision, the simulation stops.

For each forecast, you can specify precision in either absolute terms in units of the forecast, or in relative terms as percentages. Each method has its own benefits and drawbacks.

Specifying precision in absolute terms can give you greater control of the simulation when the shape and scale of the forecast distribution is roughly known. For example, for a Net Present Value (NPV) forecast that ranges from $30,000 to $65,000 dollars you can require the precision of the mean to be within plus or minus $1,000, or some other convenient measure of accuracy. However, if the forecast ranges from $30 to $65 MM dollars, an absolute accuracy of $1,000 may require an unreasonably large number of trials to reach. So, the drawback of using absolute precision is that it may require experimentation to determine reasonable values.

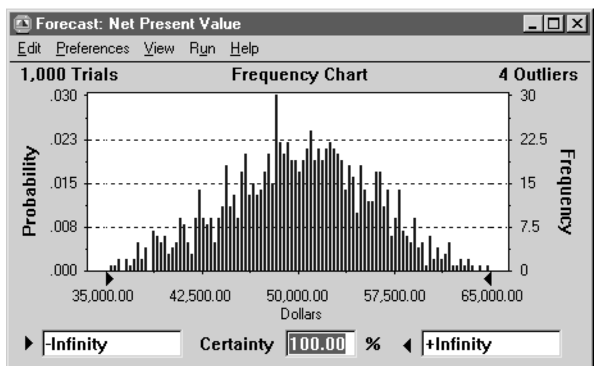


Figure 5.1 Relative precision control example

Specifying precision in relative terms can give you greater control of the simulation when the shape and scale of the forecast distribution is largely unknown and you are only interested in the accuracy as it relates to the overall distribution itself. In the NPV example above, you may not know or care that the distribution ranges $30,000 to $65,000 dollars or from $30 to $65 MM dollars. You might only require that the simulation's estimate of the mean fall within plus or minus 5% of itself.

The drawback of using relative precision is encountered when the forecast statistic is close to zero in relation to the overall distribution. In the NPV example below, the forecast has the same range width as the one above, but this time the distribution straddles the break-even point of zero. A relative precision of 5% of the mean, or roughly $60, results in a very small confidence interval that may take an unexpectedly large number of trials to satisfy.

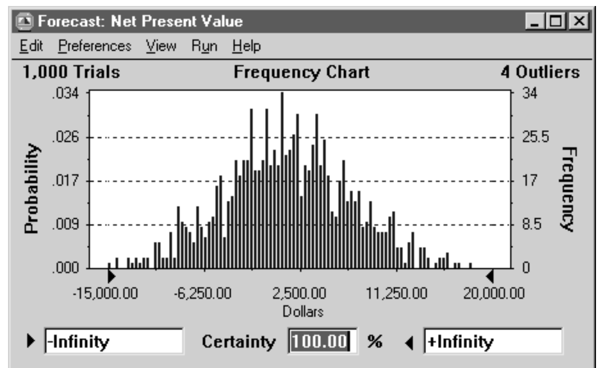


Figure 5.2 Absolute precision control example

Finally, Crystal Ball combines the individual forecast precision options with the confidence level value found in the Run Preferences > Trials dialog to calculate confidence intervals. Generally, it is a good idea to leave this value at 95% or 90% so that you can have a high degree of confidence that the precision requirements have been met. However, if you have a large number of forecasts defined with precision control set, you can adjust the confidence level up or down to globally change the accuracy of all forecasts together.

Changing the sample size

The sample size option is located on the Run Preferences > Sampling dialog. Increasing the sample size affects three Crystal Ball features:

- Correlation
- Latin hypercube sampling
- Sensitivity analysis

The sample size, which is initially set to 500, has a different significance for each of the three features:

Correlation
: Requires that a large sample of random values be generated ahead of time for each correlated assumption. Prior to being used in the simulation, the values in the samples are re-arranged by a mathematical technique to create the desired correlations. If all the values in the samples are used before the simulation stops, the simulation is temporarily suspended while a new group of samples are generated and correlated.

Latin hypercube sampling
: Divides each assumption's distribution into a number of intervals of equal probability. The sample size governs the number of intervals for each distribution. Crystal Ball generates an assumption value for each interval according to the interval's probability distribution.

Sensitivity analysis
: During a simulation, sensitivity analysis collects the generated assumption and forecast values and stores them in samples. After every so many trials, as determined by the sample size, Crystal Ball computes correlation coefficients for all pairs of samples. At the end of the simulation, Crystal Ball calculates final sensitivity coefficients by averaging the coefficients for each pair of samples.

While any sample size greater than 100 should produce sufficiently acceptable results, you can set this number higher to maximize accuracy. You can set sample size as large as 5,000.

The increased accuracy resulting from the use of larger samples, however, requires additional memory and reduces overall system responsiveness. If either of these become an issue, reduce the sample size.

Changing the sampling method

Choosing between Monte Carlo and Latin Hypercube sampling affects how the random numbers are generated for the individual assumptions.

In almost all cases, Latin Hypercube will produce more accurate forecast statistics (especially the mean) given the same number of trials as Monte Carlo since it is a more consistent sampling method. If you are primarily interested in the accuracy of your statistics, you should select Latin Hypercube as the sampling method.

If you are primarily interested in evaluating how your spreadsheet model behaves under various "what-if" scenarios, you should select Monte Carlo as the sampling method. Monte Carlo produces assumptions with the most randomness and hence will simulate real-life situations the best.

Simulation speed

Monte Carlo simulations can be very time-consuming. You can change a number of factors that affect the speed of simulations. The factors are listed below in order of importance:

1. **Change the Speed Run Preferences.**

   In the Run Preferences > Speed dialog, there are options that can dramatically increase the speed of your simulation. You should:

   - Minimize Excel during simulations
   - Suppress forecast windows during simulations
   - Increase the burst mode (see below)

2. **Use the Precision Control feature.**

   The precision control feature can be used to either increase the accuracy of your simulations or increase the speed of your simulations. If you set the absolute or relative precision to a low value (or your confidence level to a high number), your simulations will be more accurate but might run significantly longer. However, if you do not need as accurate of a result, you can set the absolute or relative precision to a high value (or the confidence level to a lower number) and the simulation speed will increase.

   Using this feature to speed up your model will require you to experiment with different absolute/relative precision control values and confidence levels.

3. **Increase the burst mode.**

   The Burst Mode option in Crystal Ball enhances the simulation speed of small models by grouping batches of data for the spreadsheet program to process. You can set this option by selecting Run > Run Preferences > Speed. The Burst Mode option is on by default.

   If you are working with large spreadsheet models that take a long time to recalculate, burst mode will have little effect. You should turn off Burst Mode to enhance responsiveness of the system to mouse clicks and menu selections. If system responsiveness is not an issue, set the burst count to 50 or 100 to maximize the simulation speed.

4. **Reduce the size of the model by reducing the number of assumptions, forecasts, and correlations.**

   Large models require more time per trial. For example, a model that takes 3 or 4 seconds per recalculation cycle will take up to an hour to simulate 1,000 trials.

   Greater numbers of assumptions and forecasts slow the simulation, especially if the assumptions and forecasts are scattered across many spreadsheets in your model. Start by examining the structure and nature of your model to locate possible efficiencies. You can also use the sensitivity feature or the Tornado Chart tool to determine which assumptions contribute the least amount of uncertainty to your bottom-line forecasts. Freeze or eliminate the least important assumptions from the simulation.

Then in the Sensitivity Preferences dialog, include only Other Forecasts. With the sensitivity chart displayed, you can decide which formulas contribute only slightly to the bottom line and can replace them with constants. The model's recalculation speed will increase with a slight loss of information.

Correlated assumptions can also consume a significant amount of processing time; the time grows geometrically as the number of correlated assumptions increases.

5. **Reduce the use of other applications.**

   Quitting other applications and closing or minimizing windows can be helpful in reducing overhead and increasing simulation speed.

6. **Increase your system's RAM.**

   The amount of RAM in your computer has a large affect on the speed of simulations. Modern operating systems give applications such as spreadsheets the appearance of additional RAM through the use of *virtual memory*.

   Virtual memory lets you run a greater number of applications than would otherwise be possible, but slows down overall processing speed because the system is frequently accessing the hard drive. If you hear your hard disk being used during a simulation, there is not enough RAM to hold all parts of the simulation. Buying more RAM or turning off virtual memory (if possible) are solutions to this problem.

7. **Consider CB Turbo.**

   CB Turbo is a separate product that speeds up the time required to perform lengthy, complex simulation by distributing the model recalculations to other workstations on the network. For more information on CB Turbo, contact Decisioneering using information in your README file, or visit the Decisioneering web site at www.decisioneering.com.

Memory use

Monte Carlo simulations can be very memory-intensive. Your computer's random access memory (RAM) must contain, at the same time: part of the operating system, your spreadsheet program, Crystal Ball, your spreadsheet files, the assumption and forecast definitions, and the trials of the simulation. So, it is important that you have a basic understanding of how Crystal Ball uses RAM.

- Each assumption consumes approximately 300 bytes. Custom distributions use a minimum of 300 bytes plus either 12 bytes per data value or 32 bytes per discrete or continuous range.
- Each forecast consumes a minimum of 2,800 bytes. During a simulation, each forecast uses 10 bytes of memory for each trial. For example, a simulation of 1,000 trials for a model with two forecasts uses a total memory of:

  $$2 \times (2,800 + 10 \times 1,000) = 25,600 \text{ bytes}$$

- If either correlation, Latin hypercube sampling, or sensitivity analysis is in use, the minimum amount of memory required is:

  $$10 \times \text{sample size} \times A$$

  $$\text{where } A = \text{number of assumptions}$$

  Correlation uses additional memory equal to:

  $$3 \times 10 \times A^2$$

  Sensitivity analysis uses additional memory equal to:

  $$10 \times (A \times F) + 10 \times F^2$$

  $$\text{where } F = \text{number of forecasts}$$

In addition to the RAM required during the simulation, memory is needed at the end of the simulation. Crystal Ball checks your system's free RAM frequently during a simulation. When the amount of free RAM dips below a certain threshold (typically between 1 and 2 megabytes), the simulation stops and the "memory limit reached" message appears. Crystal Ball needs this remaining memory for tasks such as opening forecast windows, creating reports, and performing other functions on your computer. If the alert message appears, take the steps listed in the next section.

Correlated assumptions

There is a practical limit of about 50 fully correlated assumptions (assumptions that are correlated to every other assumption) and up to about 200 serially correlated assumptions (assumptions that are correlated to one or two other assumptions) in each spreadsheet model. If you have a large number of correlated assumptions:

- For coefficients that are close to zero and are nearly independent, remove the assumption
- For coefficients that are close to one, replace the assumption with a formula in your spreadsheet.

Sensitivity analysis

The sensitivity feature in Crystal Ball has one minor limitation. If you continue a stopped simulation, Crystal Ball clears the sensitivity data and computes from the continuation point. Therefore, if you need to analyze a spreadsheet model's sensitivity to its input parameters, you should run the entire simulation without interruption.

Excel tips

This section contains ways to improve the value of Crystal Ball for various versions of Excel.

Customizing report and data layouts

Crystal Ball creates reports and data extracts as spreadsheets. This gives you a lot of control and flexibility over how Crystal Ball handles the results of a simulation. You can modify, preview, print, or save results the same way you do in other spreadsheets.

The Crystal Ball folder has two files named **Report** and **Data** that you can use to create spreadsheets. Since the files are just Excel templates, you can easily change them to suit your needs. For example, you might decide to change the default font or resize the margins and column widths.

To change the template files:

1. **Select File > Open.**
2. **In the Crystal Ball folder, find either the Report.xls or Data.xls file.**
3. **Open the file. The title of the open window should read either Report or Data.**
4. **Make any desired changes to the template. (If you are changing column widths, you should temporarily turn on gridlines and column headers using Excel's Tools > Options > View command.)**
5. **Save the template file.**

Chapter 6

Crystal Ball Tools

In this chapter

This chapter describes the Crystal Ball tools:

- Batch Fit
- Bootstrap
- Correlation Matrix
- Decision Table
- Tornado Chart
- Two-dimensional Simulation

For each tool, there is a general description, an introduction tutorial, and a description of all dialogs, fields, and options.

Overview

Crystal Ball tools are Visual Basic programs that extend the functionality of Crystal Ball. They are:

Batch Fit — Automatically fits selected continuous probability distributions to multiple data series.

Bootstrap — Addresses the reliability and accuracy of forecast statistics.

Correlation Matrix — Rapidly defines and automates correlations of assumptions.

Decision Table — Evaluates the effects of alternate decisions in a simulation model.

Scenario Analysis — Displays what inputs created particular outputs. The documentation for the Scenario Analysis tool is available online at www.decisioneering.com.

Tornado Chart — Individually analyzes the impact of each model variable on a target outcome.

Two-dimensional Simulation — Independently addresses uncertainty and variability using two-dimensional simulation.

This chapter describes each tool, provides a step-by-step example for using each tool, and describes the options for each tool.

Batch Fit tool

The Batch Fit tool fits probability distributions to multiple data series. You can select any or all of the continuous probability distributions (normal, triangular, uniform, etc.) to fit to any number of series limited only by the size of your spreadsheet.

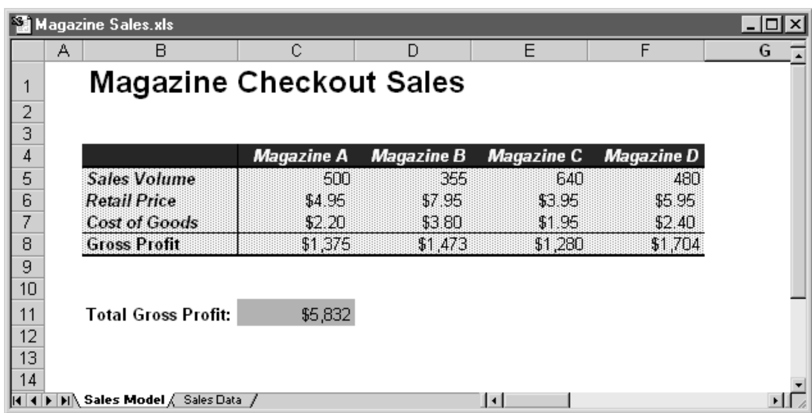
Batch Fit is intended to help you create assumptions when you have historical data for several variables. It selects which distribution best fits each series of historical data, and gives you the distribution and its associated parameters for you to use in your model. This tool can also give you a table of goodness-of-fit statistics for each distribution, so you can compare the fit of the best distribution to the fits of the other distributions.

To use the Batch Fit tool, your data must be contiguous (in adjacent rows or columns) in either rows or columns.

You can select any combination of the continuous probability distributions to fit to all your data series.

Batch Fit example

In the Crystal Ball Examples folder there is a Magazine Sales.xls spreadsheet you can use to experiment with the Batch Fit tool. This spreadsheet model shows the estimated gross profit resulting from newsstand sales of four of the company's most popular magazines.


Magazine Sales.xls

Magazine Checkout Sales

	Magazine A	Magazine B	Magazine C	Magazine D
Sales Volume	500	355	640	480
Retail Price	$4.95	$7.95	$3.95	$5.95
Cost of Goods	$2.20	$3.80	$1.95	$2.40
Gross Profit	$1,375	$1,473	$1,280	$1,704

Total Gross Profit: $5,832

Sales Model / Sales Data

Figure 6.1 Magazine Sales spreadsheet

To run Batch Fit:

1. **In Excel with Crystal Ball loaded, open the spreadsheet Magazine Sales.xls.**
2. **Select CBTools > Batch Fit.**
3. **Make sure all the distributions are in the Selected Distributions list.**
4. **Click on Next.**

5. **Click on the Select Cells icon to the right of the Location Of Data Series field.**
6. **Select the Sales Data worksheet.**
7. **Select cells A1 through D361.**

8. **Click on the return icon to return to the tool dialog.**
9. **Select Data In Columns.**

   The default is Data In Columns.

10. **Select the Fitness Criteria: Chi-Square.**

    The default is Chi-Square.

11. **Check the First Row Contains Headers option.**
12. **Click on Next.**

    The Select Output Options (3 Of 3) dialog appears.

13. **Select the Create Output On The Active Worksheet option.**

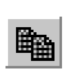
14. **Click on the Select Cells icon to the right of the Specify Upper Left Corner Of The Output field.**
15. **Select the Sales Model tab.**
16. **Select cell B13.**

17. **Click on the return icon to return to the tool dialog.**
18. **Make sure the Format Output option is checked.**
19. **Check the Show Table Of Goodness-of-fit Statistics option.**
20. **Set the Output Orientation to Data In Columns.**
21. **Click on Start.**

    The tool fits each selected distribution to each data series. The results appear on the Sales Model worksheet to the right of the existing table.

22. **Copy assumption data into the worksheet.**

    a. **Select cells C15 through F15.**



    b. **Select Cell > Copy Data.**

Crystal Ball Note: *This function copies Crystal Ball data only, not the cell value.*

c. **Select cells C5 through F5.**

d. **Select Cell > Paste Data.**

The assumptions copy to the first row of the table.

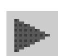
e. **Select cells C15 through F15.**

f. **Select Cell > Clear Data.**

This deletes the original Crystal Ball assumptions.

23. **In the Run > Run Preferences > Sampling dialog, set:**
    - Random Number Generation to use the Same Sequence Of Random Numbers and a seed value of 999
    - Monte Carlo simulation

    When using this tool, use these options to make the resulting simulations comparable.
24. **In the Run > Run Preferences > Trials dialog, set the Maximum Number Of Trials to 500.**
25. **Click on OK.**



26. **Select Run > Run.**

Interpreting the results

In this example, you had historical data with no trend or seasonality for all your variables.

Crystal Ball Note: *If your historical data has a time-series element, trend, or seasonality, you should use CB Predictor instead of the Fit Distribution functionality. CB Predictor is a component of the Crystal Ball 2000, Professional Edition.*

When the batch fit tool runs, it fits each column of data to each continuous distribution. For each fit of a distribution to a set of data, the tool calculates the indicated goodness-of-fit test for each distribution. The distribution with the best fit is placed into to create an assumption cell that you can copy to the appropriate location in your model.

In this example, the four columns of data resulted in four assumptions that you copied to the appropriate locations. The forecast is already defined for you, and running a simulation produces a forecast chart of the total gross profit from the

Magazine Sales workbook. If you replace the -Infinity with $5,500, you find that the chances of making this amount of profit is approximately 72.6%.

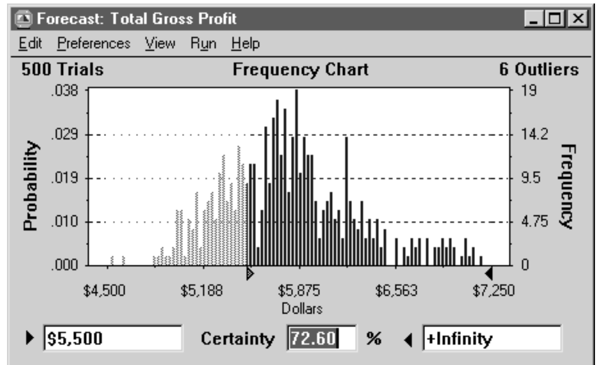


Figure 6.2 Magazine Sales profit from newsstand sales

Batch Fit dialogs

Select Distributions (Step 1 Of 3) dialog

The Select Distributions dialog lets you choose which distributions to try to fit to each set of data.

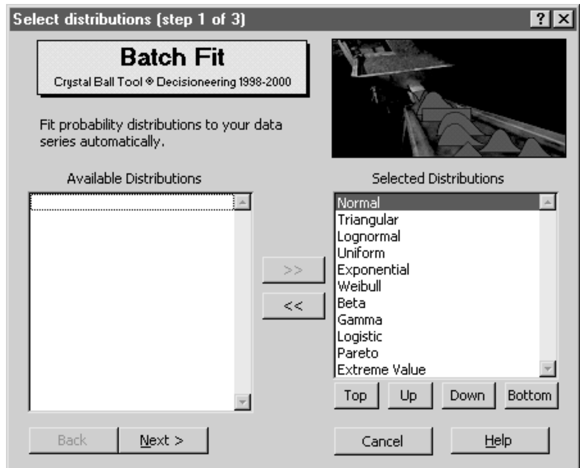


Figure 6.3 Select Distributions dialog for Batch Fit

The fields and buttons in this dialog are:

Available Distributions
: Lists all the continuous probability distributions that you can fit to data series, but that are not selected.

  By default, all the probability distributions are selected.

Selected Distributions
: Lists all the probability distributions that you selected to fit to all the data series. For each selected distribution, the tool will fit the distribution to each selected data series and calculate the goodness-of-fit statistic.

Top
: Moves the highlighted distributions to the top of the Selected Distributions list, which can affect the output order of the distributions.

Up
: Moves the highlighted distributions up one position in the Selected Distributions list, which can affect the output order of the distributions.

Down
: Moves the highlighted distributions down one position in the Selected Distributions list, which can affect the output order of the distributions.

Bottom
: Moves the highlighted distributions to the bottom of the Selected Distributions list, which can affect the output order of the distributions.

Next
: Opens the Select Input Options (Step 2 of 3) dialog.

Select Input Options (Step 2 Of 3) dialog

This dialog lets you specify the location of the data that you want to fit to the distributions selected on the previous dialog. It also lets you specify other input-related options.

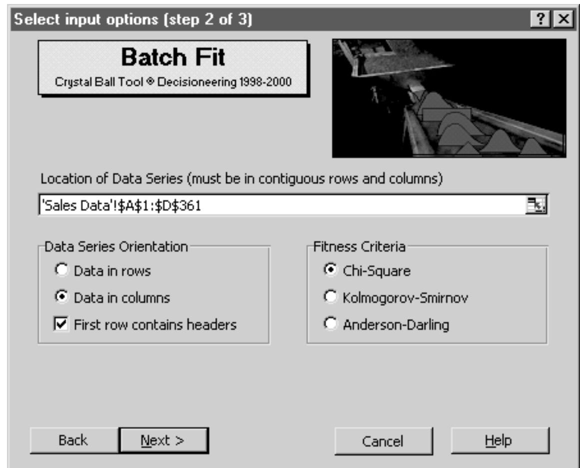


Figure 6.4 Select Input Options (Step 2 of 3) dialog for Batch Fit

The fields and options in this dialog are:

Location Of Data Series
: Lets you enter or interactively select the cells that contain data to fit. If your data has headers at the beginning of the rows or columns of data, include them in your selection and make sure you select the First Row/Column Contains Headers option.

: The data must be in adjacent rows or columns.

Data Series Orientation
: Sets whether your data is in rows or columns.

First Row Contains Headers
: Lets you include row or column titles in your selection. The headers are used in the output.

Fitness Criteria Selects which method to use to determine which distribution best fits the data. The three goodness-of-fit tests are: Chi-Square, Kolmogorov-Smirnov, and Anderson-Darling. See “How distribution fitting works” on page 138 for more information on these tests.

Back Returns to the Select Distributions (Step 1 of 3) dialog.

Next Advances to the Select Output Options (Step 3 of 3) dialog.

Select Output Options (Step 3 Of 3) dialog

This dialog lets you set output options that control the tool.

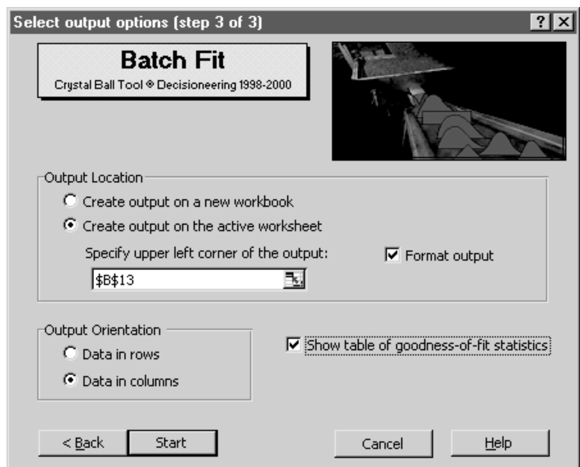


Figure 6.5 Specify Output Options dialog for Batch Fit

The options and buttons in this dialog are:

Output Location
Lets you select whether to output the results to a new workbook or in the active workbook, starting at the indicated cell.

Format Output Uses any special cell formatting for your data in your output.

Output Orientation
: Sets whether to put your output in rows or columns.

Show Table Of Goodness-of-fit Statistics
: Outputs the selected goodness-of-fit statistic for each distribution.

Back
: Returns to the Select Input Options (Step 2 of 3) dialog.

Start
: Starts the tool, automatically fitting distributions to your data and creating output of the assumptions and statistics.

Bootstrap tool

Bootstrap is a simple technique that estimates the reliability or accuracy of forecast statistics or other sample data. Classical methods used in the past relied on mathematical formulas to describe the accuracy of sample statistics. These methods assume that the distribution of a sample statistic approaches a normal distribution, making the calculation of the statistic's standard error or confidence interval relatively easy. However, when a statistic's sampling distribution is not normally distributed or easily found, these classical methods are difficult to use or are invalid.

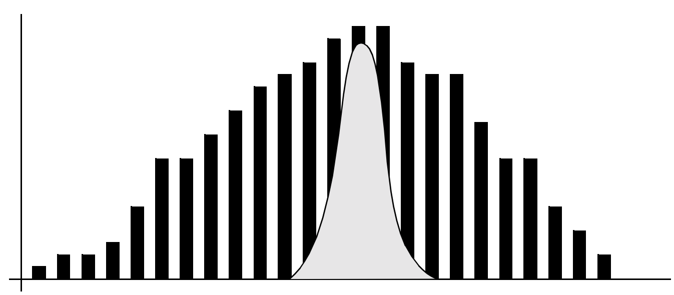


Figure 6.6 Sampling distribution of a mean statistic

In contrast, bootstrapping analyzes sample statistics empirically by repeatedly sampling the data and creating distributions of the different statistics from each sampling. The term bootstrap comes from the saying, “to pull oneself up by one’s own bootstraps,” since this method uses the distribution of statistics themselves to analyze the statistics’ accuracy.

There are two bootstrap methods available with this tool:

One-simulation method
: Simulates the model data once (creating the original sample), and then repeatedly resamples those simulation trials (the original sample values). Resampling creates a new sample from the original sample **with replacement**. It then creates a distribution of the statistics calculated from each resample.

  This method assumes only that the original simulation data accurately portrays the true forecast distribution, which is likely if the sample is large enough. This method isn’t as accurate as the multiple-simulation method, but it takes significantly less time to run.

Glossary Term:
with replacement—Returns the selected value to the sample before selecting another value, letting the selector possibly reselect the same value.

Multiple-simulation method
: Repeatedly simulates the model, and then creates a distribution of the statistics from each simulation.

  This method is more accurate than the one-simulation method, but it might take a prohibitive amount of time.

Crystal Ball Tool Note: *When you use the multiple-simulation method, the tool temporarily turns off the Use Same Sequence Of Random Numbers option.*

Statistical Note: *In statistics literature, the one-simulation method is also called the non-parametric bootstrap, and the multi-simulation method is also called the parametric bootstrap.*

Since the bootstrap technique doesn’t assume that the sampling distribution is normally distributed, you can use it to estimate the sampling distribution of any statistic, even an unconventional one such as the minimum or maximum endpoint of a forecast. You can also easily estimate complex statistics, such as the correlation coefficient of two data sets, or combinations of statistics, such as the ratio of a mean to a variance.

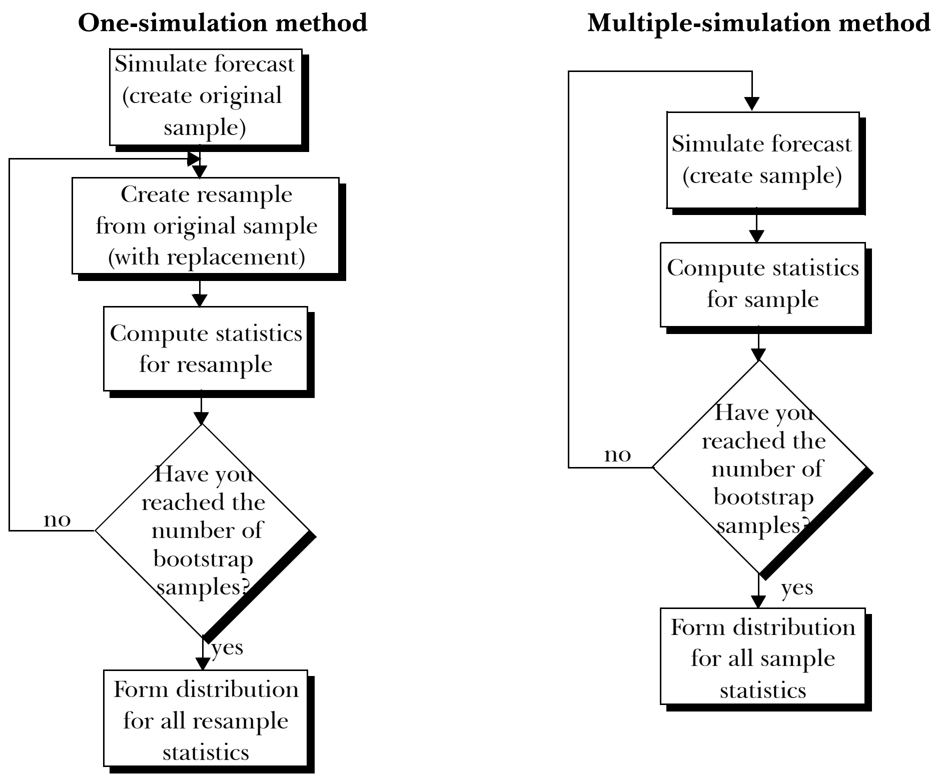


Figure 6.7 Comparison of one-simulation and multiple-simulation bootstrap methods

Statistical Note: *To estimate the accuracy of Latin hypercube statistics, you must use the multiple-simulation method.*

Bootstrap example

In the Crystal Ball Examples folder there is a Futura Apartments.xls spreadsheet you can use to experiment with the Bootstrap tool. This spreadsheet model forecasts the profit and loss for an apartment complex.

To run Bootstrap:

1. **In Excel with Crystal Ball loaded, open the spreadsheet Futura Apartments.xls.**
2. **Select CBTools > Bootstrap.**

   The Specify Target dialog appears.
3. **Set the target by selecting Profit Or Loss from the forecast list.**
4. **Click on Next.**

   The Specify Options (Step 2 of 3) dialog appears.
5. **Make sure the one-simulation method and the statistics options are selected.**
6. **Click on Next.**

   The Specify Options (Step 3 of 3) dialog appears.
7. **Set the following options:**
   - Bootstrap samples is set to 200
   - Trials per sample is set to 500
   - Show only target forecast
8. **Click on Start.**

   The bootstrap tool displays a forecast chart of the distributions for each statistic and creates a spreadsheet summarizing the data.

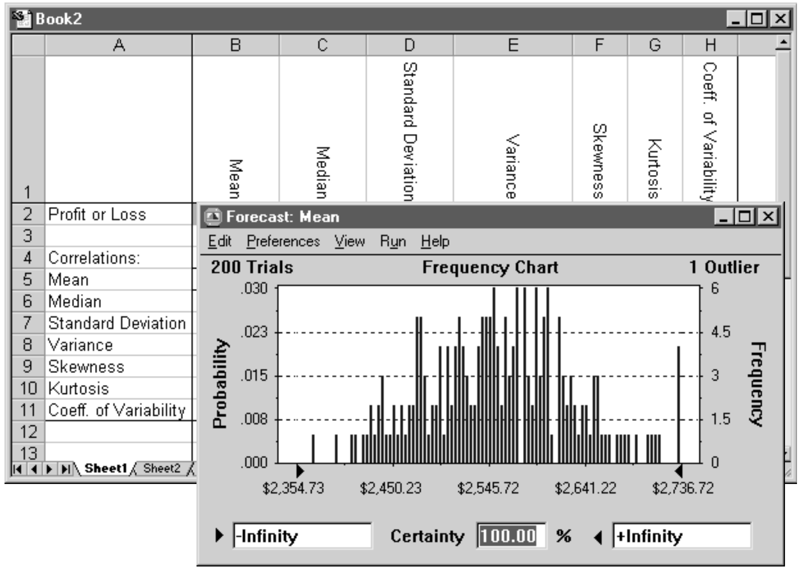


Figure 6.8 Bootstrap example results

Interpreting the results

The Bootstrap tool displays sampling distributions in forecast charts for the following statistics:

- Mean
- Median
- Standard deviation
- Variance
- Skewness
- Kurtosis
- Coefficient of variability

When you use the multiple-simulation method, the tool also displays sampling distributions for these statistics:

- Range minimum
- Range maximum
- Range width

For percentiles, the Bootstrap tool displays the percentile sampling distributions on the overlay and trend charts. To display the individual percentile forecast charts select Run > Forecast Windows.

The forecast charts visually indicate the accuracy of each statistic. A narrow and symmetrical distribution is better than a wide and skewed distribution.

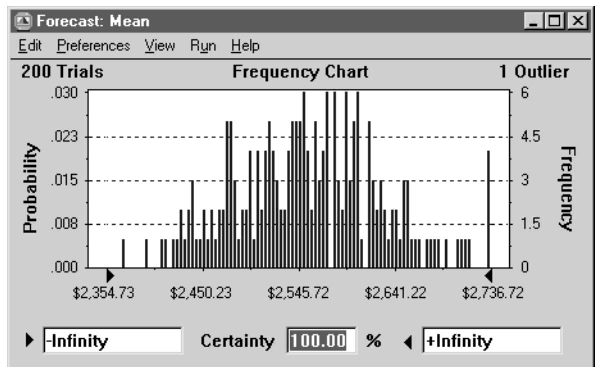
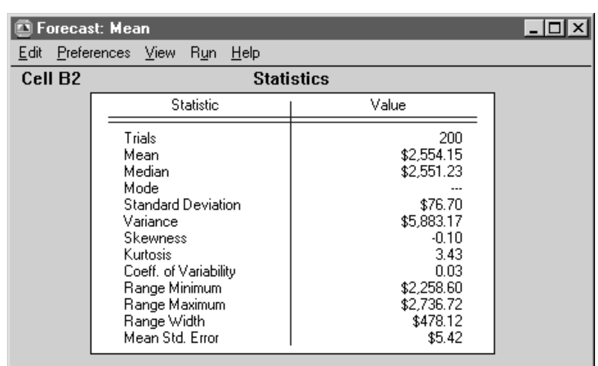


Figure 6.9 Bootstrap forecast chart of mean

The statistic view further lets you analyze the statistics' sampling distribution. If the standard deviation (standard error of the statistic) or coefficient of variability is very large, the statistic might not be reliable and might require more trials. This example has a relatively low standard error and coefficient of variability, so the forecast mean is an accurate estimate of the actual mean.



Statistic	Value
Trials	200
Mean	$2,554.15
Median	$2,551.23
Mode	---
Standard Deviation	$76.70
Variance	$5,883.17
Skewness	-0.10
Kurtosis	3.43
Coeff. of Variability	0.03
Range Minimum	$2,258.60
Range Maximum	$2,736.72
Range Width	$478.12
Mean Std. Error	$5.42

Figure 6.10 Bootstrap forecast statistics of mean

The results workbook has a correlation matrix showing the correlations between the various statistics. High correlation between certain statistics, such as between the mean and the standard deviation, usually indicates a highly skewed distribution.

You can also use the Bootstrap tool to analyze the distribution of percentiles, but you should run at least 1,000 bootstrap samples and 1,000 trials per sample to obtain good sampling distributions for these statistics.[1]

Bootstrap dialogs

Specify Target (Step 1 Of 3) dialog

The Specify Target dialog lets you analyze the statistics of a specified forecast, cell, or cell range.

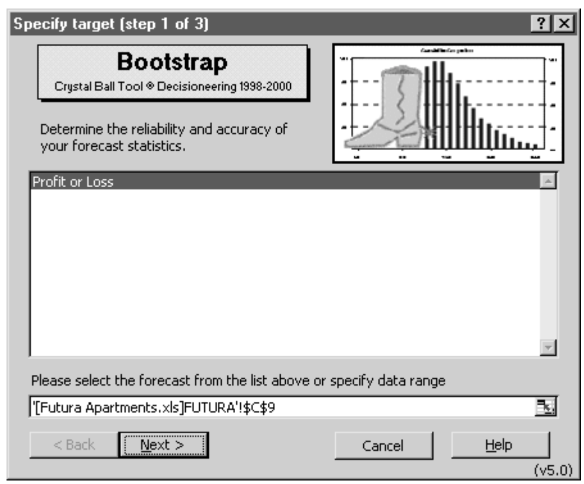


Figure 6.11 Specify Target dialog for Bootstrap

1. Efron, Bradley, and Robert J. Tibshirani. *Monographs on Statistics and Applied Probability, vol. 57: An Introduction to the Bootstrap*. New York: Chapman & Hall, 1993.

The fields for this dialog are:

Forecast List — Lists all the forecast cells in all open spreadsheets. When you select a forecast from the list, its cell information automatically appears in the Enter Target Cell field.

The first forecast is selected by default.

Enter Target Cell — Describes the cell location of the selected forecast or formula. If you select a forecast from the list above, the cell information automatically appears in this field.

You can use this field to select a formula cell instead of a forecast.

Next — Opens the next dialog for defining tool options.

Specify Options (Step 2 Of 3) dialog

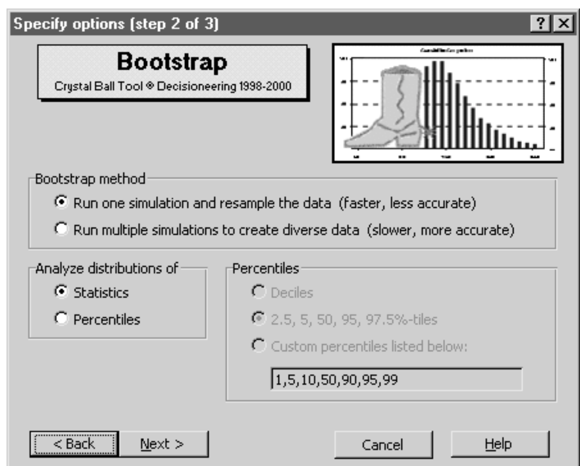


Figure 6.12 First Specify Options dialog for Bootstrap

This first options dialog has the following options, fields, and buttons:

Bootstrap Method
: Selects whether to use the one-simulation or multiple-simulation bootstrap method. For more information on these two methods, see “Bootstrap tool” on page 264.

  The default is the one-simulation method.

Analyze Distributions Of
: Selects whether to analyze distributions of statistics or percentiles. If you select Percentiles, you must complete the Percentiles options.

  The default is Statistics.

Percentiles
: Selects which target percentiles to analyze. You can select either: deciles (the 10, 20, 30, 40, 50, 60, 70, 80, and 90 percentiles); 2.5, 5, 50, 90, and 97.5 percentiles; or a custom list of percentiles you enter in the field. A custom list can have up to 10 percentiles (between 0 and 100, exclusive) separated by commas.

Back
: Opens the Specify Target dialog for specifying a target.

Next
: Opens the next dialog for defining more tool options.

Specify Options (Step 3 Of 3) dialog

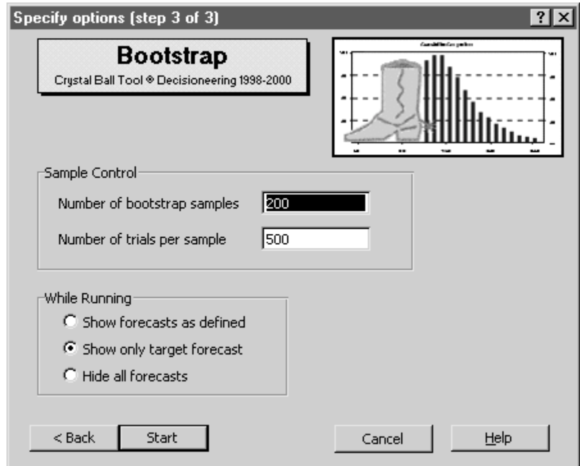


Figure 6.13 Second Specify Options dialog for Bootstrap

This second options dialog has the following options, fields, and buttons.

Sample Control
Sets the number of bootstrap samples and the number of trials per sample.

The default bootstrap samples is 200, and the default number of trials is the number set in the Crystal Ball Run Preferences dialog.

While Running
Lets you select what forecasts to show while you are running the tool. You can view all the defined forecasts, only the target forecast, or none of the forecasts.

Back
Opens the previous dialog for defining other tool options.

Start
Runs the Bootstrap tool.

Correlation matrix tool

Correlations

When the values of two variables depend on each other in any way, you should correlate them to increase the accuracy of your simulation's forecast results.

There are two types of correlations:

positive correlation
: Indicates that two assumptions increase or decrease together. The price of gasoline and shipping costs increase and decrease together.

negative correlation
: Indicates that an increase in one assumption results in a decrease in the other assumption. The more items you buy from a particular vendor, the lower the unit cost.

The correlation coefficient range is -1 to 1, where 0 indicates no correlation. The closer the coefficient is to ±1, the stronger the relationship between the assumptions. You should never use a coefficient of ±1; represent relationships this closely correlated with formulas in the spreadsheet model.

Correlation matrix

In Crystal Ball, you enter correlations one at a time using the Correlation dialog. Instead of manually entering the correlations this way, you can use the Correlation Matrix tool to define a matrix of correlations between assumptions in one simple step. This saves time and effort when building your spreadsheet model, especially for models with many correlated assumptions.

The correlation matrix is either an upper or lower triangular matrix with ones along the diagonal. When entering coefficients, think of the matrix as a multiplication table. If you follow one assumption along its horizontal row and the second along its vertical column, the value in the cell where they meet is their correlation coefficient.

	Assumption 1	Assumption 2	Assumption 3
Assumption 1	1.000		
Assumption 2		1.000	
Assumption 3			1.000

Figure 6.14 Correlation matrix

If you enter inconsistent correlations, Crystal Ball tries to adjust the correlations so they don't conflict. For more information on inconsistent correlations, see "Specifying correlations between assumptions" on page 145.

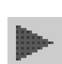
Correlation Matrix example

In the Crystal Ball Examples folder there is a Portfolio Allocation.xls spreadsheet you can use to experiment with the Correlation Matrix tool. This spreadsheet calculates the total expected return for an investment model. In this example, you will run a simulation without correlations, and then add the correlations and rerun the simulation for comparison.

To run Correlation Matrix:

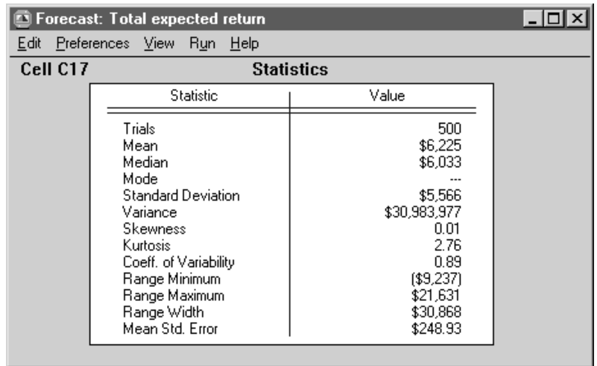
1. **In Excel with Crystal Ball loaded, open the spreadsheet Portfolio Allocation.xls.**

2. **Set the following options in the Run > Run Preferences dialog.**
   - Maximum Number Of Trials Per Simulation set to 500
   - Random Number Generation set to Use Same Sequence Of Random Numbers and An Initial Seed Value of 999
   - Use Monte Carlo simulation



3. **Run a simulation by selecting Run > Start.**

   The forecast statistics for the simulation are shown below.



Statistic	Value
Trials	500
Mean	$6,225
Median	$6,033
Mode	---
Standard Deviation	$5,566
Variance	$30,983,977
Skewness	0.01
Kurtosis	2.76
Coeff. of Variability	0.89
Range Minimum	($9,237)
Range Maximum	$21,631
Range Width	$30,868
Mean Std. Error	$248.93

Figure 6.15 Uncorrelated simulation statistics

4. **Select CBTools > Correlation Matrix.**

   The Select Assumptions dialog appears.

5. **Include all the assumptions in the correlation matrix by moving all the assumptions from the Available Assumptions field to the Selected Assumptions field by either:**

   - Double-clicking on each assumption to move.
   - Selecting each assumption to move and clicking on >> to move it.
   - Making an extended selection using the <Shift> or <Ctrl> keys.

6. **Click on Next.**

   The Specify Options dialog appears.

7. **Set the following options:**

   - Location Of Matrix set to Create A Temporary Correlation Matrix On A New Worksheet
   - Orientation set to Upper Triangular Matrix

8. **Click on Start.**

   The tool creates a temporary matrix in a new workbook.

9. Enter the following correlation coefficients into the matrix.

Load the matrix	Money Market fund	Income fund	Growth and Income fund	Aggressive Growth fund
Money Market fund	1.000	0.200	0.100	0.100
Income fund		1.000	0.300	0.200
Growth and Income fund			1.000	0.500
Aggressive Growth fund				1.000

Crystal Ball Tool Note: *Leaving a cell blank is not the same as entering a zero. Values that are not specified in the matrix will be "filled in" with estimates of appropriate values when the simulation runs.*

10. Click on Load The Matrix.

The tool loads the correlation coefficients from the matrix into your Crystal Ball model.

Crystal Ball Tool Note: *If a Matrix Successfully Loaded message doesn't appear, press <Tab> or <Return> to exit the current cell and then click on Load The Matrix again.*

11. Reset the simulation.

12. Rerun the simulation.

The forecast statistics for the correlated simulation is shown below.

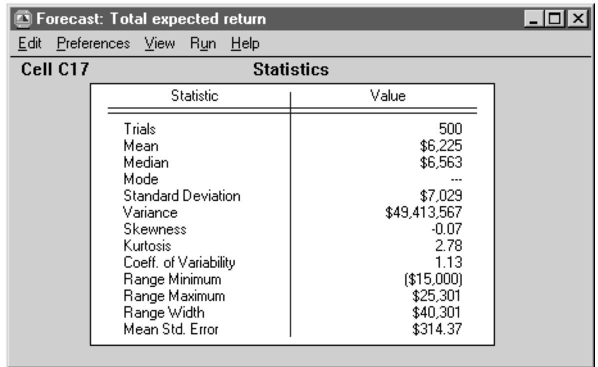


Figure 6.16 Correlated simulation statistics

The standard deviation is now much higher than the original simulation due to the correlations. The original model without the correlations ignored this risk factor and its effects.

Correlation Matrix dialogs

Select Assumptions (Step 1 Of 2) dialog

The Select Assumptions dialog lets you select which assumptions to use in the correlation matrix.

You must select at least two assumptions. There is a practical limit for one correlation matrix of about 50 fully correlated assumptions (assumptions that are correlated to every other assumption) or up to about 200 serially correlated assumptions (assumptions that are correlated to 1 or 2 other assumptions). If you have more than 50 assumptions, either eliminate very small correlations, or replace assumptions that have correlation coefficients very close to 1.0 with formulas in your spreadsheet.

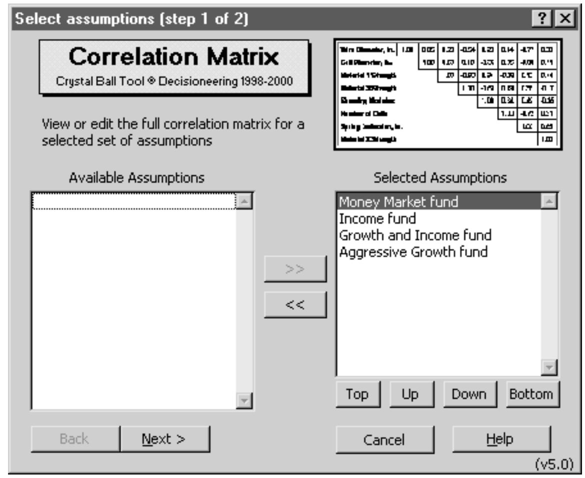


Figure 6.17 Select Assumptions dialog for Correlation Matrix

The fields and buttons in this dialog are:

Available Assumptions

Lists all the assumptions defined in the active workbook. To move an assumption into the Selected Assumptions list, either double-click on an assumption or select an assumption and click on >>. You can also make an extended selection using the <Shift> or <Ctrl> keys.

By default, all assumptions start in the Available Assumptions list.

Selected Assumptions

Lists the assumptions selected for inclusion in the matrix. To move an assumption into the Available Assumption list, either double-click on an assumption or select an assumption and click on <<. You can also make an extended selection using the <Shift> or <Ctrl> keys.

By default, no assumptions are in this list.

Top	Moves one or more selected assumptions in the Selected Assumptions list to the top of the list (therefore, first in the matrix).
Bottom	Moves one or more selected assumptions in the Selected Assumptions list to the bottom of the list (therefore, last in the matrix).
Up	Moves one or more selected assumptions in the Selected Assumptions list up one position, changing its position in the matrix.
Down	Moves one or more selected assumptions in the Selected Assumptions list down one position, changing its position in the matrix.
Next	Opens the Specify Options dialog for defining correlation matrix options.

Specify Options (Step 2 Of 2) dialog

This dialog lets you set options for the matrix.

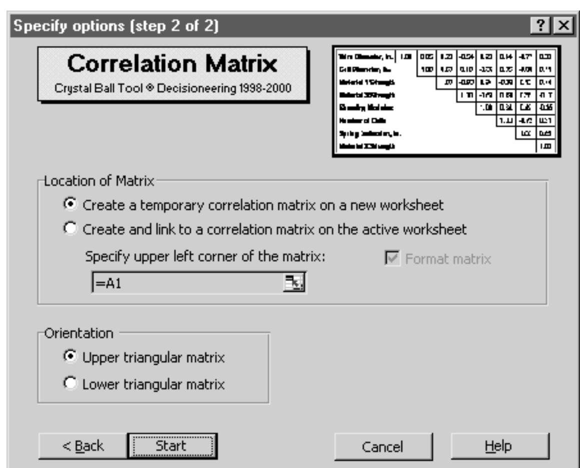


Figure 6.18 Specify Options dialog for Correlation Matrix

There are two sets of options in this dialog (described below):

- Location Of Matrix
- Orientation

This dialog also has the following buttons:

Back	Returns you to the Select Assumptions dialog.
Start	Creates the correlation matrix according to the options you specify.

Location Of Matrix options

The Location Of Matrix options lets you select whether to create the correlation matrix on the active worksheet or on a new worksheet. The options in this dialog are:

Create A Temporary Correlation Matrix On A New Worksheet

Creates the matrix on a new worksheet. Use this option if you don't want to permanently alter your spreadsheet.

After filling in the matrix, click on Load The Matrix to enter the coefficients into the model.

This is the default matrix location.

Create And Link To A Correlation Matrix On The Active Worksheet

Creates the matrix on the active worksheet and links it with cell references, so if you change the cells, the correlations update automatically. Select this option to embed the correlation matrix in your model.

If you select this option, you must indicate where to put the matrix in the model using the Specify Upper Left Corner option.

Specify Upper Left Corner

Lets you select the origin cell for the top left corner of the matrix by either entering the cell reference in the field or clicking on the cell in your worksheet.

This option is only available if you select to create the matrix on the worksheet.

Format Matrix — Adds borders and headings, and changes column widths to make the correlation matrix more readable.

This option is only available if you select to create the matrix on the active worksheet. This option is on by default.

Orientation options

The Orientation options let you select whether to make the correlation matrix an upper triangular or a lower triangular matrix.

Upper Triangular Matrix
: Creates the matrix with cells to fill in above the diagonal. This is the default.

Lower Triangular Matrix
: Creates the matrix with cells to fill in below the diagonal.

Decision Table tool

Decision variables are values that you can control, such as how much to charge for a product or how many wells to drill. But, in situations with uncertainty, it is not always obvious what effect changing a decision variable can have on the forecast results.

The Decision Table tool runs multiple simulations to test different values for one or two decision variables. The tool tests values across the range of the decision variables and puts the results in a table that you can analyze using Crystal Ball forecast, trend, or overlay charts.

The Decision Table tool is useful for investigating how changes in the values of a few decision variable affect the forecast results. For models that contain more than a handful of decision variables, or where you are trying to optimize the forecast results, use OptQuest for Crystal Ball.

OptQuest is a wizard-based program that enhances Crystal Ball by automatically finding optimal solutions to simulation models. This program is available with Crystal Ball 2000, Professional Edition.

Table 1: Comparison and contrast between the Decision Table tool and OptQuest.

	Decision Table tool	***OptQuest***
Process	Runs multiple Crystal Ball simulations for different values of decision variables	
Results	All displayed in a table	Only displays the best solutions.
Optimization	No	Yes
Number of variables	One or two	Unlimited
Variable range	Small	Small to large

Decision Table example

In the Crystal Ball Examples folder there is an Oil Field Development.xls spreadsheet you can use to experiment with the Decision Table tool. This spreadsheet model predicts how to best develop a new oil field by selecting the optimal number of wells to drill, rate of oil production, and size of the refinery to build that will maximize the net present value of the field.

To run Decision Table:

1. **In Excel with Crystal Ball loaded, open the spreadsheet Oil Field Development.xls.**

2. **In the Run > Run Preferences > Sampling dialog, set:**

    - Random Number Generation to use the Same Sequence Of Random Numbers and An Initial Seed Value of 999
    - Monte Carlo simulation

    When using this tool, use these options to make the resulting simulations comparable.

3. **In the Run Preferences dialog, click on OK.**
4. **Select CBTools > Decision Table.**

    The Specify Target dialog appears.

5. **Select the NPV forecast.**
6. **Click on Next.**

   The Select One Or Two Decisions dialog appears.

7. **Move Wells To Drill and Facility Size to the Chosen Decision Variables list.**

   a. **Select Wells To Drill in the Available Decision Variables field.**

   b. **Click on >>.**

   c. **Repeat steps 6a and 6b for the Facility Size.**

8. **Click on Next.**

   The Specify Options dialog appears.

9. **Set the following options:**
   - Number of values to test for Wells To Drill is 6
   - Number of values to test for Facility Size is 7
   - Number of trials per simulation is 500
   - Show only target forecast
10. **Click on Start.**

    The tool runs a simulation for each combination of decision variable values. It compiles the results in a table of forecast cells indexed by the decision variables.

Interpreting the results

For this example, the Decision Table tool ran 42 simulations, one for each combination of wells to drill and facility sizes. The simulation that resulted in the best mean NPV was the combination of 12 wells and a facility size of 150 mbd.

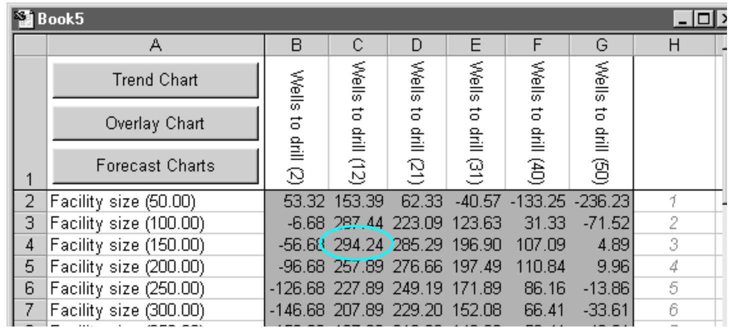


Figure 6.19 Decision table for Oil Field Development results

To view one or more of the forecasts in the decision table, select the cells and select Run > Forecast Charts. To compare one or more forecasts on the same chart, select the cells and click on Trend Chart or Overlay Chart.

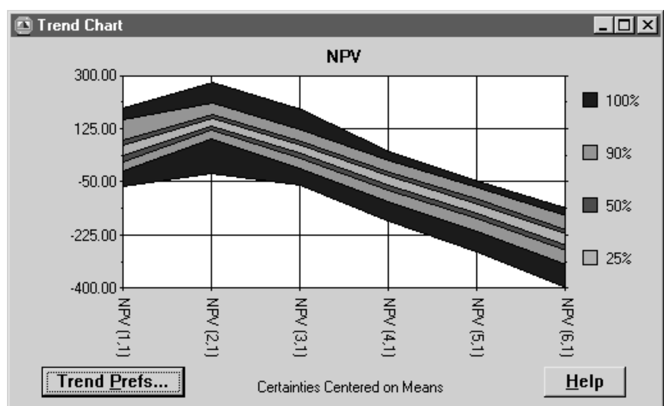


Figure 6.20 Trend chart of 150 mbd forecasts

You can create the above trend chart by selecting all the forecast cells in the Facility Size (150.00) row of the results table and clicking on Trend Chart. This chart shows that the forecast with the highest mean NPV also has the largest uncertainty compared to other forecasts with smaller NPVs of the same facility size. This indicates a higher risk that you could avoid with a different number of wells (although the lower risk is accompanied with a lower NPV).

Decision Table dialogs

Specify Target (Step 1 Of 3) dialog

The Specify Target dialog lets you choose which forecast to target.

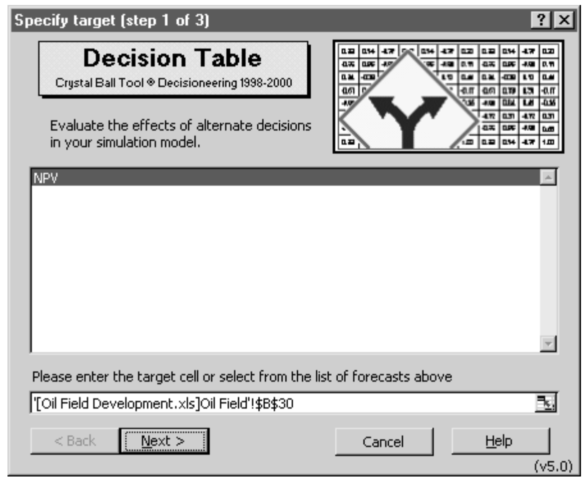


Figure 6.21 Specify Target dialog for Decision Table

The fields and buttons in this dialog are:

Forecast List — Lists all the forecast cells in all open spreadsheets. When you select a forecast from the list, its cell information automatically appears in the Enter Target Cell field.

The first forecast is selected by default.

Enter Target Cell — Describes the cell location of the selected forecast or formula. If you select a forecast from the list above, the cell information automatically appears in this field.

You can use this field to select a formula cell instead of a forecast.

Next — Opens the Select One Or Two Decisions dialog.

Select One Or Two Decisions (Step 2 Of 3) dialog

This dialog lets you select one or two decision variables to explore.

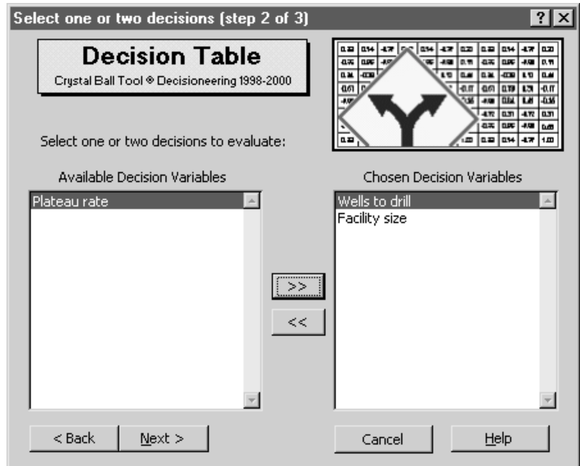


Figure 6.22 Select One Or Two Decisions dialog for Decision Table

The fields and buttons in this dialog are:

Available Decision Variables
: Lists all the defined decision variables in the open spreadsheets.

Chosen Decision Variables
: Lists one or two decision variables that the tool will test different values for.

\>> : Moves the selected decision variable in the Available Decision Variables list to the Chosen Decision Variables list.

<< : Moves the selected decision variable in the Chosen Decision Variables list to the Available Decision Variables list.

Back : Returns to the Specify Target dialog.

Next : Opens the Specify Options dialog.

Specify Options (Step 3 Of 3) dialog

The Specify Options dialog lets you select options to control how the tool works.

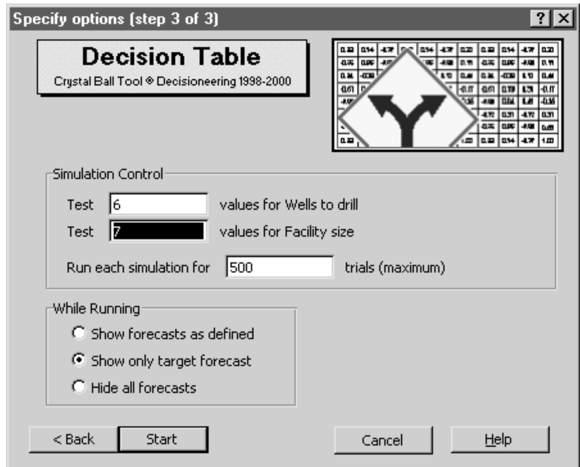


Figure 6.23 Specify Options dialog for Decision Table

The Simulation Control options in this dialog are:

Number Of Test Values For Each Decision Variable

Sets the number of values the tool will test. The tool distributes the number of values evenly across the defined decision variable range.

If you have one decision variable, the tool runs a simulation for each test value. For two decision variables, the tool runs a simulation for each combination of values, i.e., the product of the two numbers of test values.

The default is 10 values for continuous decision variables or for discrete decision variables with a step size greater than 40.

Trials Per Simulation

Sets the number of trials to run for each simulation.

The default is the number set in the Crystal Ball run preferences.

The While Running options are:

Show Forecasts As Defined
: Displays a forecast chart for each defined forecast during the simulation.

Show Only Target Forecast
: Displays only the forecast chart for the target forecast during the simulation.

Hide All Forecasts
: Displays no forecast charts during the simulation.

The buttons are:

Back	Returns to the Select One Or Two Decisions dialog.
Start	Runs the tool.

Tornado Chart tool

The Tornado Chart tool measures the impact of each model variable one at a time on a target forecast. The tool displays the results in two ways:

- Tornado chart
- Spider chart

This method differs from the correlation-based method built into Crystal Ball in that this tool tests each assumption, decision variable, precedent, or cell independently. While analyzing one variable, the tool freezes the other variables at their base values. This measures the effect each variable has on the forecast cell while removing the effects of the other variables. This method is also known as “one-at-a-time perturbation” or “parametric analysis”.

The Tornado Chart tool is useful for:

- Measuring the sensitivity of variables that you have defined in Crystal Ball.
- Quickly pre-screening the variables in your model to determine which ones are good candidates to define as assumptions or decision variables. You can do this by testing the precedent variables of any formula cell. See page 296 for more information on precedents.

Tornado chart

The tool tests the range of each variable at percentiles you specify and then calculates the value of the forecast at each point. The tornado chart illustrates the swing between the maximum and minimum forecast values for each variable, placing the variable that causes the largest swing at the top and the variable that causes the smallest swing at the bottom. The top variables have the most effect on the forecast, and the bottom variables have the least effect on the forecast.

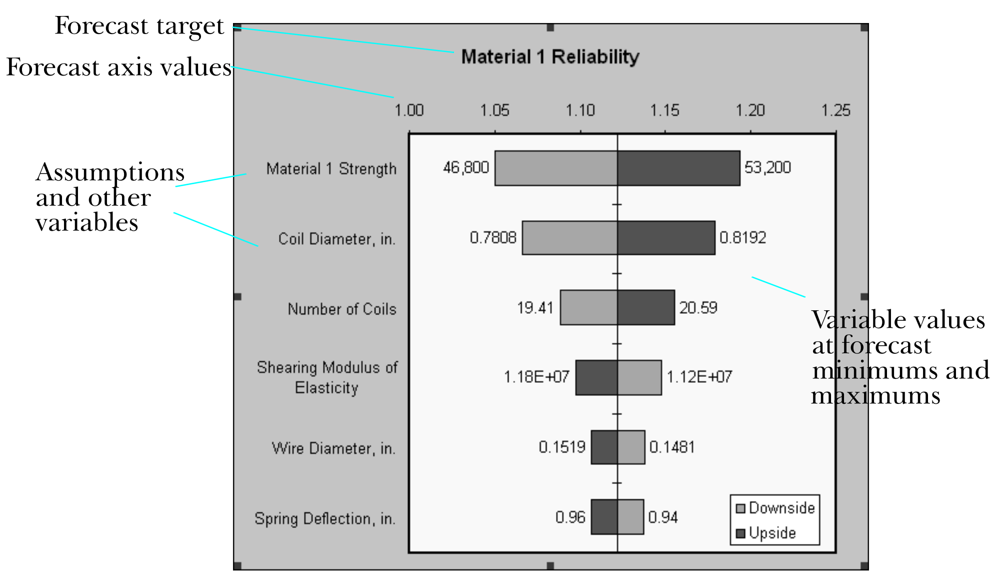


Figure 6.1 Tornado chart

The bars next to each variable represent the forecast value range across the variable tested, as discussed above. Next to the bars are the values of the variables that produced the greatest swing in the forecast values. The bar colors indicate the direction of the relationship between the variables and the forecast.

For variables that have a positive effect on the forecast, the upside of the variable (shown in blue) is to the right of the base case and the downside of the variable (shown in red) is to the left side of the base case. For variables that have a reverse relationship with the forecast, the bars are reversed.

When a variable's relationship with the forecast is not strictly increasing or decreasing, it is called non-monotonic. In other words, if the minimum or maximum values of the forecast range do not occur at the extreme endpoints of the testing range for the variable, the variable has a non-monotonic relationship with the forecast.

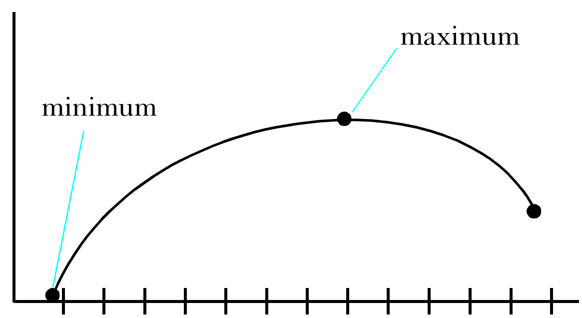


Figure 6.2 A non-monotonic variable

If one or more variables are non-monotonic, all the variable bars are the same color all the way across.

Spider chart

The spider chart illustrates the differences between the minimum and maximum forecast values by graphing a curve through all the variable values tested. Curves with steep slopes, positive or negative, indicate that those variables have a large effect on the forecast, while curves that are almost horizontal have little or no effect on the forecast. The slopes of the lines also indicate whether a positive change in the variable has a positive or negative effect on the forecast.

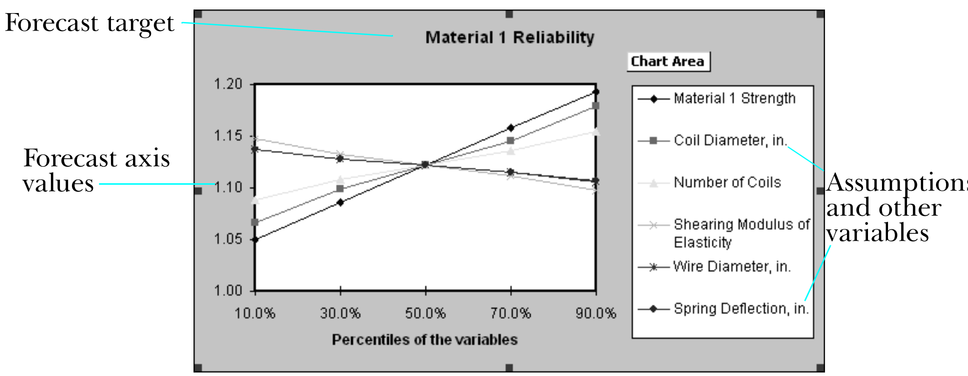


Figure 6.3 Spider chart

Crystal Ball Tool Note: *There is a maximum of 250 variables for these charts.*

Tornado Chart example

In the Crystal Ball Examples folder there is a Reliability.xls spreadsheet you can use to experiment with the Tornado Chart tool. This spreadsheet model predicts the reliability of a spring using three different construction materials.

To run Tornado Chart:

1. **In Excel with Crystal Ball loaded, open the spreadsheet Reliability.xls.**

   If you have any other worksheets open, close them first, because the tool gathers all Crystal Ball definitions from all open worksheets.

2. **Select CBTools > Tornado Chart.**

   The Specify Target (Step 1 Of 3) dialog appears. All of the forecasts from Reliable.xls appear in the list.

3. **Select the "Material 1 Reliability" forecast.**

4. **Click on Next.**

   The Specify Input Variables (Step 2 Of 3) dialog appears.

5. **Click on Add Assumptions.**

6. **Remove Material 2 Strength and Material 3 Strength.**

   a. **Select an assumption to remove.**

   b. **Click on Remove.**

   c. **Repeat steps 6a and 6b for the second assumption to remove.**

   The last two assumptions have no impact on the target forecast. If you leave them in the list, they will appear in the charts even though they are unrelated to the target forecast.

7. **Click on Next.**

   The Specify Options (Step 3 Of 3) dialog appears.

8. **Set the following options:**

   - Testing range is set to 10% to 90%
   - Testing points is set to 5
   - For Base Case is set to Use Existing Cell Values
   - Tornado Method is set to Percentiles Of The Variables
   - Both Tornado Chart and Spider Chart are selected
   - Show 20 Top Variables

9. **Click on Start.**

   The tool creates the tornado and spider charts on their own workbooks.

Interpreting the results

In this example, six assumptions are listed in the tornado chart. The first assumption, Material 1 Strength, has the highest sensitivity ranking and is the most important. A researcher running this model would investigate this assumption further in the hopes of reducing its uncertainty, and therefore its effect on the target forecast, Material 1 Reliability.

The last two assumptions, Wire Diameter and Spring Deflection, are the least influential assumptions. Since their effects on the Material 1 Reliability are very small, you might ignore their uncertainty or eliminate them from the spreadsheet.

Caveats

While tornado and spider charts are very useful, there are some caveats:

- Since the tool tests each variable independently of the others, the tool doesn't consider correlations defined between the variables.
- The results shown in the tornado and spider charts depend significantly on the particular base case used for the variables. To confirm the accuracy of the results, run the tool multiple times with different base cases.

  This characteristic makes the one-at-a-time perturbation method less robust than the correlation-based method built into Crystal Ball's sensitivity chart. Hence, the sensitivity chart is preferable, since it computes sensitivity by sampling the variables all together while a simulation is running.

Tornado Chart dialogs

Specify Target (Step 1 Of 3) dialog

The Specify Target dialog lets you choose which forecast or formula cell to target.

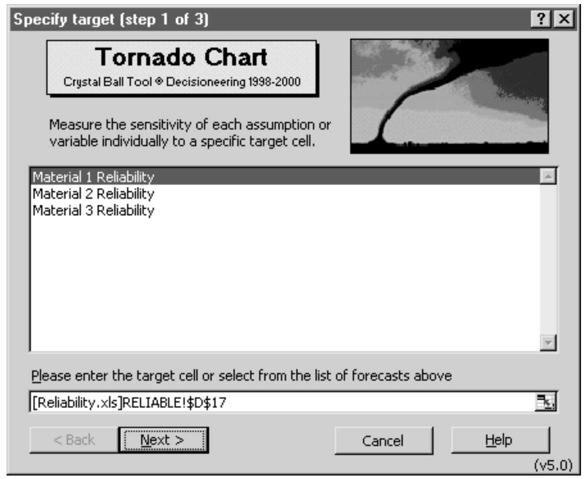


Figure 6.4 Specify Target dialog for Tornado Chart

The fields and buttons in this dialog are:

Forecast List — Lists all the forecast cells in all open spreadsheets. When you select a forecast from the list, its cell information automatically appears in the Enter Target Cell field.

The first forecast is selected by default.

Enter Target Cell — Describes the cell location of the selected forecast or formula. If you select a forecast from the list above, the cell information automatically appears in this field.

You can use this field to select a formula cell instead of a forecast.

Next — Opens the Specify Input Variables dialog.

Specify Input Variables (Step 2 Of 3) dialog

This dialog lets you select the assumptions, decision variables, and precedents to include in the tornado and spider charts.

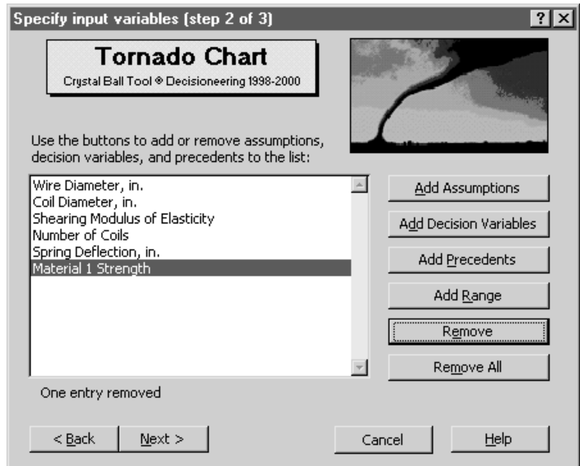


Figure 6.5 Specify Input Variables dialog for Tornado Chart

You can include any value cell in your tornado chart calculations. However, the cells are usually:

assumptions
: Cells defined as assumptions in Crystal Ball.

  For more information about assumptions, see “Defining assumptions” on page 131.

decision variables
: Cells defined as decision variables in Crystal Ball.

  For more information about defining decision variables, see “Defining decision variables” on page 151.

precedents
: All cells *within the active spreadsheet* that are referenced as part of the formula or a sub-formula of the target cell.

Crystal Ball Tool Note: *Precedents for this tool are handled differently than standard Excel precedents in that they cannot trace beyond the active* *spreadsheet. Therefore, you can only use precedents on the active spreadsheet as input variables.*

The field and buttons in this dialog are:

Input Variable List
: Lists all the variables selected for the tornado and spider charts.

Add Assumptions
: Adds all assumptions from all open worksheets to the input variable list.

Add Decision Variables
: Adds all decision variables from all open worksheets to the input variable list.

Add Precedents
: Adds all precedents of the target cell from all open worksheets to the input variable list.

Add Range
: Lets you select a range of cells from the open worksheet to add to the input variable list. If you click this button, an Input dialog appears asking you to enter a cell range or to select a range of cells from the spreadsheet. You must click on OK to accept the selected range.

Remove
: Removes the selected variable from the input variable list.

Remove All
: Removes all of the items from the input variable list.

Back
: Returns to the Specify Target dialog.

Next
: Opens the Specify Options dialog.

Specify Options (Step 3 Of 3) dialog

This dialog lets you set options that control the tool. The groups of options in this dialog are:

- Tornado Input
- Tornado Method
- Tornado Output

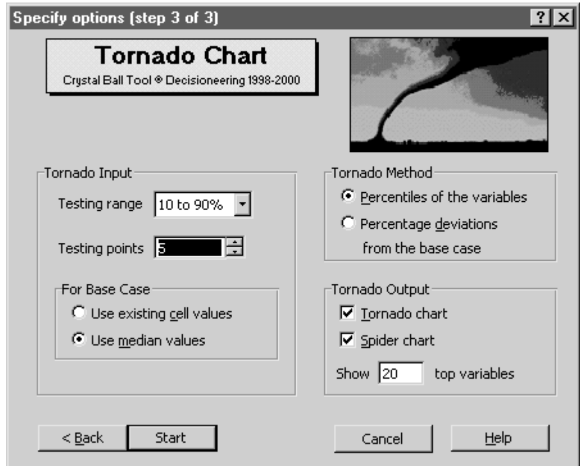


Figure 6.6 Specify Options dialog for Tornado Chart

The buttons in this dialog are:

Back	Returns to the Specify Input Variables dialog.
Start	Starts the tool, generating the tornado and spider charts.

Tornado Input options

The Tornado Input options are:

Testing Range	Defines the range in which the tool samples the variables. The choices are either: the percentile range (if the tornado method is Percentiles Of The Variables) or the percent from the base case (if the tornado method is Percentage Deviations From The Base Case). The default is 10% to 90% percentiles or -10% to 10% deviation.
Testing Points	Defines how many values to test in the testing range. The testing points are evenly distributed across the testing range. Testing more points than just endpoints detects non-monotonic variable relationships.

The default is five testing points.

For Base Case — Specifies whether to define the base case as either the existing cell values or the median values of the variables. If the tornado is based on a percentage deviation, only the cell values option is available.

The default is the median cell values.

Tornado Method options

The Tornado Method options in this dialog are:

Percentiles Of The Variables — Indicates that the tool should test the variables using percentiles of the assumption distributions or percentiles of the decision variable ranges.

This is the default.

Percentage Deviations From The Base Case — Indicates that the tool should test the variables using small changes that are specified percentages away from the base case.

This is the only option available if you selected variables other than assumptions or decision variables.

Crystal Ball Tool Note: *The tool treats discrete decision variables as continuous for this second method.*

Tornado Output options

The Tornado Output options are:

Tornado Chart — Generates a tornado chart showing the sensitivity of the variables using range bars.

Spider Chart — Generates a spider chart showing the sensitivity of the variables using sloping curves.

Show Top Variables — Indicates how many variables to include in the tornado charts if there are a lot of variables. The charts can only clearly show about 20 variables.

Two-dimensional Simulation tool

Risk analysts must often consider two sources of variation in their models:

uncertainty — Assumptions that are uncertain because you have insufficient information about a true, but unknown, value. Examples of uncertainty include the reserve size of an oil field and the prime interest rate in 12 months. You can describe an uncertainty assumption with a probability distribution.

Theoretically, you can eliminate uncertainty by gathering more information. Practically, information can be missing because you haven't gathered it or because it is too costly to gather.

variability — Assumptions that change because they describe a population with different values. Examples of variability include the individual body weights in a population or the daily number of products sold over a year. You can describe a variability assumption with a frequency distribution (or approximate it with a probability distribution).

Variability is inherent in the system, and you cannot eliminate it by gathering more information.

For many types of risk assessments, it is important to clearly distinguish between uncertainty and variability.[2] Separating these concepts in a simulation lets you more accurately detect the variation in a forecast due to lack of knowledge and the variation caused by natural variability in a measurement or population. In the same way that a one-dimensional simulation is generally better than single-point estimates for showing the true probability of risk, a two-dimensional simulation is generally better than a one-dimensional simulation for characterizing risk.

2. Hoffman, F. O. and J. S. Hammonds. "Propagation of uncertainty in risk assessments: The need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability," *Risk Analysis*, vol. 14, no. 5. pp 707-712, 1994.

The Two-dimensional Simulation tool runs an outer loop to simulate the uncertainty values, and then freezes the uncertainty values while it runs an inner loop (of the whole model) to simulate the variability. This process repeats for some small number of outer simulations, providing a portrait of how the forecast distribution varies due to the uncertainty.

The primary output of this process is a chart depicting a series of cumulative frequency distributions. You can interpret this chart as the range of possible risk curves associated with a population.

Crystal Ball Tool Note: *When using this tool, set the Seed Value option in the Crystal Ball Run Preferences dialog so that the resulting simulations are more comparable.*

Two-dimensional Simulation example

In the Crystal Ball Examples folder there is a Toxic Waste Site.xls spreadsheet you can use to experiment with the Two-dimensional Simulation tool. This spreadsheet model predicts the cancer risk to the population from a toxic waste site. This spreadsheet has two variability assumptions and two uncertainty assumptions.

To run the Two-dimensional Simulation tool:

1. **In Excel with Crystal Ball loaded, open the spreadsheet Toxic Waste Site.xls.**

2. **In the Run > Run Preferences > Sampling dialog, set:**
   - Random Number Generation to use the Same Sequence Of Random Numbers and An Initial Seed Value of 999
   - Monte Carlo simulation

   When using this tool, use these options to make the resulting simulations comparable.
3. **Select CBTools > 2D Simulation.**

   The Specify Target dialog appears.
4. **Select the Risk Assessment forecast.**
5. **Click on Next.**

   The Specify Assumptions dialog appears.

6. **Move Body Weight and Volume Of Water Per Day to the Variability list.**

   a. **Select Body Weight.**

   b. **Click on >>.**

   c. **Repeat steps 5a and 5b for Volume Of Water Per Day.**

   This separates the assumptions into the two types: uncertainty and variability.

7. **Click on Next.**

   The Specify Options dialog appears.

8. **Set the following options.**

   - Outer simulation runs set to 100
   - Inner simulation runs set to 1,000
   - Show only target forecast

9. **Click on Start.**

   The simulations start. The tool first single-steps one trial to generate a new set of values for the uncertainty assumptions. Then it freezes these assumptions and runs a simulation for the variability assumptions in the inner loop.

   The tool retrieves the Crystal Ball forecast information after each inner loop runs. The tool then resets the simulation and repeats the process until the outer loop has run for the specified number of simulations.

Interpreting the results

The results of the simulations appear in a table containing the forecast means, the uncertainty assumption values, and the statistics (including percentiles) of the forecast distribution for each simulation.

	A	B	C	D	E	F	G	H	I	J	K
1		Risk Assessment(1)	Risk Assessment(2)	Risk Assessment(3)	Risk Assessment(4)	Risk Assessment(5)	Risk Assessment(6)	Risk Assessment(7)	Risk Assessment(8)	Risk Assessment(9)	Risk Assessment(10)
2		3.94E-05	5.27E-05	5.29E-05	5.67E-05	5.88E-05	5.98E-05	6.09E-05	6.22E-05	6.31E-05	6.36E-0
3	**Assumptions:**										
4	Concentration of Contaminant in Water	101.24	92.78	111.69	116.55	102.70	102.96	93.57	115.01	103.43	110.4
5	CPF	1.3E-02	1.9E-02	1.6E-02	1.6E-02	1.9E-02	1.9E-02	2.2E-02	1.8E-02	2.0E-02	1.9E-0
6											
7	**Statistics:**										
8	Mean	3.94E-05	5.27E-05	5.29E-05	5.67E-05	5.88E-05	5.98E-05	6.09E-05	6.22E-05	6.31E-05	6.36E-0
9	Median	3.78E-05	5.05E-05	5.06E-05	5.43E-05	5.63E-05	5.73E-05	5.84E-05	5.96E-05	6.05E-05	6.09E-0
10	Standard Deviation	1.91E-05	2.55E-05	2.56E-05	2.75E-05	2.85E-05	2.90E-05	2.95E-05	3.02E-05	3.06E-05	3.08E-0
11	Variance	3.65E-10	6.52E-10	6.55E-10	7.55E-10	8.10E-10	8.40E-10	8.71E-10	9.09E-10	9.35E-10	9.49E-1
12	Skewness	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.4
13	Kurtosis	3.13	3.13	3.13	3.13	3.13	3.13	3.13	3.13	3.13	3.1
14	Coeff. of Variability	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.4
15	Minimum	4.78E-07	6.39E-07	6.41E-07	6.88E-07	7.12E-07	7.25E-07	7.39E-07	7.55E-07	7.65E-07	7.71E-0
16	Maximum	1.16E-04	1.55E-04	1.56E-04	1.67E-04	1.73E-04	1.76E-04	1.79E-04	1.83E-04	1.86E-04	1.87E-0
17	Range	1.16E-04	1.55E-04	1.55E-04	1.66E-04	1.72E-04	1.75E-04	1.79E-04	1.83E-04	1.85E-04	1.87E-0
18											
19	**Percentiles:**										
20	5%-tile	1.04E-05	1.39E-05	1.40E-05	1.50E-05	1.55E-05	1.58E-05	1.61E-05	1.65E-05	1.67E-05	1.68E-0
21	10%-tile	1.56E-05	2.08E-05	2.09E-05	2.24E-05	2.32E-05	2.36E-05	2.41E-05	2.46E-05	2.49E-05	2.51E-0

Figure 6.7 Two-dimensional Simulation results table

The tool also graphs the results of the two-dimensional simulations on an overlay chart and a trend chart. The overlay chart shows the risk curves for the simulations for different sets of uncertainty assumption values.

In the chart below, most of the risk curves are clustered densely toward the center while a few outlier curves are scattered to the right, showing the small probability of having a much greater risk.

Statistical Note: *In risk analysis literature, the curves are often called the alternate realizations of the population risk assessment.*

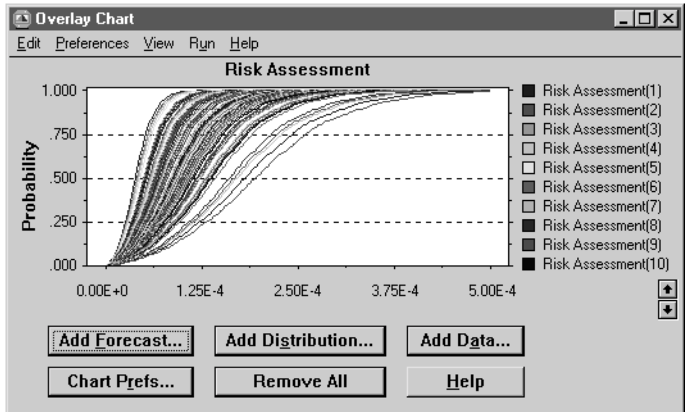


Figure 6.8 Overlay chart of risk curves

Another helpful output is a trend chart depicting certainty bands for the percentiles of the risk curves. The band width shows the amount of uncertainty at each percentile level for all the distributions.

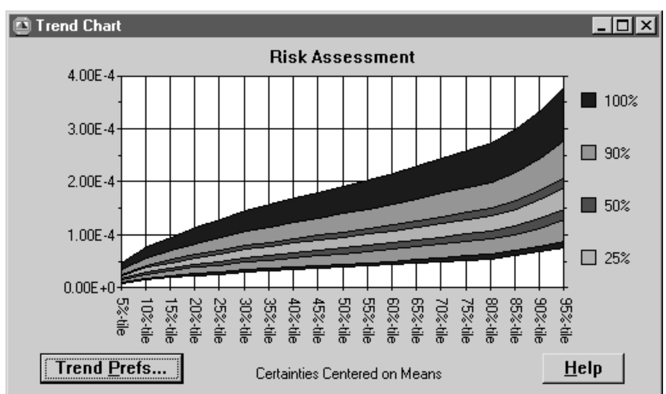


Figure 6.9 Trend chart of certainty bands

You can focus in on a particular percentile level, such as the 95th percentile, by viewing the statistics of the 95th percentile forecast.

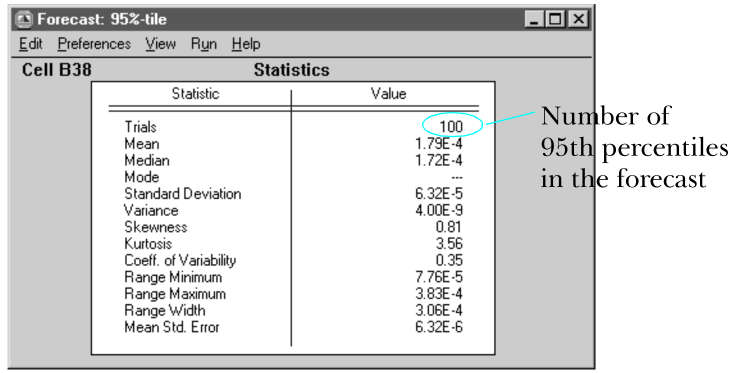


Figure 6.10 95th percentiles forecast statistics

Compare the results of the two-dimensional simulation to a one-dimensional simulation (with both uncertainty and variability co-mingling together) of the same risk model. The mean of the 95th percentiles, 1.79E-4, is lower than the 95th percentile risk of the one-dimensional simulation shown below at 2.03E-4. This indicates the tendency of the one-dimensional simulation results to overestimate the population risk, especially for highly skewed distributions.

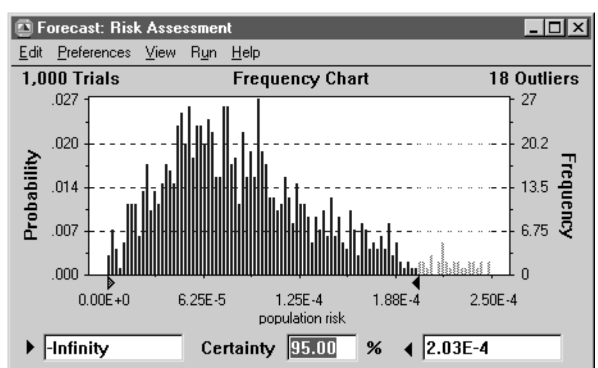


Figure 6.11 Forecast chart for one-dimensional simulation

Second-order assumptions

Some assumptions contain elements of both uncertainty and variability. For instance, an assumption might describe the distribution of body weights in a population, but the parameters of the distribution might be uncertain. These types of assumptions are called second-order assumptions (also, second-order random variables).[3] You can model these types of assumptions in Crystal Ball by placing the uncertain parameters of the distribution in separate cells and defining these cells as assumptions. You then link the parameters of the variability assumption to the uncertainty assumptions using cell references.

To illustrate this for the Toxic Waste Site.xls spreadsheet:

1. **Enter the values 70 and 10 into cells G4 and G5, respectively.**

   These are the mean and standard deviation of the Body Weight assumption in cell C4, which is defined as a normal distribution.

2. **Define an assumption for cell G4 using a normal distribution with a mean of 70 and a standard deviation of 2.**

3. **Define an assumption for cell G5 using a normal distribution with a mean of 10 and a standard deviation of 1.**

4. **Enter references to these cells in the Body Weight assumption.**

   When you run the tool for second-order assumptions, the uncertainty of the assumptions' parameters is modeled in the outer simulation, and the distribution of the assumption itself is modeled (for different sets of parameters) in the inner simulation.

---

3. Burmaster, David E., and Andrew M. Wilson. "An Introduction to Second-Order Random Variables in Human Health Risk Assessments," *Human and Ecological Risk Assessment: vol. 2, no. 4*. pp 892-919, 1996.

Crystal Ball Note: *Often, the parameters of assumptions are correlated. For example, you would correlate a higher mean with a higher standard deviation or a lower mean with a lower standard deviation. Defining correlation coefficients between parameter distributions can increase the accuracy of your two-dimensional simulation. When data is available, as in sample body weights of a population, you can use the Bootstrap tool to estimate the sampling distributions of the parameters and the correlations between them.*

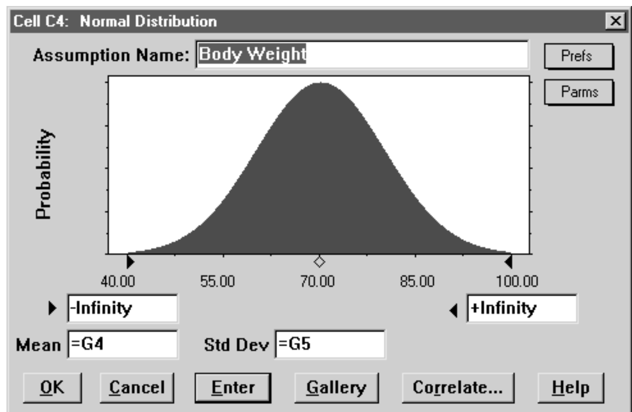


Figure 6.12 Assumption using cell references for the mean and standard deviation

Two-dimensional Simulation dialogs

Specify Target (Step 1 Of 3) dialog

The Specify Target dialog lets you choose which forecast to target.

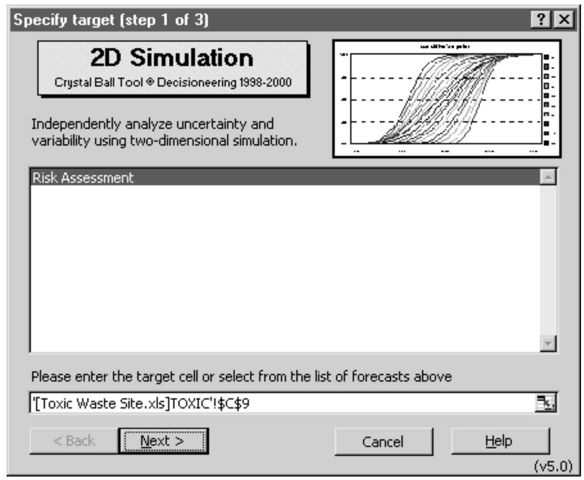


Figure 6.13 Specify Target dialog for Two-dimensional Simulation

The fields and buttons in this dialog are:

Forecast List — Lists all the forecast cells in all open spreadsheets. When you select a forecast from the list, its cell information automatically appears in the Enter Target Cell field.

The first forecast is selected by default.

Enter Target Cell — Describes the cell location of the selected forecast or formula. If you select a forecast from the list above, the cell information automatically appears in this field.

You can use this field to select a formula cell instead of a forecast.

Next — Opens the Specify Assumptions dialog.

Specify Assumptions (Step 2 Of 3) dialog

This dialog separates the assumptions into uncertainty assumptions and variability assumptions. All the assumptions from all open worksheets start in the Uncertainty list by default. You must have at least one assumption of each type.

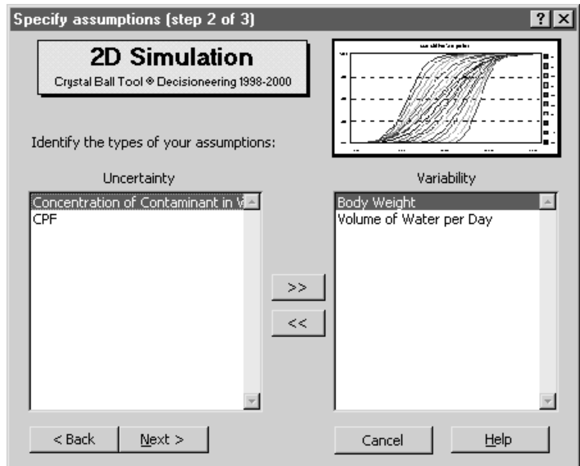


Figure 6.14 Specify Assumptions dialog for Two-dimensional Simulation

When you save your spreadsheet model, the tool remembers the assumption types for the next time you run the tool.

The fields and buttons in this dialog are:

>>	Moves any selected assumptions in the Uncertainty list to the Variability list.
<<	Moves any selected assumptions in the Variability list to the Uncertainty list.
Back	Returns to the Specify Target dialog.
Next	Opens the Specify Options dialog.

Specify Options (Step 3 Of 3) dialog

This dialog lets you set the options that control how the tool runs.

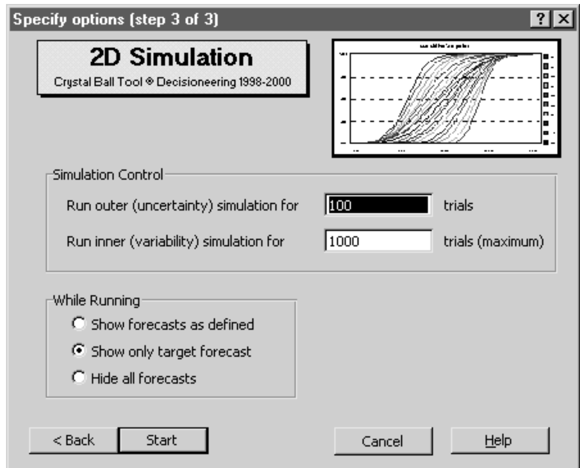


Figure 6.15 Specify Options dialog for Two-dimensional Simulation

The Simulation Control options in this dialog are:

Outer simulation runs
: Sets the number of times the tool simulates the uncertain assumptions in the outer loop. These values are then frozen during the inner simulations.

  The default is 10.

Inner simulation runs
: Sets the number of times the tool simulates the variability assumptions in the inner loop.

  The default is the number of trials set in the Crystal Ball run preferences.

The While Running options are:

Show Forecasts As Defined
: Displays a forecast chart for each defined forecast during the simulation.

Show Only Target Forecast
: Displays only the forecast chart for the target forecast during the simulation.

Hide All Forecasts
: Displays no forecast charts during the simulation.

The buttons are:

Back	Returns to the Specify Assumptions dialog.
Start	Runs the tool.

Appendix A

Error Recovery

In this appendix

This appendix lists, alphabetically by first word, messages that might appear when you are entering or selecting data, or running simulations. Either an explanation or a solution or both are provided for each message. Messages not listed in this appendix should be self-explanatory.

Crystal Ball error messages

Below is a list of the error messages for Crystal Ball. For error messages for OptQuest, the extenders, or CB Turbo, see the accompanying manual for the product.

Message	***Reason***	***Solution***
A calculation error has occurred in cell <A1>. You may wish to check your formulas and assumptions for division by 0, etc.	You have a spreadsheet calculation error.	Check your spreadsheet formulas for errors such as: square root of a negative number, division by zero, too large a number (>1.0E+307), a bad table lookup, or an invalid argument given to a trig, log, power, or financial function. (Financial functions with a guessed argument may not be converging due to a poor guess.)
Alpha plus Beta plus 1 must be less than 1750.	The sum of alpha and beta parameters has an upper limit of 1,749. This results from a restriction on the size of a number that can be factorialized.	Change the value of alpha and beta so that their sum is less than 1,749.
An assumption cannot be correlated to itself.	You entered the cell for the current assumption in the Correlate dialog. You must select an assumption cell other than the current cell to correlate with the current assumption.	Enter the cell of a different assumption or select one of the active assumptions listed under the Select Assumption button.
Assumptions and decision variables can't be defined in the same cell.	You are trying to define an assumption in a cell that already has a decision variable, or you are trying to define a decision variable in a cell with an assumption.	You must clear the assumption or decision variable data (using Cell > Clear Data) before defining the cell with a decision variable or assumption.
Assumptions can only be defined for simple value cells.	You cannot define an assumption in an empty cell or a cell with text, a formula, or a function.	You cannot define an assumption in an empty cell or a cell with text, a formula, or a function.
Cannot complete operation due to an unexpected error.	Unknown.	Contact Decisioneering technical support at www.decisioneering.com.

Message	*Reason*	*Solution*
Cannot construct distribution with the parameters given.	One or more distribution parameters are invalid.	Make sure all your parameters have an absolute value less than 1.03 E+307. Change the highlighted parameter value to a valid value.
Can't correlate assumptions that have dynamic cell references.	The cell you are trying to correlate has dynamic cell references. You cannot correlate cells with dynamic cell references.	Either change the dynamic cell references to static, remove the cell references, or correlate to a different assumption.
Can't find any spreadsheets associated with the saved run file.	Saved run files are associated with their spreadsheets. You must have the associated spreadsheet open to restore a run.	Open the associated spreadsheet and try to restore the saved run again.
Can't find main spreadsheet window.	Crystal Ball cannot find the Excel window.	Close Crystal Ball and Excel and restart the program.
Can't find or open data file. Please verify that the correct path and file name are given.	The text file you specified doesn't exist, or you didn't enter a file name.	Check the spelling and folder location of the text file and enter a valid file name.
Cell range must be in 'A1:B2' or 'A1..B2' format.	You did not properly enter a range in a field expecting a cell range.	Change the range to one of the specified formats.
Cell reference is outside the active area or refers to the wrong cell type.	A cell reference is referencing a cell with text or an empty cell.	Either enter a cell reference that has a valid value or enter a valid value in the referenced cell.
Cell references may not be entered for this distribution	You cannot use cell references with the custom distribution.	Replace cell references in the custom distribution parameter fields with values.
Cell reference must be in 'A1' format.	Crystal Ball expected a cell reference, but the reference wasn't in the correct format.	If you want to indicate a cell reference, make sure you are using the cell reference format indicated by "Entering cell references" on page 135.
Cell reference refers to a deleted cell or the wrong cell type.	The referenced cell isn't the correct cell type.	Make sure the referenced cell is the correct cell type.

Message	*Reason*	*Solution*
Certainty must fall between 0-100%	You entered an invalid value in the Certainty field. The certainty must be a percentage between 0 and 100%, inclusive.	Change your certainty value to a percentage between 0 and 100%.
Correlation Coefficient must be between one and minus one.	All correlation coefficients in Crystal Ball must be expressed as a number between minus one and plus one, inclusive.	Change your correlation coefficient to a number between -1 and 1, inclusive.
Crystal Ball cannot work with locked cells or sheets.	You cannot have locked cells or worksheets when using Crystal Ball. Or, if you have merged cells on one worksheet, Excel sometimes interprets the same cells on other worksheets of the same workbook as being locked.	Unlock all cells and worksheets. Also, check for merged cells that correspond to assumption or forecast cells on other worksheets. Unmerge those cells.
Crystal Ball found an unsaved sheet in the simulation. Please save it now.	Crystal Ball cannot save a run until its associated workbook is named and saved.	Save your current workbook and then try to save the run again.
Data contains non-positive numbers—some distributions were skipped.	Lognormal, beta, Pareto, exponential, and gamma distributions can only contain positive values. Crystal Ball eliminates these distributions from the distribution fit if the data contains negative values.	No action is necessary.
Data does not appear to be in cumulative format.	The values in the custom distribution data range are not in the right format.	Format the data in two columns: 1) the first column with data values, 2) the second column with cumulative probabilities of the data values. Both columns must be in ascending order.
Decision variable name cannot contain + - * / ^ , < = > characters.	The decision variable name you specified contains illegal special characters.	Remove special characters from your decision variable name.
Decision variables can only be defined for simple value cells.	You cannot define a decision variable in an empty cell or a cell with text, a formula, or a function.	Change the cell to contain a simple value before defining your decision variable.

Message	Reason	Solution
Forecasts can only be defined for formula or value cells.	You cannot define a forecast in an empty cell or a cell with text.	Change the cell to contain a formula, function, or simple value before defining your forecast.
Geometric Standard Deviation must be greater than one.	You entered a geometric standard deviation that is less than or equal to one.	Change the geometric standard deviation to a number greater than one.
Input data must contain at least two distinct values.	The specified range of values to fit contains less than two distinct values to fit.	Enter a valid range of data to fit that contains at least two distinct values.
Likeliest value must fall between Min and Max.	Your likeliest parameter must be between the minimum and maximum parameter values.	Change your likeliest value to fall between the minimum and maximum values or change the minimum or maximum values to define a range that includes the likeliest value.
Line or dot width must be less than 16.	The values associated with the line and dot chart type must be less than 16.	Change the values in the Line and Dot fields to be less than 16.
Location must be between zero and 1000.	The location parameter value is outside the accepted range.	Enter a new value for the location parameter that is between zero and 1000.
Lognormal mean must be less than 11250.	The log mean parameter for the lognormal distribution has an upper limit of 11,250 due to a restriction on the size of a number that can be exponentiated.	Change the lognormal log mean to a number less than 11,250.
Lower percentile must be greater than upper percentile.	The lower percentile value is less than the higher percentile value.	Either change the lower percentile value to be above the higher percentile or change the higher percentile value to be below the lower percentile value. If you want the upper percentile to be greater than the lower percentile, in the Forecast window, select Preferences > Statistics and change the option to show percentiles as the Probability Below A Value.

Message	*Reason*	*Solution*
Maximum number of trials reached.	The number of trials has reached the maximum number you specified in the Run Preferences dialog.	To keep running the simulation, increase the maximum number of trials specified in the Run Preferences dialog.
Maximum report length reached.	You are generating a report that is reaching the maximum report length.	Report your data in smaller sets of data, e.g., only report assumption information and then, separately report forecast information.
Mean or median does not lie within the display range.	You specified a confidence interval that doesn't have the mean or median in the display range.	No action required. Just make sure the certainty is displaying what you expected.
Memory limit reached.	You are running a simulation with a large or complicated spreadsheet and your computer doesn't have enough RAM.	• Close other applications • Turn on virtual memory • Consult your *Windows User's Guide* for information on optimizing your memory • If this is a recurring problem, install more RAM in your computer
Min must be less than Max.	Your minimum value is greater than your maximum value.	Change your minimum to be less than your maximum or your maximum to be greater than your minimum.
Must be a whole number.	The highlighted number must be a positive integer, for example, 1, 2, or 3.	Replace the highlighted number with a positive integer.
No assumptions chosen.	You selected the Chosen... option but didn't select any assumptions.	Select at least one assumption using the All or Chosen option.
No cells have been selected.	You are trying to paste Crystal Ball data, but there are no cells selected. You might have a graphic item or text box selected.	Select the cell or cells where you want to paste data, and then paste.

Message	Reason	Solution
No distributions chosen.	You selected to fit to chosen distributions, but didn't select any distributions from the Choose Distributions dialog.	Make sure you select at least one distribution. Click on the Chosen option to open the Choose Distribution dialog and select distributions.
No distributions tested fit the data.	Either you selected to fit your data to a discrete distribution (which Crystal Ball doesn't support) or only to distributions that can't use negative numbers (lognormal, beta, Pareto, exponential, and gamma) when your data contains negative numbers.	Select only continuous distributions, and, if your data contains negative values, select at least one continuous distribution that can use negative numbers.
No forecast cells have been defined.	You must define at least one forecast to run a simulation.	Go back to your spreadsheet and define at least one cell as a forecast.
No forecasts chosen.	You selected the Chosen... option but didn't select any forecasts.	Select at least one forecast using the All or Chosen option.
No worksheet active.	You must have a workbook open to run a simulation.	Open a workbook and try to run a simulation again.
Not a valid date—must be in mm/dd/yy format.	You entered a parameter value using the special character /. Crystal Ball is trying to interpret the number as a date, but the values are invalid for the format mm/dd/yy.	Either remove the / character from your parameter or make sure your date is in the proper format.
Not a valid number.	You entered a non-numeric character in a field that requires a number.	Make sure all numeric fields only contain numbers or cell references to numbers.
Not enough memory available to complete the operation.	You are running an operation other than simulation, and your computer does not have enough RAM.	• Close any open forecast windows • Set the Maximum Number Of Trials in the Run Preferences dialog to a lower number and rerun the simulation • Try solutions for the message: Memory limit reached • If this is a recurring problem, install more RAM

Message	***Reason***	***Solution***
Note: this option may significantly alter the results of the simulation.	Selecting the option to turn off correlations can change results.	No action is required.
Note: this option may significantly increase the amount of memory required to run a simulation.	Selecting the Latin Hypercube simulation method uses more memory than Monte Carlo.	No action is required. However, if your memory is low, you might want to consider solving any memory problems before using Latin Hypercube.
Note: this option may significantly increase the time and memory required to run a simulation.	Selecting the Calculate Sensitivity option uses more memory than running a simulation without this option.	No action is required. However, if your memory is low or your model particularly large, you might want to consider solving any memory problems before using this option.
Number is too large.	You must use numbers between 5 E-324 and 1.7 E308 (positive numbers) or between -5 E-324 to -7 E308 (negative numbers).	Change your value to be in one of the two valid ranges.
Number must be greater than zero.	The highlighted number cannot be a negative value.	Replace the negative value with a positive value.
Number must not be infinity.	You entered infinity for a field that requires a number with an absolute value less than 1.03E307.	Enter a number with an absolute value less than 1.03E307.
Number of trials must be less than or equal to 1750.	The trials parameter has an upper limit of 1,750 due to a restriction on the size of a number than can be factorialized.	Enter a number of trials less than 1,750. In some cases, you might be able to use the normal distribution to approximate the binomial or hypergeometric distributions when the number of trials is large.
Number of trials must not exceed population size.	The trials parameter cannot be greater than the population parameter.	Either change the trials parameter to be less than the population or increase the population to be greater than the trials.
Number of values entered exceeds capacity of custom distribution.	The custom distribution has a capacity of approximately 1,400 single values or 450 continuous and discrete ranges.	Divide your assumption into at least two different custom assumptions with less than the maximum number of values or ranges.

Message	Reason	Solution
One or more values resulting from calculation errors had to be discarded. You may wish to check your formulas and assumptions for division by 0, etc.	A forecast had some calculation errors during the simulation, and you didn't have your run preferences set to stop the simulation on a calculation error.	Check your spreadsheet formulas for errors such as: • Square root of a negative number • Division by 0 • Too large a number $(\geq 1.03E+307)$ • Bad table lookup • Invalid argument given to a trig, log, power, or financial function. (Financial functions with a poor guess argument might not be converging.)
Percentage must fall between 0-100%.	You entered a value that was not between 0 and 100.	Change the percentile to a number between 0 and 100.
Population size must be less than 1750.	The population parameter has an upper limit of 1,750 due to a restriction on the size of a number than can be factorialized.	Enter a population size less than 1,750. In some cases, you might be able to use the normal distribution to approximate the binomial or hypergeometric distributions when the population size is large.
Probability must be greater than zero.	You must express all probabilities in Crystal Ball as fractions between 0 and 1, or as percents (e.g., 8%) between 0% and 100%.	Enter a probability between 0 and 1 or as a percent between 0% and 100%.
Probability must be less than one.	You must express all probabilities in Crystal Ball as fractions between 0 and 1, or as percents (e.g., 8%) between 0% and 100%.	Enter a probability between 0 and 1 or as a percent between 0% and 100%.
Rate must be greater than zero.	You entered a negative value or zero for the rate, which must be positive.	Change the rate value to a positive number.

Message	***Reason***	***Solution***
Rate must be less than or equal to 1500.	The rate parameter has an upper limit of 1,500 due to a restriction on the size of a number that can be raised to a power.	Enter a rate less than 1500. In some cases, you might be able to use the normal distribution to approximate the Poisson distribution when the rate is large.
Sensitivity value must be between zero and one.	You entered a sensitivity value to use as a cut-off that was not between zero and one.	Change the value in the Display Only Sensitivities Greater Than field to a number between 0 and 1.
Shape divided by Probability must be less than 1750.	The trials parameter has an upper limit of 1,750 due to a restriction on the size of a number than can be factorialized.	Enter either a smaller shape or a larger probability until the shape divided by the probability is less than 1750. In some cases, you might be able to use the normal distribution to approximate the negative binomial distribution: when the shape is large or the probability is small.
Shape must be less than 1500.	You entered a number greater than 1500 for the shape parameter.	Change the shape parameter value to a number less than or equal to 1500.
Sorry, that file is not a Crystal Ball saved run.	You tried to open a files with a .RUN extension that is not a saved Crystal Ball run.	Check the file to make sure that the file is a saved Crystal Ball .RUN file.
Standard deviation must be greater than zero.	You entered zero or a negative number for the standard deviation.	Change the standard deviation value to a positive number.
Step must be positive.	The custom distribution Step parameter must be a positive number.	Change the Step parameter to a positive value.
Step must be zero or positive.	The step size for a decision variable cannot be negative.	Replace the step size with a non-negative number.
System resources are low. Try closing windows or quitting other applications.	You have less than 8% of your system resources available.	Try rebooting your computer. If you start Crystal Ball right away and still have this problems, try removing programs from automatically starting and reboot again.

Message	Reason	Solution
The amount of data requested exceeds 256 columns. The data sheet will be incomplete.	You are extracting more than 256 columns worth of forecast data. Excel has a limit of 256 columns. The data will be truncated.	Extract data in smaller chunks by choosing a smaller number of forecasts to extract data for at one time.
The correlation coefficients are inconsistent. Crystal Ball can adjust the coefficients by an average of <value> (worst case of <value2>) to allow the simulation to continue.	You have specified a situation that cannot exist or is beyond Crystal Ball's capability. Correlation coefficients cannot be specified without regard to the total set of correlations. Contradictions arise when assumptions are related to each other by large positive and negative correlation coefficients. Also, as the number of correlated assumptions increases, the ability to use large negative correlation coefficients decreases.	Either: • Let Crystal Ball adjust the coefficients just for this simulation • Let Crystal Ball adjust the coefficients permanently • Click on cancel and recheck your correlation coefficients. In particular, try to make large correlation coefficients smaller. Or try to restructure your spreadsheet so that assumptions with large correlation coefficients can be calculated using formulas instead.
The Crystal Ball greetings file cannot be found or has been corrupted. Cannot continue.	The Crystal Ball greetings file is missing or damaged.	Reinstall Crystal Ball.
The Crystal Ball license file cannot be found or has been corrupted. Cannot continue.	The Crystal Ball license file is missing or damaged.	Reinstall Crystal Ball.
The number of values in both cell ranges must be the same.	To calculate the correlation coefficient, you must enter two different cell ranges with the same number of cells in each range.	Re-enter cell ranges that have the same number of cells.
The sample size must be greater than or equal to 10.	You entered a sample size in the Run Preferences dialog that is less than 10.	Change the sample size value to a number greater than or equal to 10.
The sample size must be less than or equal to 5000.	You entered a sample size in the Run Preferences dialog that is greater than 5000.	Change the sample size value to a number less than or equal to 5000.

Message	***Reason***	***Solution***
There was an unrecoverable error while restoring the file.	Crystal Ball could not restore the run because it didn't have access to the run file or the file was damaged.	Make sure you have read permission for the file and folder. If the file is damaged or corrupted, look for a backup copy.
There was an unrecoverable error while saving the file.	Crystal Ball could not restore the run because it didn't have access to the run file or the file was damaged.	Make sure you have read and write permission for the file and folder. If the file is damaged or corrupted, look for a backup copy.
There was an unrecoverable error while specifying the file.	Crystal Ball could not restore the run because it didn't have access to the run file or the file was damaged.	Make sure you have read permission for the file and folder. If the file is damaged or corrupted, look for a backup copy.
Too many cells selected (greater than 256). Try again with a smaller selection.	You can only copy 256 assumptions or forecasts at a time.	Select less than 256 cells to copy.
Unable to find user macro <macro_name>. Make sure the macro name is correct in the Run Preferences dialog.	Crystal Ball could not find the macro specified in the Run Preferences > Macros dialog.	Make sure the macro name is spelled correctly and is in the proper format. See "Macros preferences" on page 168 for the correct format.
Unable to generate assumption value—distribution is too small. Error occurred in assumption <Name> in cell <Cell>.	Crystal Ball attempted to generate a value for one of the assumptions, but a severely truncated distribution prevented the value generation.	Recheck the assumptions and reduce the degree of truncation.
Unable to open any more windows.	Crystal Ball required at least 3 KB of memory to open a window. You have less than that much RAM and virtual memory left.	• Try solutions for the message: Memory limit reached • If this is a recurring problem, install more RAM
Unable to read complete custom distribution record.	You are trying to open a saved run file that has been corrupted.	Check your disk for errors with a disk utility, and try to repair them.

Message	Reason	Solution
Upper percentile must be greater than lower percentile.	The higher percentile value is less than the lower percentile value.	Either change the higher percentile value to be above the lower percentile or change the lower percentile value to be below the higher percentile value. If you want the upper percentile to be less than the lower percentile, in the Forecast window, select Preferences > Statistics and change the option to show percentiles as the Probability Above A Value.
Value or range overlaps an existing value or range.	Ranges and single values cannot overlap each other. However, the ending value of a continuous range can begin at the starting value of another continuous range.	Check your values and re-enter to avoid overlapping values.
Warning: one or more sheets open during the simulation have been closed or renamed—some data may not be saved.	You closed or renamed sheets between running a simulation and extracting data.	Rerun the simulation and extract data before you close or rename any sheets.

Appendix B

Equations and Methods

In this appendix

This appendix provides formulas for the various distribution types and describes the methods used for random generation of numbers. It also provides the following formulas:

- Mean
- Variance
- Standard Deviation
- Coefficient of Skewness
- Coefficient of Kurtosis
- Mean Standard Error
- Coefficient of Variability
- Correlation Coefficient
- Random number generator

Formulas for probability distributions

This section contains the formulas used in calculating the probability distributions. It also describes the random generation methods used in Crystal Ball.

Beta distribution

Parameters: Alpha (α), Beta (β), Scale (s)

Formula:

$$f(x) = \begin{cases} \dfrac{\left(\frac{x}{s}\right)^{(\alpha-1)}\left(1-\frac{x}{s}\right)^{(\beta-1)}}{\beta(\alpha, \beta)} & \text{if } 0 < x < s;\ \alpha > 0;\ \beta > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\beta(\alpha, \beta) = \dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

and where $\Gamma(\beta)$ is the Gamma function.

Method 1: Gamma Density Combination

Comment: The Beta variate is obtained from: $\left(\dfrac{x}{x+y}\right) \cdot s$

where x = Gamma $(\alpha, 1)$ and y = Gamma $(\beta, 1)$.

Method 2: Rational Fraction Approximation method with a Newton Polish step

Comment: This method is used instead of Method 1 when Latin Hypercube Sampling is in effect.

Binomial distribution

Parameters: Probability (p), Trials (n)

Formula: $P\{x = i\} = \binom{n}{i} p^i (1-p)^{(n-i)}$

for $i = 0, 1, 2, \ldots$

where $\binom{n}{i} = \dfrac{n!}{i!(n-i)!}$

Method: Direct Simulation

Comment: Computation time increases linearly with number of trials.

Exponential distribution

Parameters: Rate (λ)

Formula:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

Method: Inverse Transformation

Extreme value distribution

The extreme value distribution is a skewed distribution. Besides the two parameters (Mode and Scale), Crystal Ball selects the direction of skewness with Min and Max.

Parameters: Mode (m), Scale (α)

Formula: $f(x) = \dfrac{1}{\alpha} \cdot z \cdot e^{-z}$ for $-\infty < x < \infty$

where $z = e^{-\left(\frac{x-m}{\alpha}\right)}$

Method: Inverse Transformation.

Gamma distribution

Parameters: Location (L), Shape (α), Scale (β)

Formula:

$$f(x)=\begin{cases}\dfrac{\left(\dfrac{x-L}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x-L}{\beta}\right)}}{\Gamma(\alpha)\ \beta} & \text{if } x > L \\ 0 & \text{if } x \le L\end{cases}$$

where $\Gamma(\alpha)$ is the gamma function

Note: some textbook Gamma formulas use:

$\beta = \dfrac{1}{\lambda}$

Method 1: When $\alpha < 1$, Vaduva's rejection from a Weibull density.

When $\alpha > 1$, Best's rejection from a *t* density with 2 degrees of freedom.

When $\alpha = 1$, inverse transformation.

Method 2: Rational Fraction Approximation method with a Newton Polish step

Comment: This method is used instead of Method 1 when Latin Hypercube Sampling is in effect.

Geometric distribution

Parameters: Probability (p)

Formula: $P\{x = i\} = (1-p)^{i-1}p \qquad \text{for } i = 1, 2, \ldots$

Method: Inverse Transformation

Hypergeometric distribution

Parameters: Probability (p), Trials (n), Population Size (N), Number of items in the population with the characteristic x (N_P).

Formula:

$$P\{x = i\} = \frac{\binom{N_p}{i}\binom{N-N_p}{n-i}}{\binom{N}{n}}$$

where $\binom{n}{i} = \frac{n!}{i!(n-i)!}$

for $i = Max(n-(N-N_p), 0)\ldots \quad Min(n, N_p)$

Method: Direct Simulation

Comment: Computation time increases linearly with Population Size.

Logistic distribution

Parameters: Mean (μ), Scale (α)

Formula:

$$f(x) = \frac{z}{\alpha(1+z)^2} \qquad \text{for } -\infty < x < \infty$$

where $z = e^{-\left(\frac{x-\mu}{\alpha}\right)}$

Method: Inverse Transformation.

Lognormal distribution

Parameters: Log Mean (m), Log Standard Deviation (s)

Formula: $$f(x) = \frac{1}{x\sqrt{2\pi}\,s} e^{-(\log_e(x) - m)^2/2s^2} \quad \text{for } 0 < x < \infty$$

Method: Polar Marsaglia

Comment: A Normal variate in log space is generated and then exponentiated.

The following parameter relationships hold:

$$\text{geometric mean} = e^m$$
$$\text{geometric std. dev.} = e^s$$

$$\text{arithmetic median} = \text{geometric mean}$$

$$\text{arithmetic mean} = \text{median} \cdot e^{\left(\frac{s^2}{2}\right)}$$

$$\text{arithmetic std. dev.} = \text{median} \cdot \sqrt{e^{s^2} \cdot (e^{s^2} - 1)}$$

Negative binomial distribution

Parameters: Probability (p), Shape (β)

Formula: $$P\{x = i\} = \begin{cases} \binom{x-1}{\beta-1} p^{\beta}(1-p)^{x-\beta} & i = \beta, \beta+1, \beta+2\ldots \\ 0 & \text{otherwise} \end{cases}$$

where $$\binom{x-1}{\beta-1} = \frac{(x-1)!}{(\beta-1)!(x-\beta)!}$$

Method: Direct Simulation through summation of Geometric variates.

Comment: Computation time increases linearly with Shape.

Normal distribution

Parameters: Mean (m), Standard Deviation (s)

Formula: $f(x) = \frac{1}{\sqrt{2\pi}\, s} e^{-(x-m)^2/2s^2}$ for $-\infty < x < \infty$

Method 1: Polar Marsaglia

Comment: This method is somewhat slower than other methods, but its accuracy is essentially perfect.

Method 2: Rational Fraction Approximation

Comment: This method is used instead of the Polar Marsaglia method when Latin Hypercube sampling is in effect.

This method has a 7-8 digit accuracy over the central range of the distribution and a 5-6 digit accuracy in the tails.

Pareto distribution

Parameters: Location (L), Shape (β)

Formula:

$$f(x) = \begin{cases} \frac{\beta \cdot L^{\beta}}{x^{(\beta+1)}} & \text{if } x > L \\ 0 & \text{if } x \le L \end{cases}$$

and $L > 0$

Method: Inverse Transformation.

Poisson distribution

Parameters: Rate (λ)

Formula: $P\{x = i\} = e^{-\lambda} \frac{\lambda^i}{i!}$ for $i = 0, 1, 2, \ldots$

Method: Direct Simulation through summation of Exponential variates

Comment: Computation time increases linearly with Rate.

Triangular distribution

Parameters: Min, Likeliest, Max

Formula:

$$f(x) = \begin{cases} \dfrac{h(x - Min)}{Likeliest - Min} & \text{if } Min < x < Likeliest \\ \dfrac{h(x - Likeliest)}{Likeliest - Max} + h & \text{if } Likeliest < x < Max \\ 0 & \end{cases}$$

where $h = \dfrac{2}{Max - Min}$

Method: Inverse Transformation

Uniform distribution

Parameters: Min, Max

Formula:

$$f(x) = \begin{cases} \dfrac{1}{(Max - Min)} & \text{if } Min < x < Max \\ 0 & \end{cases}$$

otherwise

Method: Multiplicative Congruential Generator
This routine uses the iteration formula:

$$r <\!-(630, 360, 016 \cdot r) \bmod (2^{31} - 1)$$

Comment: The generator has a period of length $2^{31} - 2$, or 2,147,483,646. This means that the cycle of random numbers repeats after several billion trials. This formula is also implemented in the widely used IMSL mathematical package.

Weibull distribution

Parameters: Location (L), Scale (α), Shape(β)

Formula:

$$f(x) = \begin{cases} \left(\frac{\beta}{\alpha}\right)\left(\frac{x-L}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x-L}{\alpha}\right)^{\beta}} & \text{if } x \geq L \\ 0 & \text{if } x < L \end{cases}$$

Method: Inverse Transformation

Custom distribution

Formula: The formula consists of a lookup table of single data points, continuous ranges, and discrete ranges. Each item in the table has a distinct probability relative to the other items. In addition, ranges might be positively or negatively sloped, giving values on one side or the other a higher probability of occurring.

Method: Sequential search of relative probabilities table

Comments: A Uniform variate is generated in the range (0, total relative probability). A sequential search of the relative probabilities table is then performed. The Inverse Transformation method is used whenever the uniform variate falls within a continuous or discrete range that is sloped in one direction or the other.

Additional comments

All of the non-uniform generators use the same uniform generator as the basis for their algorithms.

The Inverse Transformation method is based on the property that the cumulative distribution function for any probability distribution increases monotonically from zero to one. Thus, the inverse of this function can be computed using a random uniform variate in the range (0, 1) as input. The resulting values then have the desired distribution.

The Direct Simulation method actually performs a series of experiments on the selected distribution. For example, if a binomial variate is being generated with Prob =. 5 and Trials = 20, then 20 uniform variates in the range (0, 1) are generated and compared with Prob. The number of uniform variates found to be less than Prob then becomes the value of the binomial variate.

Distribution fitting methods

During distribution fitting, Crystal Ball computes Maximum Likelihood Estimators (MLEs) to fit most of the probability distributions to a data set. In effect, this method chooses values for the parameters of the distributions that maximize the probability of producing the actual data set. Sometimes, however, the MLEs do not exist for some distributions (e.g., gamma, beta). In these cases, Crystal Ball resorts to other natural parameter estimation techniques.

When the MLEs do exist, they exhibit desirable properties:

- They are minimum-variance estimators of the parameters.
- As the data set grows, the biases in the MLEs tend to zero.

For several of the distributions (e.g. uniform, exponential), it is possible to remove the biases after computing the MLEs to yield minimum-variance unbiased estimators (MVUEs) of the distribution parameters. These MVUEs are the best possible estimators.

Precision control formulas

The Precision Control feature of Crystal Ball lets you stop the simulation when the specified precision of the chosen statistics is reached. Crystal Ball periodically checks whether the confidence interval is less than the specified precision. The following sections describe how Crystal Ball calculates the confidence interval for each statistic.

Mean confidence interval

Formula: $$z \cdot \frac{s}{\sqrt{n}}$$

where:

s is the standard deviation of the forecast
n is the number of trials
z is the z-value based on the specified confidence level (in the Run Preferences > Trials dialog).

Standard deviation confidence interval

Formula: $$z \cdot s \cdot \sqrt{\frac{k-1}{4 \cdot (n-1)}}$$

where:
s is the standard deviation
k is the kurtosis
n is the number of trials
z is the z-value based on the specified confidence level (in Run Preferences > Trials dialog).

Percentiles confidence interval

To calculate the confidence interval for the percentiles, instead of a mathematical formula, Crystal Ball uses an analytical bootstrapping method.

Statistical Note: *Patent pending for this bootstrapping method for calculating precision.*

Formulas for other statistical terms

This section contains the mathematical formulas used in calculating the descriptive statistics.

Mean

Formula: $$\frac{1}{n}\sum_{i=1}^{n} x_i \qquad (\bar{x})$$

Variance

Formula: $$\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 \qquad (s^2)$$

Standard deviation

Formula: $$\sqrt{\frac{1}{n-1}\sum(x_i-\bar{x})^2} \qquad (s)$$

Coefficient of skewness

Description: Skewness is computed by finding the third moment about the mean and dividing by the cube of the standard deviation.

Formula: $$\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{s^3}$$

Coefficient of kurtosis

Description: Kurtosis, or peakedness, is computed by finding the fourth moment about the mean and dividing by the quadruple of the standard deviation.

Formula: $$\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^4}{s^4}$$

Mean standard error

Description: This statistic is a method for determining the accuracy of the mean statistic. It tells you the probability that the estimated mean will deviate from the true mean by less than a specific amount. The probability that the true mean of the defined problem is within the estimated mean plus or minus the mean standard error is approximately 68%.

Crystal Ball Note: *This applies to the Monte Carlo method. The mean standard error for Latin hypercube sampling is typically much smaller.*

Formula: $\frac{s}{\sqrt{n}}$

where: s = Standard Deviation
n = Number of Trials

The error estimate might be inverted to show the number of trials needed to yield a desired error (ε).

$$n = \frac{s^2}{\varepsilon^2}$$

Coefficient of variability

Description: This statistic gives an indication of the variability of the forecast value on an absolute scale. This coefficient is independent on the units of forecast and typically ranges between 0 and 1. The coefficient might exceed 1 in a small number of cases when the standard deviation of the forecast is unusually high. This statistic is computed by dividing the standard deviation by the mean.

Formula: $\frac{s}{\bar{x}}$

Correlation coefficient

Description: This statistic indicates the degree of linear relationship between two variables (assumption cells). A positive correlation coefficient means that as the values of one variable increase or decrease, so do the values of the other variable. A negative correlation coefficient means that as the values of one variable increase, the values of the other variable decrease by a similar amount, and vice-versa. A correlation coefficient of zero means that there is no linear relationship between the two variables, and they are considered to be uncorrelated.

Crystal Ball uses rank correlation to correlate assumption values. This means that assumption values are replaced by their rankings from lowest to highest value by the integers 1 to N, prior to computing the correlation coefficient. This method allows distribution types to be ignored when correlating assumptions.

Formula:

$$\frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \cdot \sqrt{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2}}$$

Random number generator

The random number generator given below is used as the basis for all of the non-uniform generators. For no starting seed value, Crystal Ball takes the value of the number of milliseconds elapsed since Windows started.

Method: Multiplicative Congruential Generator
This routine uses the iteration formula:

$$r <-(630, 360, 016 \cdot r) \bmod (2^{31}-1)$$

Comment: The generator has a period of length $2^{31}-2$, or 2,147,483,646. This means that the cycle of random numbers repeats after several billion trials. This formula is also implemented in the widely used IMSL mathematical package, and is discussed in detail in the *Simulation Modeling & Analysis* reference in the bibliography.

.

Appendix C

Keyboard Commands

In this appendix

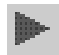
- Command key combinations
- Cell, Run, and CBTools menu commands
- Forecast windows commands
- Forecast chart hot keys
- Overlay/Comparison chart hot keys
- Trend chart hot keys
- Sensitivity chart hot keys

Command key combinations

Use the following <Alt>-key combinations to execute the spreadsheet Cell, Run, and CBTools menu commands and forecast windows commands without using the mouse.

Cell, Run, and CBTools menu commands

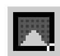
Commands	*Keystrokes*
About Crystal Ball	<Alt>-u, a
Batch Fit tool	<Alt>-b, a
Bootstrap tool	<Alt>-b, b
Cell Preferences	<Alt>-c, l
Clear Data	<Alt>-c, e
Close Crystal Ball	<Alt>-u, c
Copy Data	<Alt>-c, c
Correlation Matrix tool	<Alt>-b, x
Create Report	<Alt>-u, r
Decision Table tool	<Alt>-b, c
Define Assumption	<Alt>-c, a
Define Decision	<Alt>-c, d
Define Forecast	<Alt>-c, f
Extract Data	<Alt>-u, d
Forecast Windows	<Alt>-u, f
Freeze Assumptions	<Alt>-c, z
Paste Data	<Alt>-c, p
Reset	<Alt>-u, e

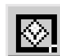
 Run, Continue .... <Alt>-u, u

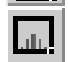
 Run Preferences .... <Alt>-u, p

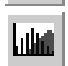
Save Run .... <Alt>-u, v

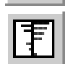
Select Some .... <Alt>-c, s

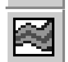
 Select All Assumptions .... <Alt>-c, m

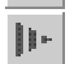
 Select All Decisions .... <Alt>-c, i

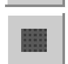
 Select All Forecasts .... <Alt>-c, r

 Open Overlay Chart .... <Alt>-u, l

 Open Sensitivity Chart .... <Alt>-u, n

 Open Trend Chart .... <Alt>-u, t

 Single Step .... <Alt>-u, s

 Stop .... <Alt>-u, o

Tornado Chart tool .... <Alt>-b, t

Two-dimensional (2D) Simulation tool .... <Alt>-b, d

Forecast windows commands

Commands *Keystrokes*

Bring windows to front .... <Alt>-<Tab>

Chart .... <Alt>-p, c

Help Commands .... <Alt>-h, c

Copy Chart, Copy Statistics .... <Alt>-e, c

Cumulative Chart .... <Alt>-v, c

Difference .... <Alt>-v, i

Display Range .... <Alt>-p, d

Frequency Chart .... <Alt>-v, f

Format .... <Alt>-p, f

Help Glossary .... <Alt>-h, g

Help Index .... <Alt>-h, i

Help Messages .... <Alt>-h, m

Help Procedures .... <Alt>-h, p

Percentiles .... <Alt>-v, p

Reverse Cumulative Chart .... <Alt>-v, r

Statistics (Preferences)................................ <Alt>-p, s

Statistics (View) .... <Alt>-v, s

Using Help .... <Alt>-h, h

Hot keys

Use the Crystal Ball hot keys to bypass the Chart Preferences dialogs for the forecast, overlay, trend and sensitivity charts. With this feature, you can change the chart display quickly as you interpret the results of the simulation.

Crystal Ball Note: *Occasionally, these keys do not work properly while a simulation is running. When this happens, use the appropriate Preferences dialog or wait until the simulation stops.*

Forecast chart hot keys

To change a forecast chart, use these keys while the forecast window appears in the front:

- Type **d** to change from a frequency distribution, to a cumulative distribution, to a reverse cumulative distribution. You can also press the spacebar to alternate among these types.
- Type **g** to change the number of groups from 300, to 150, 100, 75, 50, 25, 15 for small charts or from 450, to 225, 150, 90, 75, 50, 25 for large charts.

- Type **l** to toggle the horizontal grid lines between on and off.
- Type **t** to change the chart type from an area chart, to an outline chart, to a column chart.

Overlay/Comparison chart hot keys

To change an overlay or comparison chart, use these keys while the window appears in front:

- Type **d** to change the distribution display type from frequency, to cumulative, to reverse cumulative.
- Type **g** to change the number of groups from 300, to 150, 100, 75, 50, 25, 15 for small charts or from 450, to 225, 150, 90, 75, 50, 25 for large charts.
- Type **t** to change the chart type from area, to outline, to column, to dots and lines.
- Type **i** to toggle between comparison mode and difference mode.
- Type **l** to toggle the horizontal grid lines between on and off.
- Type **n** to toggle the legend display on the overlay chart between on and off.

Trend chart hot keys

To change a trend chart, use these keys while the trend chart appears in the front:

- Type **f** to toggle the forecast names between horizontal and vertical formats.
- Type **d** to change the placement of the display bands from centered around, to above, to below a given value.
- Type **l** to change the grid lines from horizontal, to vertical, to horizontal and vertical, to none.
- Type **v** to change the value axis endpoints from relative, to minimum through maximum, to zero based.

Sensitivity chart hot keys

To change a sensitivity chart, use these keys while the sensitivity chart appears in the front:

- Type **t** to toggle the measurement of sensitivity between rank correlation and contribution to variance.
- Type **d** to change the type of data included from only assumptions, to only forecasts, to both assumptions and forecasts.
- Type **g** to toggle the value cutoff between on and off.
- Type **h** to toggle the count cutoff between on and off.

Appendix D

Index of Commands

In this appendix

Crystal Ball **menu**s

- Cell
- Run
- CBTools

Forecast menus

- Edit
- Preferences
- View
- Run
- Help

Crystal Ball menus

This appendix provides information about the drop down menus used in Crystal Ball.

The Cell menu

Define Assumption

Defines the selected cell as an assumption. Also see “Defining assumptions” on page 131.

Define Decision

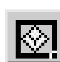
Defines selected cell as a decision variable. Also see “Defining decision variables” on page 151.

Define Forecast

Defines selected cell as a forecast. Also see “Defining forecasts” on page 153.

Select All Assumptions

Selects all of the assumptions in the current spreadsheet.



Select All Decisions

Selects all of the decision variables in the current spreadsheet.

Select All Forecasts

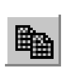
Selects all of the forecasts in the current spreadsheet.

Select Some

Views, selects or defines assumptions or forecasts from a list. Also see “Selecting and reviewing your data” on page 159.

Freeze Assumptions

Excludes a chosen set of assumptions from a simulation. Also see “Freezing assumptions” on page 150.



Copy Data

Copies the selected assumption or forecast cell to the clipboard. Also see “Editing Crystal Ball data” on page 157.

Paste Data

Pastes the copied assumption or forecast cell(s) into selected cell(s). Also see “Editing Crystal Ball data” on page 157.

Clear Data

Removes the selected assumption or forecast cell from Crystal Ball, leaving the initial value. Also see “Editing Crystal Ball data” on page 157.

Cell Preferences

Used for changes to the highlighting of assumption and forecast cells, and adds automatic note fields on the spreadsheet.

The Run menu

Run/Stop/Continue

This command has three modes, depending on the status of the simulation.

Run	Begins the simulation if it has not yet been started or has just been reset.
Stop	Pauses the simulation if it is running.
Continue	Restarts the simulation if it has been paused or stopped by a calculation error. Also see “Running the simulation” on page 172.

Reset

Resets the current simulation by closing open forecast windows and clearing all the forecast values. Also see “Running the simulation” on page 172.

Single Step

Allows running the simulation one step at a time. Helps to locate where an error occurred.

Run Preferences

Specifies certain options and parameters of the simulation.

Forecast Windows

Opens or closes all or selected forecast windows.

Open Overlay Chart

Opens an overlay chart that allows for comparison of forecasts to each other or to standard distributions. Also see “Creating the overlay chart” on page 206.

Open Trend Chart

Opens a trend chart that summarizes information scattered over multiple forecasts. Also see “Creating the trend chart” on page 216.

Open Sensitivity Chart

Opens a sensitivity chart that shows the influence your assumptions have on a forecast cell. Also see “Understanding sensitivity and using the sensitivity chart” on page 223.

Create Report

Requests a formatted report of the simulation results.

Extract Data

Exports data produced by the simulation.

Save Run

Saves a simulation.

Restore Run

Opens a simulation saved earlier from disk.

Close Crystal Ball

Closes the Crystal Ball program and removes the Cell and Run menus from the application menu bar. Also see “Closing Crystal Ball” on page 51.

About Crystal Ball

Displays information about the program (for example, company name, version, author, and so on.).

The CBTools menu

Batch Fit

Starts the Batch Fit tool, which automatically fits probability distributions to multiple data series.

Bootstrap

Starts the Bootstrap tool, which addresses the reliability and accuracy of your forecast statistics.

Correlation Matrix

Starts the Correlation Matrix tool, which rapidly defines and automates correlations between your assumptions.

Decision Table

Starts the Decision Table tool, which evaluates the effects of alternate decisions in a simulation model.

Tornado Chart

Starts the Tornado Chart tool, which individually analyzes the impact of each model variable on a target outcome.

2D Simulation

Starts the 2D Simulation tool, which independently addresses uncertainty and variability in your models using two-dimensional simulation.

Forecast menus

These menus are found on the forecast chart, which displays after a simulation has been run.

Edit

Cut

Not active. Use the Copy and Clear commands on the Cell menu. Also see “Editing Crystal Ball data” on page 157.

Copy

Copies the current view of the forecast chart to the clipboard.

Paste

Not active. Use the Paste Assumptions/Forecasts command on the Cell menu. Also see “Editing Crystal Ball data” on page 157.

Clear

Not active. Use the Clear Assumptions/Forecasts command on the Cell menu. Also see “Editing Crystal Ball data” on page 157.

Preferences

Format

Permits changes to be made to the way numbers are formatted on the forecast windows. Also see "Formatting forecast charts" on page 200.

Chart

Permits changes to be made to the appearance of the forecast chart. Also see "Customizing the forecast chart" on page 194.

Display Range

Permits focus on a particular range of the forecast chart. Also see "Focusing on the display range" on page 191.

Statistics

Permits changes to be made to the way certain descriptive statistics are calculated for a forecast. Also see "Changing how statistics are calculated" on page 203.

View

Statistics

Displays a full range of statistics for the selected forecast. Also see "Interpreting the statistics" on page 202.

Percentiles

Displays percentile information in 10% increments for the selected forecast. Also see "Interpreting the statistics" on page 202.

Frequency Chart

Displays the number or frequency of values occurring in a given interval for the selected forecast. Also see "Changing the distribution type" on page 197.

Cumulative Chart

Displays the number or proportion (or percentage) of values less than or equal to a given amount for the selected forecast. Also see "Changing the distribution type" on page 197.

Reverse Cumulative Chart

Displays the number or proportion (or percentage) of values greater than or equal to a given amount. Also see “Changing the distribution type” on page 197.

Difference

Active only with the overlay chart. Displays the frequency differences between the first distribution on the overlay chart and each of the remaining distributions. Also see “Comparison Chart - Difference” on page 144.

Run

Same as the Crystal Ball Run menu on the application menu bar with the following exception:

Restore Run

Not active. It is only available when you reset a simulation, and when you reset a simulation, all the forecast windows disappear.

Close Crystal Ball

Not active. Use Close Crystal Ball from the Run menu on the application menu bar.

Help

Index

Displays the Help categories in Crystal Ball for Windows.

Using Help

Presents the basic features of Windows Help. Also includes practice exercises.

Procedures

Displays and describes all Crystal Ball procedures.

Commands

Displays and describes Crystal Ball menu commands.

Glossary

Displays and defines terms used in Crystal Ball.

Messages

Displays Crystal Ball program messages.

Appendix E

Default Names and Distribution Parameters

In this appendix

Naming Defaults

Distribution Parameter Defaults

- Beta
- Binomial
- Custom
- Exponential
- Extreme Value
- Gamma
- Geometric
- Hypergeometric
- Logistic
- Lognormal
- Negative Binomial
- Normal
- Pareto
- Poisson
- Triangular
- Uniform
- Weibull

This first section of this appendix describes the process Crystal Ball uses to name distribution and forecast charts. The second section shows the values it assigns to each of the distribution types.

Naming defaults

When a distribution or forecast dialog appears, Crystal Ball uses the following sequence to generate a name for the distribution or forecast:

1. **Checks the cell immediately to the left of the selected cell. If it is a text cell, Crystal Ball uses that text as the distribution or forecast name.**
2. **Checks the cell above the selected cell. If it is a text cell, Crystal Ball uses that text as the distribution or forecast name.**
3. **If there is no applicable text or range name, Crystal Ball uses the cell coordinates for the name (for example, B3 or C7).**

Distribution parameter defaults

This section lists the initial values Crystal Ball provides for the primary parameters in the assumption dialog.

If an alternate parameter set is selected as the default mode, the primary parameters are still calculated as described below before conversion to the alternate parameters.

Beta

Alpha 2.0

Beta.. 3.0

Scale.. cell value

Binomial

If the cell value is between 0 and 1:

Probability (Prob) ....cell value

Trials 50

If the cell value is between 1 and the maximum number of binomial trials (1,750):

Probability (Prob) 0.5

Trials ....cell value

Otherwise:

Probability (Prob) 0.5

Trials 50

Custom

Initially empty

Exponential

Rate ....1/absolute value of cell value

Extreme value

Mode ....cell value

Scale 1

Gamma

Location (Loc) ....cell value

Scale 1.0

Shape 2.0

Geometric

If the cell value is between 0 and 1:

Probability (Prob) ....cell value

Otherwise:

Probability (Prob) 0.2

Hypergeometric

If the cell value is between 0 and 1:

Probability (Prob) ....cell value

Trials 50

Population Size (Pop) 100

If the cell value is between 1 and the maximum number of Hypergeometric trials (1,750):

Probability (Prob) 0.5

Trials ....cell value/2.0

Population Size (Pop) ....cell value

Otherwise:

Probability (Prob) 0.5

Trials 50

Population size (Pop)...................................100

Logistic

Mean ....cell value

Scale...1.0

Lognormal

If the cell value is greater than 0:

Mean ....cell value

Standard Deviation (Std Dev)......................cell value/10.0

Otherwise:

Mean ....e

Standard Deviation (Std Dev)......................1.0

Negative binomial

If the cell value is between 0 and 1:

Probability (Prob) ....cell value

Trials 10

Otherwise:

Probability (Prob) 0.2

Trials 10

Normal

Mean ....cell value

Standard Deviation (Std Dev)......................cell value/10.0

Pareto

If the cell value is between 1.0 and 1,000:

Location .... cell value

Shape 2.0

Otherwise:

Location 1.0

Shape 2.0

Poisson

If the cell value is between 0 and the maximum rate (1,500):

Rate .... cell value

Otherwise:

Rate 10.0

Triangular

Likeliest .... cell value

Minimum (Min) .... cell value - cell value/10.0

Maximum (Max) .... cell value + cell value/10.0

Uniform

Minimum (Min) .... cell value - cell value/10.0

Maximum (Max) .... cell value + cell value/10.0

Weibull

Location (Loc)..cell value

Scale..1.0

Shape 2.0

Bibliography

In this bibliography

Bibliography entries by subject

Bootstrap

Efron, Bradley, and Robert J. Tibshirani. "An Introduction to the Bootstrap," *Monographs on Statistics and Applied Probability, vol. 57.* New York: Chapman & Hall, 1993.

Mooney, C. Z., and R. D. Duval. "Bootstrapping: A nonparametric approach to statistical inference," *Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-095*. Newbury Park, CA: Sage, 1993.

Monte Carlo simulation

Burmaster, D.E. and P.D. Anderson, "Principles of Good Practice for the Use of Monte Carlo Techniques in Human Health and Ecological Risk Assessments," *Risk Analysis* 14(4):477-481 (1994)

Hammersley, J.M. and D.C. Handscomb. *Monte Carlo Methods*. London: Methuen & Co Ltd, 1964.

Kalos, Malvin H. and Paula A. Whitlock. *Monte Carlo Methods.* New York: John Wiley & Sons, 1986.

Morgan, Byron J. T. *Elements of Simulation*. Portland, ME: Chapman and Hall, 1984.

Rubinstein, Reuven Y. *Simulation and the Monte Carlo Method.* New York: John Wiley & Sons, 1981.

Thompson, K.M., D.E. Burmaster, and E.A.C. Crouch, "Monte Carlo Techniques for Quantitative Uncertainty Analysis in Public Health Risk Assessments," *Risk Analysis* 12(1):53-63 (1992)

Yakowitz, Sidney J. *Computational Probability and Simulation.* Reading, MA: Addison-Wesley Publishing Co., 1977.

Probability theory and statistics

Alder, Henry L. and Edward B. Roessler. *Introduction to Probability and Statistics.* San Francisco: W.H. Freeman & Co, 1972.

Biswas, Suddhendu. *Topics in Statistical Methodology.* New York: John Wiley & Sons, Inc., 1991.

Fraser, D.A.S., *Probability and Statistics: Theory and Applications.* North Scituate, MA: Duxbury Press, 1976.

Johnson, Norman L. and Samuel Kotz. *Distributions in Statistics: Continuous Multivariate Distributions.* New York: John Wiley & Sons, Inc., 1972.

Kapur, K.C. and L.R. Lamberson. *Reliability in Engineering Design.* New York: John Wiley & Sons, 1977.

Larson, Harold J. *Introduction to Probability Theory and Statistical Inference.* New York: John Wiley & Sons, 1974.

Logothetis, N. and V. Rothschild. *Probability Distributions.* New York: John Wiley & Sons, Inc., 1986.

Spurr, William and Charles P. Bonini. *Statistical Analysis for Business Decisions.* Homewood, IL: Richard D. Irwin, Inc., 1973.

Wadsworth, George P. and Joseph G. Bryan. *Introduction to Probability and Random Variables.* New York: McGraw-Hill Book Co., 1960.

Winkler, Robert L. and William L. Hays. *Statistics—Probability, Inference, and Decision.* New York: Holt, Rinehart, and Winston, 1975.

Random variate generation methods

Devroye, Luc. *Non-Uniform Random Variate Generation*. New York: Springer-Verlag, 1986.

Iman, Ronald L. and Conover, W.J., “A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables,” *Communications in Statistics*, Vol B11, No. 3. 1982.

Iman, Ronald L. and Davenport, J. M., “Rank Correlation Plots for Use With Correlated Input Variables,” *Communications in Statistics*, Vol. B11, No. 3. 1982.

Iman, Ronald L. and Davenport, J. M. *Latin Hypercube Sampling (A Program User’s Guide).* Technical Report SAND79-1473, Sandia National Laboratories, 1980.

Kennedy, William J., Jr. *Statistical Computing.* New York: Marcel Dekker, Inc., 1980.

Knuth, Donald E. *The Art of Computer Programming*, Vols I-III. Reading, MA: Addison-Wesley Publishing Co., 1969.

Newman, Thomas G. and Odell, Patrick L. *The Generation of Random Variates.* New York: Hafner Publishing Co., 1971.

Press, William H., et al. *Numerical Recipes in C.* Cambridge, England: Cambridge University Press, 1988.

Odeh, R.E. and J.O. Evans, "Percentage Points of the Normal Distribution," *Applied Statistics*. London: Royal Statistical Society, 1974.

Sedgewich, Robert. *Algorithms*. Reading, MA: Addison Wesley Publishing Co., 1983.

Specific distributions

Extreme value distribution

Castillo, Enrique. *Extreme Value Theory in Engineering.* London: Academic Press, 1988.

Lognormal distribution

Aitchison, J. and J.A. Brown. *The Lognormal Distribution.* New York: Cambridge University Press, 1973.

Weibull distribution

King, James R. *Probability Charts for Decision Making.* New York: Industrial Press, Inc., 1971.

Henley, Ernest J. and Hiromitsu Kumamoto. *Reliability Engineering and Risk Assessment.* Englewood Cliffs, NJ: Apprentice-Hall, Inc., 1981.

Tornado charts and sensitivity analysis

Clemen, Robert T. *Making Hard Decisions: An Introduction to Decision Analysis*. 2nd ed. Belmont, CA: Duxbury Press, 1996.

Morgan, M. Granger, and Max Henrion; with a chapter by Mitchell Small. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. New York: Cambridge University Press, 1990.

Two-dimensional simulation

Burmaster, David E., and Andrew M. Wilson. "An Introduction to Second-Order Random Variables in Human Health Risk Assessments," *Human and Ecological Risk Assessment: vol. 2, no. 4*. pp 892-919, 1996.

Frey, H. Christopher, and David S. Rhodes. "Characterizing, Simulating, and Analyzing Variability and Uncertainty: An Illustration of Methods Using an Air Toxics Emissions Example," *Human and Ecological Risk Assessment: vol. 2, no. 4*. pp 762-797, 1996.

Hoffman, F. O., and J. S. Hammonds. "Propagation of uncertainty in risk assessments: The need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability," *Risk Analysis*, vol. 14, no. 5. pp 707-712, 1994.

Rai, S. N., D. Drewski, and S. Bartlett. "A General Framework for the Analysis of Uncertainty and Variability in Risk Assessment," *Human and Ecological Risk Assessment: vol. 2, no. 4*. pp 972-989, 1996.

Uncertainty analysis

Henrion, Max and M. Granger Morgan. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis.* Cambridge: Cambridge University Press, 1990.

Hoffman, F.O. and J.S. Hammonds, "Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability" *Risk Analysis* 14(5): 707-712. 1994

International Atomic Energy Agency, “Evaluating the Reliability of Predictions Made Using Environmental Transfer Models,” *IAEA Safety Series* No. 100:1-106. STI/PUB/835. Vienna, Austria 1989

Kelton, W. David and Averill M. Law. *Simulation Modeling & Analysis.* New York: McGraw-Hill, Inc. 1991.

Glossary

In this glossary

A compilation of terms specific to Crystal Ball as well as statistical terms used in this manual.

Glossary

assumption

An estimated value or input to a spreadsheet model.

assumption cell

A value cell in a spreadsheet model that has been defined as a *probability distribution*.

CDF

Cumulative distribution function that represents the probability that a variable will fall at or below a given value.

certainty bands

In a trend chart, a graphic depiction of a particular *certainty range* for each forecast.

certainty level

The percentage of values in the certainty range compared to the number of values in the entire range.

certainty range

The linear distance for the set of values between the certainty grabbers on the forecast chart.

coefficient of variability *also coefficient of variance or coefficient of variation*

A measure of relative variation that relates the standard deviation to the mean. Results can be represented in percentages for comparison purposes. (see "Coefficient of variability" on page 120)

continuous probability distribution

A *probability distribution* that describes a set of uninterrupted values over a range. In contrast to the Discrete distribution, the Continuous distribution assumes there are an infinite number of possible values.

correlation

In Crystal Ball, a dependency that exists between *assumption cells*. (see "Understanding other statistical terms" on page 116)

correlation coefficient

A number between -1 and 1 that specifies mathematically the degree of positive or negative correlation between *assumption cells*. A *correlation* of 1 indicates a perfect positive correlation, minus 1 indicates a perfect negative correlation, and 0 indicates there is no *correlation*. (see "Correlation coefficient" on page 122)

cumulative frequency distribution

A chart that shows the number or proportion (or percentage) of values less than or equal to a given amount.

decision variable

A Crystal Ball variable in your model that you can control.

deterministic model

Another name for a *spreadsheet model* which yields single-valued results.

discrete probability distribution

A *probability distribution* that describes distinct values, usually integers, with no intermediate values. In contrast, the Continuous distribution assumes there are an infinite number of possible values.

display range

The linear distance for the set of values displayed on the forecast chart.

dominant

A relationship between distributions in which one distribution's values for all percentile levels are higher than another's. *(see also Subordinate)*

entire range

The linear distance from the minimum *forecast value* to the maximum *forecast value*.

forecast

A statistical summary of the assumptions in a spreadsheet model, output graphically or numerically.

Glossary

forecast cell

A formula cell that has been defined as a forecast and refers either directly or indirectly to *assumption cells*.

forecast definition

The forecast name and parameters assigned to a cell in a Crystal Ball dialog.

forecast formula

A formula that has been defined as a *forecast cell*.

forecast value *also trial*

A value calculated by the forecast formula during an *iteration*. These values are kept in a list for each forecast, and are summarized graphically in the forecast chart and numerically in the descriptive statistics.

formula cell

A cell that contains a mathematical formula.

frequency *also frequency count*

The number of times a value recurs in a *group interval*.

frequency distribution

A chart that graphically summarizes a list of values by subdividing them into groups and displaying their frequency counts.

goodness-of-fit

A set of mathematical tests performed to find the best fit between a standard probability distribution and a data set.

grabber *also certainty grabber and truncation grabber*

A control that lets you use the mouse to change values and settings.

group interval

A subrange of a distribution that allows similar values to be grouped together and given a frequency count.

iteration *also trial*

A three-step process in which Crystal Ball generates *random numbers* for *assumption cells*, recalculates the spreadsheet model(s), and displays the results in a Forecast Chart.

kurtosis

The measure of the degree of peakedness of a curve. The higher the kurtosis, the closer the points of the curve lie to the *mode* of the curve. A normal distribution curve has a kurtosis of 3. (see "Kurtosis" on page 120)

Latin hypercube sampling

In Crystal Ball, a sampling method that divides an assumption's *probability distribution* into intervals of equal *probability*. The number of intervals corresponds to the Minimum Sample Size option available in the Run Preferences dialog. A *random number* is then generated for each interval.

Compared with conventional *Monte Carlo sampling*, Latin hypercube sampling is more precise because the entire range of the distribution is sampled in a more even, consistent manner. The increased accuracy of this method comes at the expense of added memory requirements to hold the full Latin hypercube sample for each *assumption*. (see "Latin hypercube" on page 114)

mean

The familiar arithmetic average of a set of numerical observations: the sum of the observations divided by the number of observations. (see "Mean" on page 116)

mean standard error

The *Standard Deviation* of the distribution of possible sample *means*. This statistic gives one indication of how accurate the *simulation* is. (see "Mean standard error" on page 121)

median

The value midway (in terms of order) between the smallest possible value and the largest possible value. (see "Median" on page 116)

mode

That value which, if it exists, occurs most often in a data set. (see "Mode" on page 117)

model sensitivity

The overall effect that a change in an *assumption cell* produces in a *forecast cell*. This effect is solely determined by the formulas in the spreadsheet model.

Monte Carlo simulation

a system which uses random numbers to measure the effects of uncertainty in a spreadsheet model.

outliers *also outlying values*

Values generated during a simulation on the extreme end of a distribution and are excluded from the display range.

PDF

Probability density function that represents the probability that an infinitely small variable interval will fall at a given value.

probabilistic model

A system whose output is a distribution of possible values. In Crystal Ball, this system includes a *spreadsheet model* (containing mathematical relationships), *probability distributions*, and a mechanism for determining the combined effect of the *probability distributions* on the model's output (*Monte Carlo Simulation*).

probability

(Classical Theory) The likelihood of an event.

Probability Distribution *also Distribution*

A set of all possible events and their associated *probabilities*.

random number

A mathematically selected value which is generated (by a formula or selected from a table) to conform to a *probability distribution*.

random number generator

A method implemented in a computer program that is capable of producing a series of independent, *random numbers*.

range

The difference between the largest and smallest values in a data set. (see "Range (also range width)" on page 121)

rank correlation ***also Spearman's rank correlation***

A method whereby Crystal Ball replaces assumption values with their ranking from lowest value to highest value using the integers 1 to N prior to computing the *correlation coefficient*. This method allows the distribution types to be ignored when correlating assumptions. (see "Rank correlation" on page 123)

relative probability ***also relative frequency***

A value, not necessarily between 0 and 1, that indicates probability when used in a proportion.

reverse cumulative frequency distribution

A chart that shows the number or proportion (or percentage) of values greater than or equal to a given amount.

risk

The uncertainty or variability in the outcome of some event or decision.

seed value

The first number in a sequence of random numbers. A given seed value produces the same sequence of random numbers every time you run a simulation.

sensitivity

The amount of uncertainty in a *forecast cell* that is a result of both the uncertainty (*probability distribution*) and *model sensitivity* of an *assumption cell*.

sensitivity analysis

The computation of a forecast cell's sensitivity with respect to the assumption cells.

skewed

An asymmetrical distribution.

skewed, negatively

A distribution in which most of the values occur at the upper end of the range.

skewed, positively

A distribution in which most of the values occur at the lower end of the range.

skewness

The measure of the degree of deviation of a curve from the norm of a asymmetric distribution. The greater the degree of *skewness*, the more points of the curve lie to either side of the peak of the curve. A normal distribution curve, having no *skewness*, is symmetrical. (see “Skewness” on page 119)

spreadsheet model

Any spreadsheet that represents an actual or hypothetical system or set of relationships.

standard deviation

The square root of the *variance* for a distribution. A measurement of the variability of a distribution, i.e., the dispersion of values around the mean. (see formulas in “Standard deviation” on page 117 and page 339)

subordinate

A relationship between distributions in which one distribution’s values for all percentile levels are lower than another’s. *(see also Dominant)*

trial *also iteration*

A three-step process in which Crystal Ball generates *random numbers* for *assumption cells*, recalculates the spreadsheet model(s), and displays the results in a Forecast Chart.

trial *as used to describe a parameter in certain probability distributions*

The number of times a given experiment is repeated.

value cell

A cell that contains a simple numeric value.

variable

A quantity that can assume any one of a set of values and is usually referenced by a formula.

variance

The square of the *standard deviation*; i.e., the average of the squares of the deviations of a number of observations from their mean value.

Variance can also be defined as a measure of the dispersion, or spread, of a set of values about a mean. When values are close to the mean, the variance is small. When values are widely scattered about the mean, the variance is larger. (see formulas in "Variance" on page 118 and on page 339)

virtual memory

Memory which uses your hard drive space to store information after you run out of random access memory. Virtual memory supplements your random access memory.

Index

In this index

A comprehensive index designed to give you quick access to the information in this manual.

Index

Index

Index

E

F

Index

G

H-K

Index

U

V

W

Index

Credits

Crystal Ball User Manual

Written and edited by Carol Werckman, Terry Hardy, and Eric Wainwright.

Parts written by and edited by Roy Sargent.

Layout design by Mike Marsh.

Illustrations and screen captures by Tracy Ford using Jasc, Inc. Paint Shop Pro.

Screen captures for the Windows version were performed using Desert color scheme.

This document was created electronically using Adobe® Framemaker® Release 5.5 for Microsoft Windows®.

Typeset using NewBskvll BT and Univers fonts.

Plate-ready film was output on the Scitex® Dolev 200 Imagesetter at 2540 dpi.

Blueline proof printed on the Omnivac® Theimer.

Cover design and production by Crossroads Communication, Nashua, NH.

Printing and binding by CGPress, Broomfield, Colorado.