RiskAMP User Guide

Structured Data LLC
www.riskamp.com
info@riskamp.com

Contents

Random Distribution Functions

The RiskAMP Add-in includes more than 40 random distribution functions. The parameters or arguments to these functions differ depending on the distribution and its requirements, but they generally adhere to standard statistical conventions. This section won't address all of them, but individual descriptions are available in the manual and Excel's function help will identify the required parameters.

Random distribution functions are included to provide source values for running a Monte Carlo simulation, but they are incidental to the simulation itself. Any Excel cell can be analyzed following a simulation, and any random value function – such as Excel's built-in random function – can be used to inject uncertainty (note, however, that Excel's function has a very short period).

Normal and Lognormal

The most commonly used distributions are the normal and lognormal. The RiskAMP function for the normal distribution, NormalValue, takes parameters for the mean and standard deviation of the distribution.

The basic lognormal function, LognormalValue, represents the standard lognormal distribution – that is, $e^{N(x)}$ where N(x) represents a normal distribution with a given mean and standard deviation. The parameters to this function represent the underlying normal distribution, and not the mean and standard deviation of the generated lognormal curve.

The add-in includes an additional function, ScaledLognormalValue, to sample from a lognormal distribution with a given mean and standard deviation.

The beta-PERT Distribution

The beta-PERT distribution is a special four-parameter version of the beta distribution, used to model expert data given most likely, minimum, and maximum likelihood values. This distribution is commonly used in project time and cost estimation. The PertValue function takes these values as parameters, and returns a sample from the generated beta-PERT distribution. The optional fourth parameter, the lambda value, represents the peakedness of the generated distribution, similar to kurtosis.

Spreadsheet Selection Functions

Several functions can be used to select values from ranges of cells. There are two functions for sampling uniformly from a spreadsheet range: ChoiceValue samples from a range with replacement, and UniformRangeSample samples from a range without replacement (although replacement will occur if the number of samples exceeds the range size). An additional function, WeightedChoiceValue, takes a range of choices and a range of weights, which affect the likelihood of selecting a particular cell value.

Multivariate Distributions

The RiskAMP Add-in includes several functions for sampling from multivariate distributions. The simplest function is CorrelatedNormalValue, which generates a single normally-distributed value which is correlated with another normally-distributed value. The function takes as an argument the coefficient of correlation, and generates a random value such that the two values, over the course of a simulation, will exhibit the given correlation.

For cross-correlated values, the MultiNormalValue generates a set of normally-distributed values, using a matrix to represent internal correlation among the sample values. This function takes arguments including the correlation matrix; the range of returned values; and optionally, a mean and standard deviation for each value to scale the returned sample. The MultiNormalValue is constructed using a Cholesky decomposition.

Simulation Analysis Functions

Simulation analysis functions are used to find statistics about, or particular values from, individual cells in the spreadsheet following a simulation. For example, if a spreadsheet cell contained a normal distribution, following a simulation these functions could be used to find the mean and standard deviation of the values in that cell during the simulation.

Retrieving Ranges and Percentiles

Two of the most useful functions for analyzing the results of a simulation can find the range of values matching a particular likelihood, or the portion of simulation results that fall within some bounds.

The SimulationInterval function can find the portion of simulation results that fall between a particular minimum and maximum value, or find the portion of simulation results that are above or below some value.

For example, if a particular spreadsheet cell contained a normal distribution with mean 0 and standard deviation 1, the SimulationInterval function could be used to find the portion of values above zero (expected to be 50%), or between -1 and 1 (expected to be 68.27%). If both a minimum and maximum value are provided to the function, it returns the portion between values; if one of the minimum or maximum are omitted, then the function returns the range falling above or below the provided value.

The SimulationPercentile function returns the value of a particular cell corresponding to a given percentile. For example, if a cell contained a normal distribution with mean 0 and standard deviation 1, the value at the 50th percentile is zero. In essence, this means that in 50% of the trials of the simulation, the value in the given cell was less than or equal to zero.

Similarly, if a cell value is normally distributed, the value at the 84th percentile should be roughly plus-one standard deviation from the mean, and the value at the 16th percentile should be roughly minus-one standard deviation from the mean. The SimulationPercentile function sorts results in order of increasing value. Therefore to find the values within 5% of the minimum and maximum, you would find the 5th and 95th percentile values, respectively.

Statistics Functions

Most of the analysis functions return some statistics about the values in a particular cell. Functions are available for finding mean (arithmetic or geometric), variance, standard error of the mean, and so on, from particular cells. Some functions reference multiple cells; for example, functions are available to find the correlation and covariance between two cells during a simulation.

Retrieving Individual Values

Individual values from the simulation can be retrieved for any cell in the spreadsheet. The SimulationValue function returns the value of a particular cell at a particular iteration of the simulation. Another function, SortedSimulationIndex, finds the inverse: it can be used to find the particular iteration number at which some value occurred.

The latter function is useful for selecting iterations to analyze. For example, if a particular spreadsheet model has several inputs and one output, you can find the minimum and maximum output value using the statistics functions (SimulationMin and SimulationMax). These values can be used with the SortedSimulationIndex to find the iterations within the simulation at which the minimum and maximum

values occurred; and then that iteration number can be used with the SimulationValue function to examine the values of each input cell in the minimum and maximum output cases.

Chart Functions

Several functions are used when the add-in generates automatic tables and charts. The SimulationHistogram and SimulationHistogramBin functions are similar to the SimulationInterval function, but are more flexible when values change significantly, because they use relative and variable interval sizes. In most cases, it's easiest to use the add-in's wizard to generate tables rather than use these functions manually.

General Simulation Information

Some additional functions return information about the simulation. The SimulationTrials function returns the number of trials from the most recent simulation, and the SimulationTime function returns the elapsed processing time of the most recent simulation.

Running a Simulation

The RiskAMP Add-in runs Monte Carlo simulations by calling Excel's calculation engine repeatedly, in sequence, and storing the results of output cells for later analysis.

To run a simulation, select "Run Simulation" from the Monte Carlo menu. This will launch the Simulation dialog, which controls the running simulation.

In the simulation dialog, enter the desired number of iterations. In most cases, it's useful to start from a small number of iterations to see if the model functions as expected. Simulations can be run as many times as is necessary.

The "Allow screen updates" checkbox controls whether or not to show calculations while the simulation is running. Depending on the complexity of the spreadsheet, turning off screen updates can dramatically improve performance.

Press "Start" to run the simulation, or "Cancel" to close the dialog. When the simulation is running, click "Stop" or press the "Escape" key to stop and reset the simulation.

Please note that Excel 2007 charts are very slow. In most cases, it's preferable to turn off screen updates when using Excel 2007 with charts.

Run Simulation Again

The menu option "Run Simulation Again (Same Settings)" runs a simulation without requiring you to click the "Start" button. The simulation will use the number of iterations from the last simulation. This is useful mostly as a shortcut when developing or testing a model, or when re-running a commonly-used model.

Re-Play Last Simulation

Just before running every simulation, the add-in stores the random number generator's internal state. The replay simulation command resets the random number generator to the state of the previous simulation, and then re-runs the simulation.

The effect of this will be that all random numbers sampled using the RiskAMP random functions will be identical to the last time the simulation was run. This is useful primarily for demonstration or presentation purposes, when a specific, known result is desired.

Please note that the re-play function will not work as expected when using multiple processors or multiple threads in Excel 2007.

Histogram and Chart Wizard

The Histogram and Chart Wizard can quickly generate tables and charts to display statistical and distribution information from a simulation. The wizard is available from the Monte Carlo menu.

Source Cell

The Source Cell page selects the cell used as a base for the generated tables and charts. This should be a cell that has varied over the simulation, such as a random variate or the output of some model based on random data.

Target Range

The chart wizard can be used to insert a table of values into any worksheet, by selecting a target range. The range can be either horizontal or vertical. If the selected range contains two columns (or two rows when creating a horizontal table), labels will be included. If the range contains only a single column (or row), labels will be omitted.

Simulation Results Sheet

Instead of using a target range, if the "Create a new simulation results sheet" is checked, the wizard will insert a new worksheet into the workbook with tables, charts, and some additional statistical data.

Chart Types

If a target range is selected (instead of a new results sheet), the wizard offers a selection of table/chart types. The resulting table will be inserted in the target range. The "Column (Row) Titles" checkbox controls whether to include a title in the first row of the table.

Random Number Generator

The RiskAMP Add-in uses a random number generator based on the Mersenne Twister algorithm, suited for running complex Monte Carlo simulation models with thousands of sample values generated per iteration.

Re-seed the Random Number Generator

To reset the random number generator, select "Advanced Properties" from the Monte Carlo menu. On the "Random Number Generator" tab, enter a new seed value or zero to use the system time as a seed. Then click "OK".

Latin Hypercube Sampling

Latin Hypercube Sampling (LHS) is an alternative method of sampling random values from random distributions, and is provided as an option in the RiskAMP Add-in.

In Latin Hypercube sampling, the random distribution is divided into a number of intervals of equal probability. When creating a sample value, the program first randomly selects one of these intervals (without replacement). A sample value from the underlying distribution is then randomly selected from within that interval.

Because the distribution is divided into intervals of equal probability, samples selected over the course of the simulation will conform to the given distribution. Because samples within each interval are randomly selected during the simulation, and because the order of intervals is randomly selected, each simulation will be unique. Taken together, these properties mean that using Latin Hypercube sampling when running a simulation can result in better distribution of random variates and lower error, with a smaller number of iterations.

To turn on LHS in the RiskAMP Add-in, select "Advanced Properties" from the Monte Carlo menu. On the "Sampling Method" tab, check the box for Latin Hypercube sampling to turn it on.

Number of Samples

The default behavior when using LHS is to use the number of simulation iterations as the number of samples. This helps to ensure broader distribution of samples over the course of a simulation. To change this behavior, select the "Fixed number of samples" radio button and enter a number. Note that if the number of samples is lower than the number of iterations, intervals in the LHS model will be re-used.

Locking Simulation Functions

Spreadsheets constructed using the RiskAMP Add-in contain functions and data that are only available if the add-in is installed. In some cases it may be useful to distribute spreadsheets to users who do not have the add-in available. The add-in includes functions for "locking" simulation functions, which replaces simulation functions with their output values. The spreadsheet can be "unlocked", which restores the original simulation and random number functions.

To lock simulation functions in a spreadsheet, select "Lock Simulation Functions" from the Monte Carlo menu. When this function has completed, the resulting spreadsheet can be saved and distributed to other users. Select the "Unlock Simulation Functions" command to restore the functions.

Please note that if you lock simulation functions twice in succession, the first set of functions will be lost. A warning dialog will be displayed if you attempt to do this.