

THE RISK AGENDA

Natural Monte Carlo User Manual

Natural Monte Carlo is a Microsoft Excel spreadsheet which contains built in functions and other routines to enable you to carry out simple risk modelling using the Monte Carlo method. We have found that *Natural Monte Carlo* works well with light to medium duty risk models which we have developed as part of our consulting activities. We think it is a lot more convenient for compiling these models and exploring the results than the heavy duty proprietary systems we also use.

The underlying concept is that instead of containing single values, spreadsheet cells can contain probability distributions. The distributions update automatically as the model is developed. Needless to say, as with any spreadsheet model, beyond a certain size it makes sense to switch automatic calculation off and then force a calculation when the model has been constructed.

This manual provides the specification of the new built in functions in Section 1 and describes how they can be put to use in Section 2. *Natural Monte Carlo* is supplied with a model set up which displays most of the capabilities. You can have a look at this and play with it, as a tutorial, before you delete it all and create your own models. You can create as many copies as you want: each model will run on its own as it will contain the necessary procedures. Section 3 contains a few hints and tips.

It is assumed that you are familiar with the basic principles of risk modelling and Monte Carlo. If not, there is more information on our website www.riskagenda.com. If you have any comments we would be glad to hear them at software@riskagenda.com.

1 Risk Modelling Built In Functions

There are three of these functions which are used to:

- define input distributions
- identify the cells for which the output distributions are required
- allow correlation to be modelled by more advanced users.

The Input Function ID()

This is a built in function which is used to set up an input distribution associated with a cell. The syntax, with optional arguments in square brackets, is:

=ID(Target, Ident, Name, DType, Parameters, [Scale], [Prob], [Units], [CC], [CDist])

| | | |
|--------|--------|---|
| Target | Cell | The location of the input distribution, which would normally be the address of the cell containing the ID() function itself. It will work if it is not, but this is potentially confusing for the user. |
| Ident | String | An identifier for the input distribution. |
| Name | String | The name of the input distribution. |
| DType | String | The type of input distribution which at present is one of Uniform, Normal, Triangle, Fixed or Data - see the Distribution Library appendix. |

THE RISK AGENDA

Natural Monte Carlo User Manual

| | | |
|------------|--------|--|
| Parameters | Range | The range contains the parameters of the distributions, the number of which depends on DType - also see the Distribution Library appendix. |
| Scale | Number | A scale factor which is applied to the distribution. Optional with default value 1. |
| Prob | Number | The probability with which the distribution is sampled; otherwise the value is 0. Optional with default 1. |
| Units | String | The units of the input distribution (used for drawing graphs etc). Optional with default value of "None". |
| CC | Number | Correlation factor -see section on correlation below. Optional with default value zero. |
| CDist | Cell | The cell which contains the correlating distribution - see section on correlation below. Optional. |

Commas should be used where there are optional arguments except that any trailing commas can be omitted.

When the ID() function is evaluated (for example when first input into the spreadsheet) a probability distribution is calculated and held in memory. The value returned is the mean of the distribution (as estimated from the random sample, not the exact mean).

The distributions are held in memory and are not physically stored in the spreadsheet. This means they must be recreated when the spreadsheet is opened, for example.

The Output Function OD()

This is a built in function which is used to identify an output distribution associated with a cell. The syntax, again with the optional argument in square brackets, is:

=OD(Target, ID, Name, [Units])

| | | |
|--------|--------|--|
| Target | Cell | The location of the output distribution, which will normally contain a formula which is the last stage in the calculation of the output from the inputs. Since a cell cannot contain both a formula and the OD() function, the target cannot be the location of the function itself. This contrasts with ID(). |
| Ident | String | An identifier for the output distribution. |
| Name | String | The name of the output distribution. |
| Units | String | The units of the output distribution (used for drawing graphs etc). Optional with default value of "None". |

Again, when this function is calculated, the corresponding distribution is created in memory. The value returned is simply confirmation of Target, the location of the output.

THE RISK AGENDA

Natural Monte Carlo User Manual

The Correlating Function CD()

The Risk Agenda supports the basic principle that correlation is best modelled explicitly, rather than being inserted through a correlation coefficient. Nonetheless we recognise that this functionality can be useful, and we supply it through the concept of correlating distributions. These are simple underlying distributions and two (or more) input distributions can be correlated by correlating them both with one of the correlating distributions. (This comes close to our underlying philosophy of explicit correlation modelling.)

To apply correlation, a correlating distribution is defined using CD(). The final two arguments of the ID() function are then used to determine the degree of correlation (CC) and the relevant correlating distribution (this final argument can be omitted if there is only one such distribution).

Without going into detail, the method ensures that a particular value of the rank correlation coefficient is obtained. This is approximately the product of the two CC values. If you need the exact value you can get it by using the built in function RCC(CC1,CC2).

The correlating distribution function has syntax

=CD(Target, Ident, Name)

| | | |
|--------|--------|---|
| Target | Cell | The location of the correlating distribution, which would normally be the address of the cell containing the CD() function itself, as for ID(). |
| Ident | String | An identifier for the correlating distribution. |
| Name | String | The name of the correlating distribution. |

2 Putting the Functions to Use

The risk model is constructed in a worksheet called *Model*. You cannot change the name of this worksheet which should contain all the CD(), ID() and OD() functions. There are two sets of buttons on the left hand side of the sheet, one for control of *Natural Monte Carlo* and the other for output.

Control Buttons

| | |
|-----------------------|---|
| Force Recalculation | Re-samples and calculates the whole model. It is always best to do this before using the output. |
| Automatic Calculation | Switches Excel to automatic calculation mode. This should allow the model to be kept constantly up to date, but for larger models (and smaller computers) at the expense of waiting for the calculation to complete each time something is changed. |
| Manual Calculation | Switches Excel to manual calculation mode. Force Recalculation then recalculates the model when you want to. This is the default mode on opening the workbook. |

THE RISK AGENDA

Natural Monte Carlo User Manual

| | |
|------------------|---|
| Hide Diagnostics | This hides the diagnostic messages: messages are restricted to errors. This is the default setting on opening the workbook. |
| Show Diagnostics | This shows a series of messages as each element of the model is calculated. Should not be necessary for the general user. |

Output Buttons

| | |
|-----------------|--|
| Generate Report | This provides a text report on the model. This is based on the sample values, matching them up with the spreadsheet formulas. You will therefore get warnings if these are not consistent, for example if the model has not been sampled at all. If this is the case do a forced recalculation. |
| Draw Graphs | This provides a simple dialog box from which the input and output distributions can be selected and various graphs can be drawn. The available types are S-curves, sensitivity charts, scatter diagrams and approximate probability density functions. |
| Show Graph Data | If for any reason you want to extract data from <i>Natural Monte Carlo</i> you can do so by using this button which exposes the graph data sheet. If you have just drawn a scatter chart it is the raw simulation data; if S-curves/sensitivity it is ordered (thereby losing the data for each simulation); if PDFs it is the approximate PDF data. |

3 Hints and Tips

1. *Natural Monte Carlo* uses the Excel status bar to provide messages about what is going on.
2. *Natural Monte Carlo* always uses 1000 samples to define the probability distributions.
3. Use the report button to check your models. Any anomalies will show up here - see tips on error messages below.
4. The Scale argument in ID() is not strictly necessary. You could get the same effect using the Parameters, but it is something we find useful.
5. Be careful when you are using correlation with a Prob value not equal to 1. The correlation is applied to the whole distribution, including the zero which results with probability 1-Prob. *Natural Risk Manager* assumes this is the lowest ranked part of the overall distribution, that is, the non-zero parts are positive. If you have the opposite situation you will need to reverse signs.
6. (Especially in connection with the previous point) always look at the scatter charts when you use correlation to check it is providing the desired effect.
7. Although we noted that the model should be confined to the *Model* worksheet, you can acquire data from other worksheets (or workbooks) if you wish. However all cells containing the CD(), ID() and OD() function calls should be in the *Model* worksheet and data should not flow from the *Model* worksheet to other locations and back again. This is

THE RISK AGENDA

Natural Monte Carlo User Manual

because the Monte Carlo simulation do not use the *Model* sheet, but another sheet which mirrors it. (This sheet is called *Mirror* if you really need this functionality.)

8. When you delete CD(), ID() and OD() functions from the model, the associated distributions remain in memory until the next recalculation. This can give surprising results in that they can be charted after having been deleted from the model.
9. Don't forget that changing the calculation mode using the buttons will affect the application as a whole, and therefore the recalculation behaviour of any other open workbooks. On exit *Natural Monte Carlo* changes the calculation mode back to what it was before it was opened. This may or may not be what you intend.
10. The CD(), ID() and OD() functions have fairly complicated arguments. If you get them wrong, particularly if you have the wrong number of arguments or you mix up range and value arguments, Excel itself will generate #VALUE! error messages. *Natural Monte Carlo* also traps other errors, either DType Err if the distribution type is not recognised, warning messages if the parameters are anomalous and an Error message for other issues. The report generator also creates an error message if the Target arguments are not correctly specified.
11. However, because *Natural Monte Carlo* uses the underlying Excel sheet calculation methods with built in functions some controllability is lost compared with a situation where programme control is totally in the hands of the software developer. This may mean other errors can arise and information may be lost from memory, though not from the model itself. If this happens and the control buttons do not recover it, save the model and close Excel. Then reopen and check your model carefully using the report button.

THE RISK AGENDA

Natural Monte Carlo User Manual

Appendix: Distribution Library

At present there is a relatively small number of basic distributions available. If there is demand this can be expanded. Each of the Sections below sets out the DType keyword and describes the parameters. Note that as the parameter argument for ID() is a range you can provide the range in any form which fits the number of parameters. For example for a normal distribution you have range which is either one column with two rows or one row with two columns. The cells in a range are counted off by columns and then by row.

Uniform

Two parameters: (1) is the minimum and (2) is the maximum.

Triangle

Three parameters: (1) is the minimum, (2) is the most likely and (3) is the maximum.

Normal

Two parameters: (1) is the mean and (2) is the standard deviation.

Data

The parameters are any number of probability impact pairs: (1) is zero, (2) is the minimum impact, (3) is some other probability, (4) is the associated impact, (n-1) is one, (n) is the maximum impact. This would normally be encoded on the spreadsheet with an array of two columns, for example:

| | |
|-----------|---------------------|
| $p_0 = 1$ | x_0 - the minimum |
| p_1 | x_1 |
| p_2 | x_2 |
| p_3 | x_3 |
| $p_4 = 1$ | x_4 - the maximum |

In principle: $0 = p_0 \leq p_1 \leq p_2 \leq p_3 \leq p_4 = 1$; $x_0 \leq x_1 \leq x_2 \leq x_3 \leq x_4$. The distribution is uniform between each of the x_i values.

Fixed

One parameter: (1) the single, fixed value. This would generally be used with a probability to represent a one-off event.