SIPmath™

MODELER TOOLS FOR EXCEL

REFERENCE MANUAL

DISTRIBUTION PROCESSING IN EXCEL

Sam Savage (savage@stanford.edu)

Marc Thibault (marc@smpro.ca)

Rev. 2012-10-30

SIPmath Modeler Tools Version 1.0

TABLE OF CONTENTS

1 INTRODUCTION

The SIPmath™ Modeler Tools for Excel facilitate the development of Distribution Processing applications, based on SIP Libraries or interactive random simulations with the Excel RAND formula. Once created, such applications will run on their own within Excel.

Probability Management is a decision support framework for dealing with uncertainties much in the way that traditional data management deals with numbers (see www.ProbabilityManagement.org). It is based on three principles[i]

1 **Distributions as Data**. Probability distributions are represented as data. Specifically, as Stochastic Information Packets or SIPs. These contain hundreds or thousands of simulated or historical realizations of the distributions in questions.

   A central goal of Probability Management.org is to promote standards for formatting libraries of SIPs (Stochastic Libraries) along with metadata relating to summary statistics and provenance.

2 **Distribution Processing**. Distributions stored as SIPs may be used in interactive calculations (using an approach we call SIP math) to yield output distributions also stored as SIPs.

   A SIP contains realizations of a distribution

   SIP math may be performed in Microsoft Excel using nothing but the Index formula and Data Table. The SIPmath Modeler Tools for Excel from ProbabilityManagement.org facilitate this process by automating several of the steps. In addition they may be used to create interactive random simulations based on the RAND formula in Excel.

3 **Managerial Accountability**. There must be managerial accountability for the provenance of the Stochastic Libraries. We refer to this function as the Chief Probability Officer (CPO).

Stochastic Libraries – The basic elements in a stochastic library are vectors of possible values of an uncertainty (think rows or columns in Excel, or sequential elements in a data base). That is, instead of representing the future demand for a product as a single "best guess," it's represented by 1,000 or 10,000 possible future demands. We refer to the set of values for a single uncertainty as a Stochastic Information Packet or SIP, and calculation with SIPs as *SIP math*.

Coherence – When there are multiple uncertainties, it's important to make sure they are coherent. For example, consider two SIPs of 1,000 trials, one representing the value of your house, and the other representing the value of your fire insurance policy. If there is one chance in 1,000 that your house will burn down, then one element of the house SIP is zero, while the other 999 take on its the current value. The Fire Insurance SIP would have one element in which it pays out to rebuild the house, while the other 999 contain the negative of your insurance premium. These two SIPs are said to be coherent, if the payoff on the insurance SIP lines up with the fire in the house SIP. A coherent set of SIPs is also known as a Stochastic Library Unit with Relationships Preserved or SLURP. If the house and fire SIPs were not coherent, it would be as if you had bought fire insurance on a stranger's house.

A SIP might be formatted in any of several data representations. At the very least it contains a list of values drawn from the possible values of the uncertain variable. In most cases, a SIP will also include metadata about the distribution's statistical attributes and about its provenance.

Distribution Processing illuminates the consequences of uncertainty and cures the Flaw of Averages[ii] without the need for statistical training. From the perspective of the Excel user it adds a third dimension to your worksheet representing thousands of potential outcomes. It is an outgrowth of a technique known as Monte Carlo simulation in which random inputs are run through spreadsheets or other modeling environments, while monitoring the resulting outputs. There are three primary differences between Distribution Processing and earlier forms of simulation.

See Distribution Processing and the Arithmetic of Uncertainty at www.analytics-magazine.org

First, Distribution Processing does not involve the generation of random numbers (although the **SIPmath** Modeler Tools also include this capability). Instead it draws its values from libraries of Stochastic Information Packages (SIPs). SIPs hold arrays of pre-computed trial values created by statistical experts and subject matter experts using traditional simulation or real-world data. SIPs also provide for data provenance, an audit trail, and reproducible results. We refer to calculating with SIPs as SIP math.

Second, unlike most simulation methods, Distribution Processing with SIP math may be performed directly in Microsoft Excel without add-ins or macros. SIP math uses the same operations as numerical calculation, and provides the same sort of immediate feedback. Instead of outputting numbers, however, SIP math outputs SIPS from which you may derive statistics, including the chances that you will "make your numbers."

See SIPmath Tutorial.xlsx at www.ProbabilityManagement.org

Third, Distribution Processing enables aggregation. If various Distribution Processing applications are based on the same set of coherent input SIPs the results may be rolled up through simple addition with SIP math.

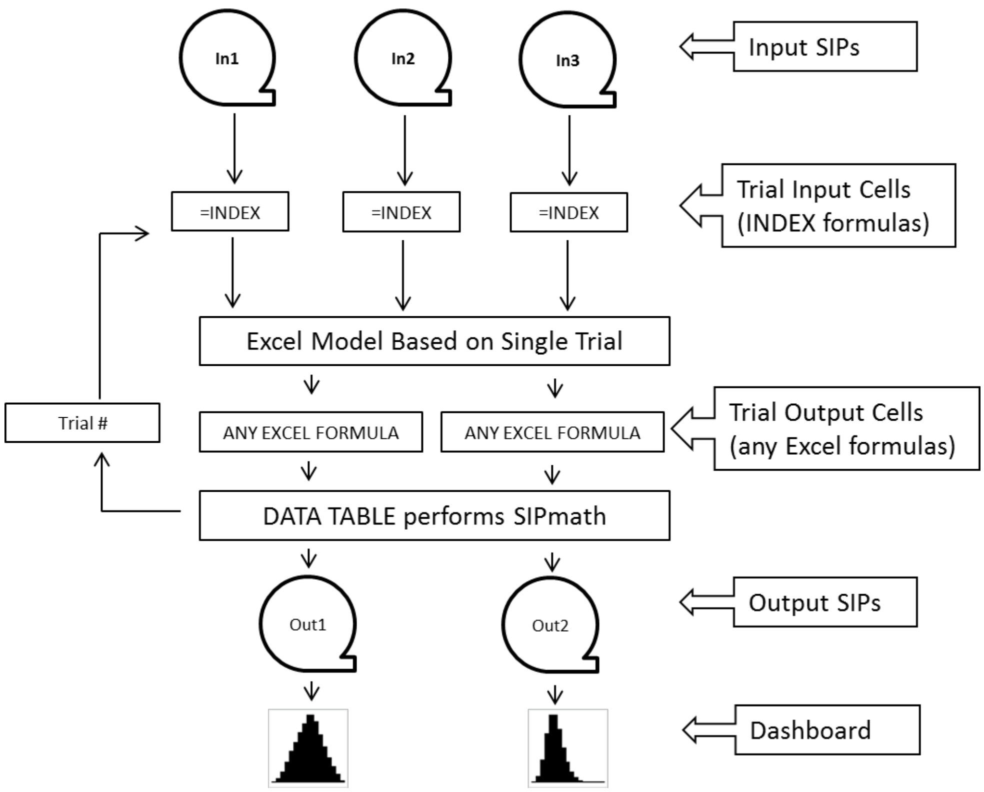
The SIPmath Modeler Tools facilitate the creation of Distribution Processing applications in Excel. They may also be used to create interactive random simulations. When SIPmath.xlam is loaded, a SIPmath ribbon appears in Excel and presents a set of tool buttons. These are used to open stochastic libraries, specify inputs and outputs, and assist in creating graphs. The resulting models perform distribution processing with raw Excel; no add-ins are needed to run the models or to interact with them.

These tools facilitate the creation of applications, but are not required to run them.

2 SIP MATH IN EXCEL ARCHITECTURE

2.1 OVERVIEW

Distribution Processing Model



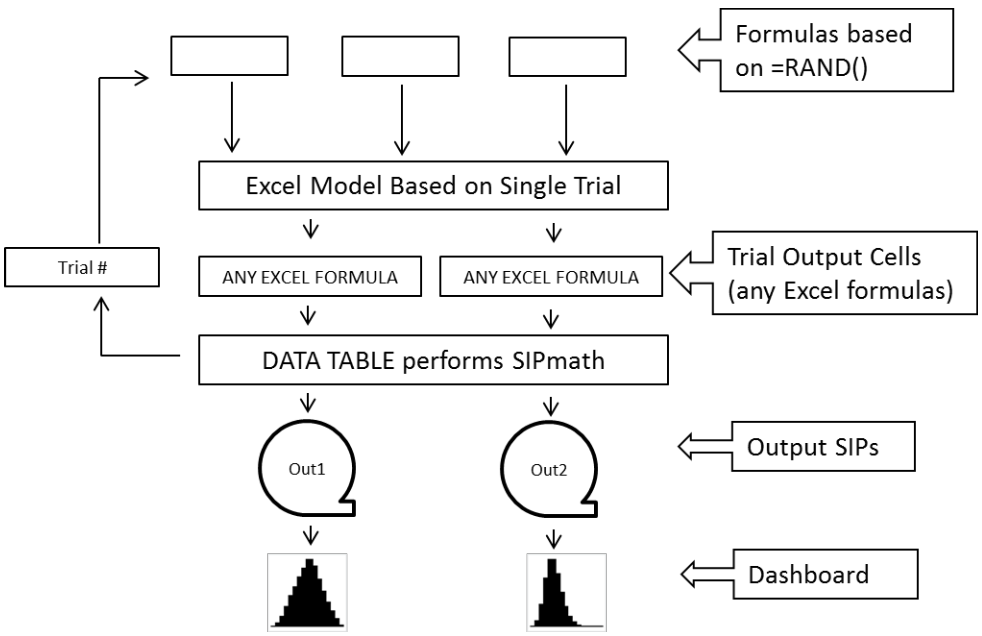
SIPs In

SIPs Out

Distribution Processing with SIP math supports a model architecture that uses Excel's *Index* function and *Data Table* tool to step the model through the trials of input SIPs and collect the results into output SIPs. This may be accomplished entirely within Excel without macros or add-ins. The SIPmath Modeler Tools are a set of macros that merely facilitate this process by automatically creating a Data Table on a sheet called PMTable, and assisting in setting up the inputs, outputs, graphs, etc. The model calculations are formulated as if the inputs and outputs are simple numerical cells in Excel. Then the Excel Data Table causes the simulation to run the model for the number of iterations stored in the PM_Trials cell of the stochastic library. Data Table computations are very efficient, and for medium sized models, thousands of calculations may be performed nearly instantaneously.

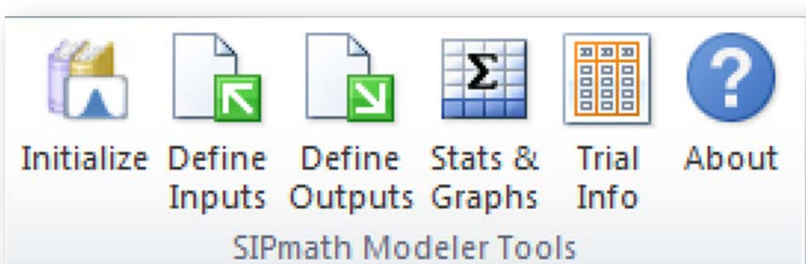
Random Mode

When applied in random simulation mode, no stochastic library is needed, and expressions based on the RAND() formula replace the INDEX formulas as shown below.



NOTE that for interactive models in either mode, Excel's calculation mode must be set to automatic.

2.2 THE SIPMATH RIBBON



SIPmath Modeler Tools

There are five tools on the **SIPmath** Ribbon:

The *Initialize* tool is where the modeler identifies the Stochastic Library containing the Input SIPs for use in Distribution Processing mode and specifies the number of trials to run if creating a model to run in Random mode.

The *Define Inputs* tool is where the modeler identifies the model's input cells and links them to the desired input SIPs. This tool is not required in Random Mode.

The *Define Outputs* tool is where the modeler identifies the model's output cells and links them to the data table to create the Output SIPs, which are created in either mode.

The *Stats & Graphs* tool provides an easy way to create graphs and statistics from the output SIPs in either mode.

The *Trial Info* tool provides a simple way to step through the input SIPs one trial at a time or view optional metadata in Distribution Processing mode only.

2.3 STOCHASTIC LIBRARIES

Input SIPs come from libraries

Input distributions come from *Stochastic Libraries* – Excel worksheets containing SIPs and SLURPs (coherent sets of SIPs). There are some specifics that must be met for the **SIPmath** Modeler Tools to use them:

1 There must be a named cell, PM_Trials, containing the number of trials per simulation. The **SIPmath** tools will use this to set up the data table on the PM_Table sheet. (Note when the SIPmath tools are applied in random mode, even if no stochastic library is used, a PM_Trials cell must still be present in the model to specify the number of trials).

2 The SIPs can be in rows or in columns, but each one must be defined as a named range.

3 Optionally, there can be metadata Index Values appended to a SIP's trial values. These are useful for observing the results of your models under specified scenarios, such as averages values of the inputs. If used, there must be a range named PM_IV containing the names of the metadata elements, and another named

PM_IV_Index with the displacements to the metadata. These displacements must be common to all the SIPs in the library. For example, assuming a PM_Trials value of 1000 the two ranges might appear as below:

PM_IV_Index	PM_IV
1001	Average
1002	Median
1003	Minimum
1004	Maximum

These names will be used by the *Trial Info* dialog to let the modeler see model instances other than specific trials. The named ranges for the input distributions should include the PM_IV elements so that *Trial Info* can use the range name with an INDEX() formula to find them. The values of the metadata are not calculated automatically, and must be specified when creating the library.

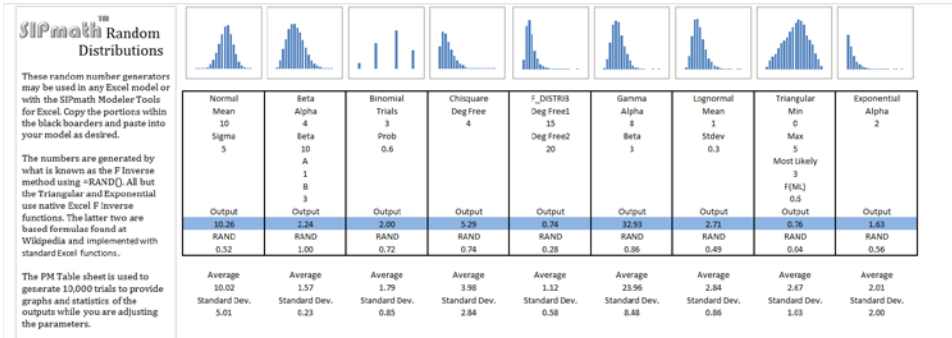
2.4 INPUT CELLS

Input cells interface between the input SIPs and the model

In Distribution Processing mode, the interface between the model and the SIPs in the library is a pair of ranges – one for the SIP names and another for the SIP trial values as specified by the user with the *Define Inputs* tool. Each input distribution must input value cell and optionally one input name cell. **SIPmath** Input tool will initialize the value cells with INDEX() formulas pointing to the first trial in the SIPs. As the data table tool increments PM_Index, these will evaluate to successive values from the corresponding input SIPs. The model refers to these cells to get its input data. In Random mode, the Index formulas are replaced with formulas using the RAND() formula, and the input cells do not require names.

2.5 PARAMETRIC DISTRIBUTION INPUTS

The *Random Distribution Generators.xlsx* workbook included with **SIPmath** Modeler Tools provides a number of parametric distributions, which may be used to either create stochastic libraries or drive random simulations.



To use these in your model in random mode, simply copy and paste the appropriate rectangular region into your worksheet. If you want replicable results, create a small model in random mode, specifying the distributions as outputs, and save the resulting SIPs.

2.6 OUTPUT CELLS

Output cells interface between the model and the data table

With the *Define Outputs* tool, the modeler will specify the cells that will evaluate to the results of the model calculation at the end of each iteration. These results will be written into the data table, collectively making up the output distributions. The use of this tool is the same in both modes. However, in Random mode, the output SIPs will change every time a calculation is performed, or the F9 Key is pressed.

2.7 THE DATA TABLE

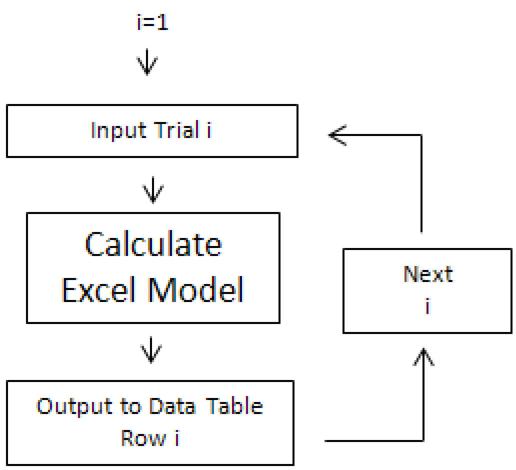
SIPmath sets up the data table on a worksheet called `PMTable` as part of the *Open Library* action. This includes defining `PM_Index` to be used as the trial number by the input cell formulas.

The data table indexes the input and records the output.

It also initializes the input column of the data table with a count from 1 to `PM_Trials`. These will be fed, one at a time, to `PM_Index` which in turn will be used by the input cell formulas to select specific trials from the input distributions.

The *Define Output* tool completes the picture by setting up the output column or columns of the data table. It sets the formula driving the Data Table to a simple assignment that copies the model output cell.

2.8 THE PROCESS



The Simulation Loop

A sheet recalculation event causes the data table to sequentially copy all the numbers from 1 to PM_Trials into PM_Index. This results in re-evaluating the model for each of the trials in the stochastic library. The Data Table has been optimized to do this very efficiently; the process is nearly instantaneous for medium sized models.

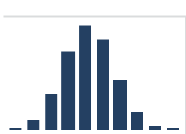
The results appear in the corresponding rows of the output columns. Because these calculations are all internal to Excel they are much faster than would occur if the numbers were pasted into Excel by a macro.

The process ends when the number of iterations specified by PM_Trials is reached.

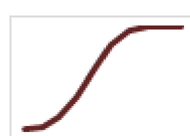
2.9 GRAPHICAL PRESENTATION

The *Stats & Graphs* tool provides an easy way to create histograms and cumulative probability graphs with the simulation data. It assumes a bunch of defaults to keep the amount of work you have to do to a minimum.

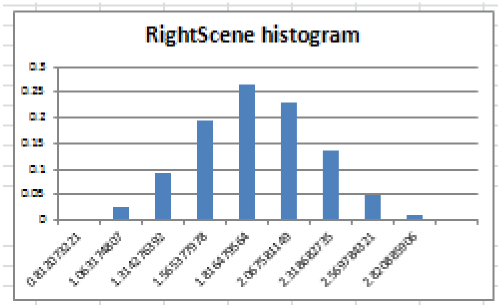
Sparklines



If you have Excel 2010 or later, you can create sparklines – Edward Tufte's tidy little graphs that sit in one cell.



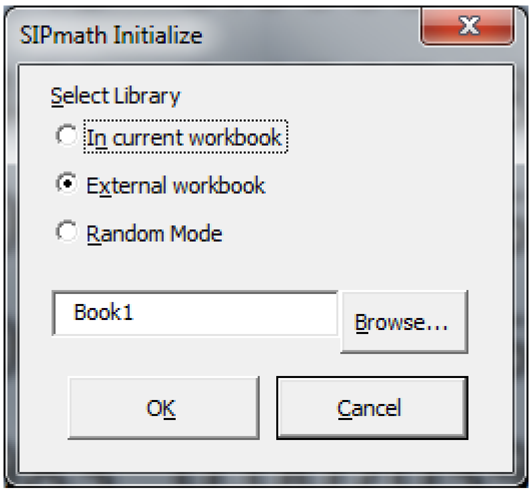
Excel Charts



You can also create the usual Excel charts. The automatic formatting is fairly basic, so you'll need to do some touch-up.

3 THE SIPMATH RIBBON DETAIL

3.1 INITIALIZE



The *Initialize* tool must be used to create a model in either Distribution Processing or Random mode. In Random mode the *Initialize* tool queries the user as to the number of trials, and the location of the PM_Trials cell which will contain them. In Distribution Processing mode the input distributions are assumed to be named ranges in an Excel Workbook. Use the *Initialaize* tool to tell **SIPmath** Modeler Tools where to find the distributions the model will be using for input. They can be in the current workbook or some other workbook. To specify another workbook, click **Browse** and use the file dialog to select the library file.

Get a stochastic library

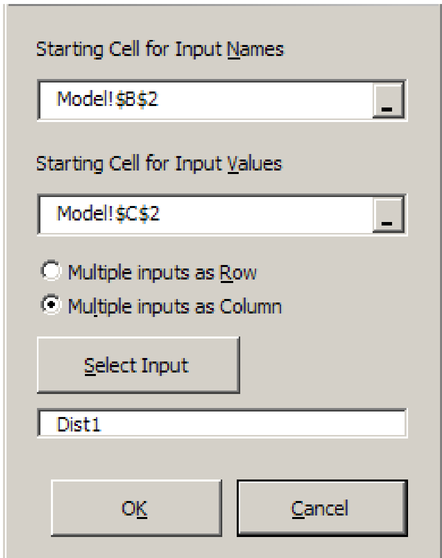
Make sure the distributions' named ranges include any metadata extensions specified in a PM_IV table.

The model will be linked to the library file at this stage. If you want to distribute a model widely you may want to place the library in the same workbook as the model. This can be accomplished by opening an existing library, right clicking on the sheet tab then using the Move or Copy command to copy it to the desired workbook.

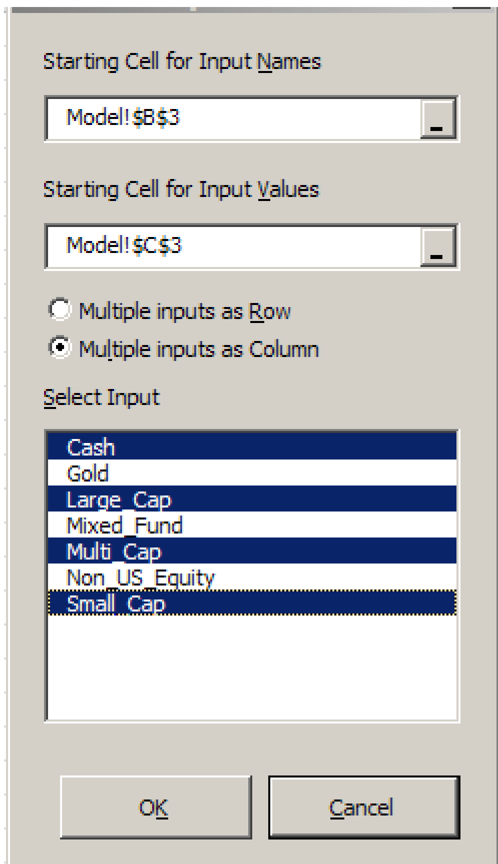
NOTE: When returning to work on an incomplete model linked to an external library, OPEN THE LIBRARY FILE BEFORE the incomplete model to properly re-establish the links. This is not an issue for models that contain internal libraries.

3.2 DEFINE INPUTS

For use in Distribution Processing mode only. Use the *Define Inputs* tool to tell **SIPmath** Modeler Tools which distributions to use from the library chosen during initialization, and which cells the model will use to get input trial values.



Click on **Select Input** to get the distribution selection dialog:



Link the library distributions to the model

Use the shift and control keys in the usual way to select multiple distributions. Click **OK** when you’re done. Notice that you select the distributions by name; this is why the library specification calls for named ranges.

If you’ve added distributions from other sources (e.g. the Random Distribution Generators workbook) and given them range names, those names will be included in the input choices.

The “Starting Cells” assume multiple distributions. If you select “Multiple Inputs as Row” the cell will be the leftmost cell. If you select “Multiple Inputs as Column”, the cell will be the top cell.

Multiple Inputs as Row:

Name1	Name2	Name3
.982736894	.35674313	.6546569

Multiple Inputs as Column:

Name 1	.982736894
Name 2	.35674313
Name 3	.6546569

The values in the table are from the SIPs in the library, usually the first trial; the PM_Index default value is 1.

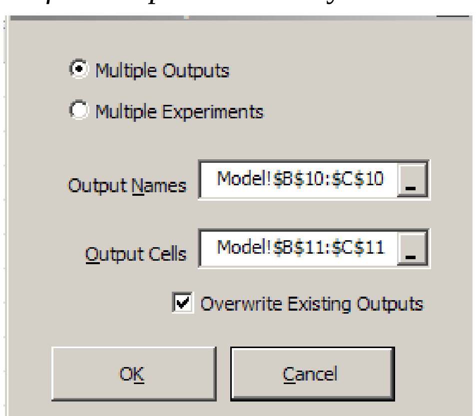
If you look at the formulas under the values, you'll see that the *Define Inputs* tool has set them to something like

=INDEX(Library01.xlsx!Distribution3, PM_Index)

Where Library01.xlsx is the library file name, Distribution3 is the distribution name and PM_Index selects the element in the distribution's range.

3.3 DEFINE OUTPUTS

Define Outputs is actually two different dialogs, depending on how your



model is structured, but works the same way in both Distribution Processing and Random mode.

Link the model outputs to the data table

The first, selected as "Multiple Outputs", is straightforward; the simulation runs once and generates one or more distributions as outputs. This dialog lets you specify names for the output distributions and the cells containing the formulas to be simulated.

Define Outputs will put a named output column range in the data table for each cell and associated name that you specify. These names can then be used to refer to the output distributions for creating graphs and summary statistics anywhere in your model.

If you check “Overwrite Existing Outputs,” *Define Outputs* will create a new table, otherwise it will append your selections to the right of the existing output columns. This means you don’t have to specify all your outputs in one go. Also, the output cells don’t have to be contiguous.

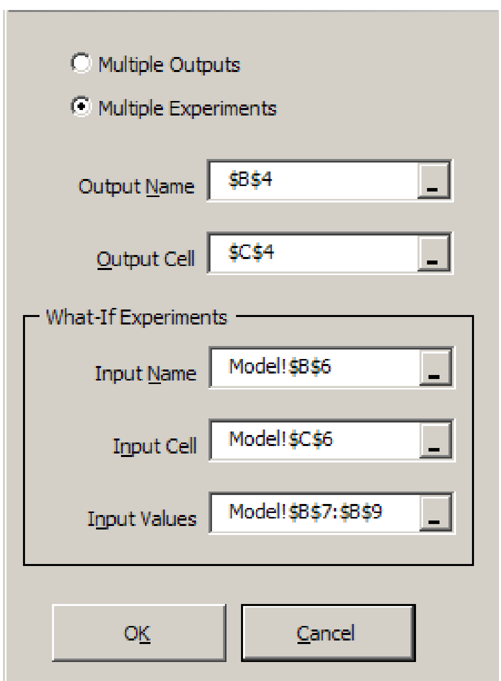
You can add to the existing outputs.

What-If

You get the second dialog by selecting “Multiple Experiments.” This is for “What If” experiments where you take a numeric cell in your model and experiment with different values to see what happens to the output.

Note ‘output’ –singular. The output name and output value cells work the same way as the previous case – specifying where to get the model results and what name to give the output columns in the data table. The difference is that there will be an output column for each of the experimental values specified. So you’ll have one output cell but a different output distribution for each of the experiment input values.

Test your assumptions with What-If

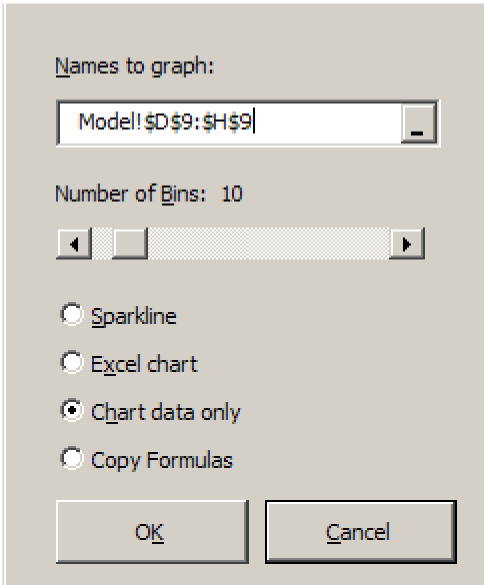


Note that the experiment input values are not the same as the input distributions you specified in the “Define Inputs” dialog, which will still drive each simulation. The “What-If” values will drive multiple simulations and the output simulation columns are the distributions of the result of each experiment.

The importance of the names is that *Define Outputs* will concatenate the output name, experiment input name and experiment input values to give each output column in the data table a unique range name you can use for graphs and summaries.

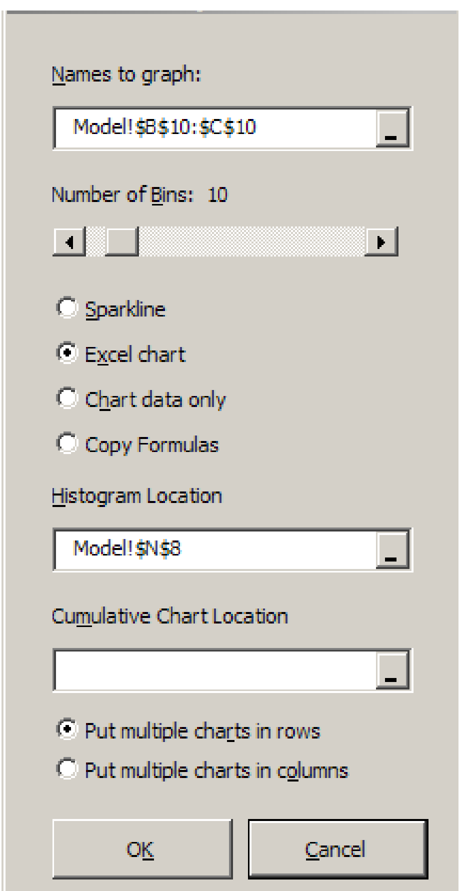
3.4 STATS & GRAPHS

This tool performs the same function in both Distribution Processing and Random modes. The opening dialog for Stats & Graphs lets you choose one or more distributions to plot. You identify the distributions with references to their *names*, not the *output cells* that defined them. Suppose for example that we want to graph distributions called Output1 through Output4. Typically these names will have been typed into your worksheet before the Define Outputs stage. Here we assume that they are in cells D9:H9. The number of bins field sets the number of histogram bars and the number of points on the cumulative curve.



Graphs present clear results

To make a chart, the *Stats and Graphs* tool must first set up a small table with a row for each bin. The columns are the bin values and the percentiles that, respectively, become the X and Y axis inputs to the graph.



Selecting “Chart Data Only” causes *Stats and Graphs* to make the table without creating a chart. You can then use this to create your own charts using Excel’s *Insert Chart* tool. The table can be found below the last values in the data table in suggestively named ranges (e.g. Profit_bins, Profit_freq).

If you select “Sparkline” or “Excel Chart” the dialog will expand to ask you which graphs you want and where you’d like them put.

The “Copy Formulas” option is useful for entering formulas for summary statistics, such as AVERAGE or STDEV.

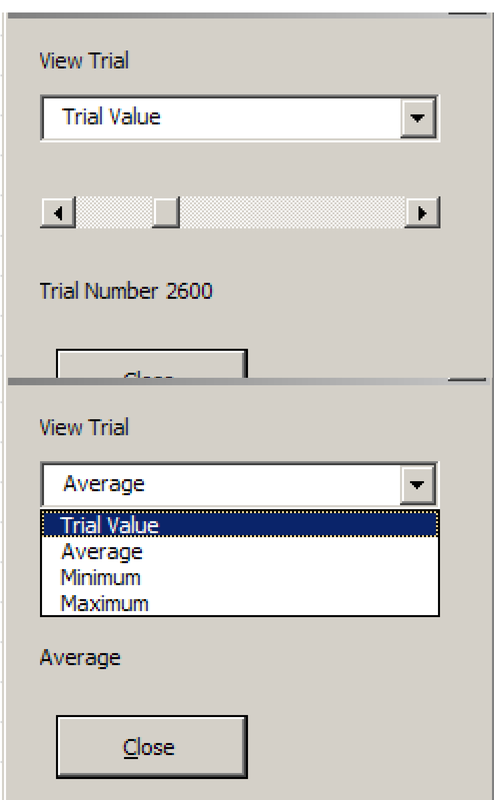
Assume your outputs are data per region, and the output names are East, Central, West, Canada. To get the mean of East, you can simply type =AVERAGE(East). But what if you want the mean for each of the four outputs? Since East is a range name, you can’t simply copy|drag the formula to get AVERAGE(Central).

Instead:

1. Set up one or more formulas referring to East.
2. In “Names to Graph” define the range of names you want to use in the formula (probably the row of names over the data).
3. Click the “Copy Formulas” radio button.
4. In “Formulas to Copy” put the range containing your formula(s) referring to East.
5. Select whether you want the copy to go down a column or across a row and click “OK”.

Stats and Graphs will replicate your formulas, replacing ‘East’ with each of the other names you selected.

3.5 TRIAL INFO



The *Trial Info* tool is only applicable in Distribution Processing mode. It lets you step through the model simulation trials one at a time.

Wherever you set the trial number, you can see the corresponding values in the input cells and the corresponding result values in the output cells.

If you’ve set up metadata extensions to the distributions, you can also use those as model instances. Keep in mind that the choice is only for your inputs. For example, if you choose “Average”, you’ll see the outputs given average inputs, not average outputs.

If your spreadsheet model is non-linear, you will find an interesting difference between the *output* given *average* inputs and the *average output*. This is due to Jensen’s Inequality, or what Sam Savage calls the Flaw of Averages.

Keep in mind that this tool won’t be there when the model is run without the SIPmath Modeler Tools add-in. If you want your user to be able to step through the trials, you’ll need to expose the `PM_Index` range and include a control to access the metadata at the ends of the SIPs if desired.

4 EXERCISE: BUILD A DISTRIBUTION PROCESSING MODEL

4.1 THE OBJECTIVE

We will now demonstrate how to build a distribution processing application using Excel and SIPmath Modeler Tools. The application can be run, without the use of macros, in Excel 2007 or later. You can either start with an existing model and wrap distributions and simulation around it, or as we describe, start with a blank workbook, and build a model with input and output distributions.

You're going to build a financial portfolio simulation so that you can compare the expected risks and returns for a variety of investment portfolio choices. You'll present these as a graph plotting the mean against the variance of the return for each portfolio.

A financial portfolio assessment application

In financial circles, variance is a proxy for risk, so the best portfolio is the one with the greatest return for the smallest variance. Unfortunately, risk and return are trade-offs, so there isn't just one "best". That's why you need a picture.

You'll simulate five different portfolios and be able to compare them while making changes to their mix of assets.

We could have you start by developing a model with single value inputs and outputs, then test and debug it before wrapping it in stochastic inputs and outputs. In this case, the model is pretty simple, so you'll start with the inputs and work your way to the outputs and graphical presentation.

Start by opening a blank workbook in Excel. In our example we renamed "Sheet1" to "Model". Give the file a name of your choice.

4.2 DEVELOPER NOTE

Open the library first

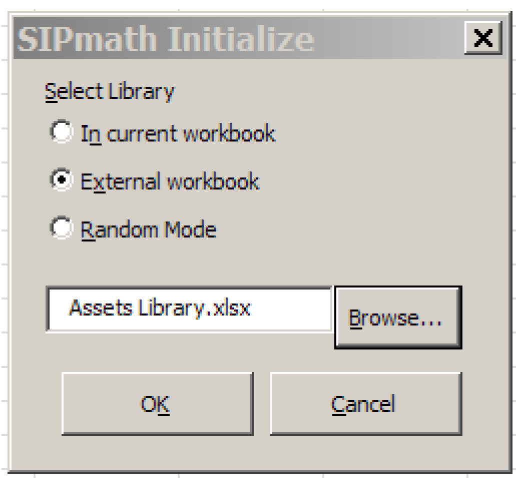
If you're using an external library, and developing the application involves multiple sessions, you **must** open the library workbook before you open your application workbook to continue development. This is NOT required when you are starting a model from scratch, or have a library built into your model.

4.3 INPUT DISTRIBUTIONS

The first thing you need is a source of data about assets you can invest in. The data you want is probability distributions of their annual returns. In a real-world application, these would come from a *Stochastic Library* maintained by a financial organization with the appropriate credentials. Since this is just an exercise, you'll use the Stochastic Library workbook that comes with the **SIPmath** Modeler Tools.

To get a stochastic library, you use the **Initialization** tool on the **SIPmath** ribbon. When you click the Open Library button on the **SIPmath** ribbon, you are prompted to open a stochastic library file.

Input distributions come from stochastic libraries



Instead of linking your model to an external Library file, you have the option to use a stochastic library within the current workbook. This is used for creating self-contained models, not linked to other files, to distribute as a single Excel file. You can easily copy a worksheet from a desired library to the application workbook with the sheet Move or Copy command in Excel.

For this exercise, click on **Browse** and pick the file named *Assets Library.xlsx* that came with the **SIPmath** Modeler Tools.

When you click **OK**, *Initialize*

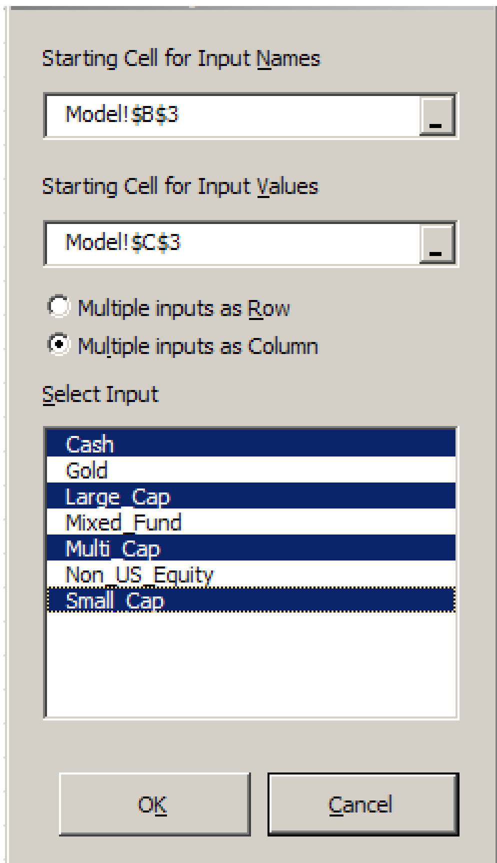
1. Opens the library file.
2. Creates a "PMTable" sheet.

3 Names cell A1 of the PMTable sheet PM_Index.

Cell A1 will be the Column Input Cell of the Data Table that will run the simulation.

The next thing you need to do is set up the input variables and link them to both the data table and the distributions from the library.

Inputs are asset return probability distribution SIPs



For this you'll use the  **Define Inputs** tool. We need to specify where to put the input cells for the distributions you'll be using – in this case you'll put the names in a column starting at B3 and the values in a column starting at C3.

On the Model sheet, select cell C3 and click **Define Inputs** on the **SIPmath** ribbon. In the dialog that comes up, with the cursor in the "Starting Cell for Input Names", select cell B3. Make sure the "Multiple inputs as Column" is ticked. Finally, select some assets to use in the portfolios. Use the Shift and Ctrl keys in the usual ways to select multiple entries from the list.

Here we've chosen Cash, Large_Cap, Multi_Cap, and Small_Cap. You don't need to know what the names mean to enjoy the exercise.

Verdana 10

PM_Index

	A	B
1	1	
2		Index
3		Values
4		1
5		2
6		3
7		4
8		5
9		6
10		7

=INDEX('Assets Library.xlsx'!Cash, PM_Index)

B	C	D	E	F
Cash	0.019945			
Large_Cap	-0.14445			
Multi_Cap	-0.13153			
Small_Cap	-0.13284			

When you click OK, *Define Inputs* will put the names of the assets you chose into a column below B3 and the index formulas in a column below C3. The numbers in column C are the first trial of each asset distribution.

4.4 THE MODEL

The model will consist of a few portfolios and their calculated returns.

Assume the total investment is 1.0. Allocate to assets.

=1-SUM(D4:D6)

B	C	D
		Port1
Cash	0.019945	0.4
Large_Cap	-0.14445	0.2
Multi_Cap	-0.13153	0.2
Small_Cap	-0.13284	0.2

Define a portfolio in column D with its name in D2 and a weight for each asset. The weights must add up to 1.0 because you want them to be the fraction of total investment allocated to each asset. To enforce this, use a formula in D3 to set the Cash weight so that the portfolio sum is 1.0.

Then, to calculate the return from this portfolio, multiply each asset's return by its weight and add up all the results. Excel's *SUMPRODUCT()* function does that in one operation.

=SUMPRODUCT($C$3:$C$6,D3:D6)

B	C	D
		Port1
Cash	0.019945	0.4
Large_Cap	-0.14445	0.2
Multi_Cap	-0.13153	0.2
Small_Cap	-0.13284	0.2
		-0.07379

That gives you the model for one portfolio. Note that $'s appear on C3:C6 in the formula but not on D3:D6). This will ensure that when we create multiple portfolios, they will all be pointing at the same inputs. We will make four more portfolios now. Select D2 through to D7 and drag/copy them over to column H.

Excel will automatically increment the name "Port1" as it copies and it will adjust the formula for each column. Now, change the portfolio weights so that each portfolio is different. You should have something like this:

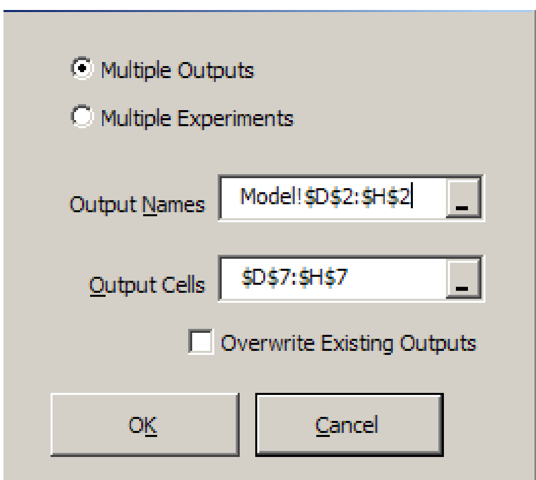
	A	B	C	D	E	F	G	H
1								
2				Port1	Port2	Port3	Port4	Port5
3		Cash	0.01995	1	0.6	0.3	0	0
4		Large_Cap	-0.1444	0	0.2	0.2	0.4	0
5		Multi_Cap	-0.1315	0	0.2	0.3	0.4	0
6		Small_Cap	-0.1328	0	0	0.2	0.2	1
7				0.02	-0.043	-0.089	-0.137	-0.133

Interesting portfolio choices

That's the model.

The formulas in column C hook the model to the input distributions, and now you need to hook it up to the data table and start the simulation.

4.5 THE OUTPUT DISTRIBUTIONS



For this you'll use the **Define Outputs** tool on the **SIPmath** ribbon. Select the cells with the model results, D7:H7 and click **Define Outputs**. In the dialog box that comes up, set the Output Names field to D2:H2 by selecting the cells containing "Port1" to "Port5" while the cursor is in the Names field.

When you click **OK**, *Define Outputs* will create a named range in the data table for each portfolio, using the output names you selected. Each of these will be an output column holding the results of running the model with all the different values in the input distributions. If you switch to the PMTable sheet, you should see something like this:

Data Table does all the work

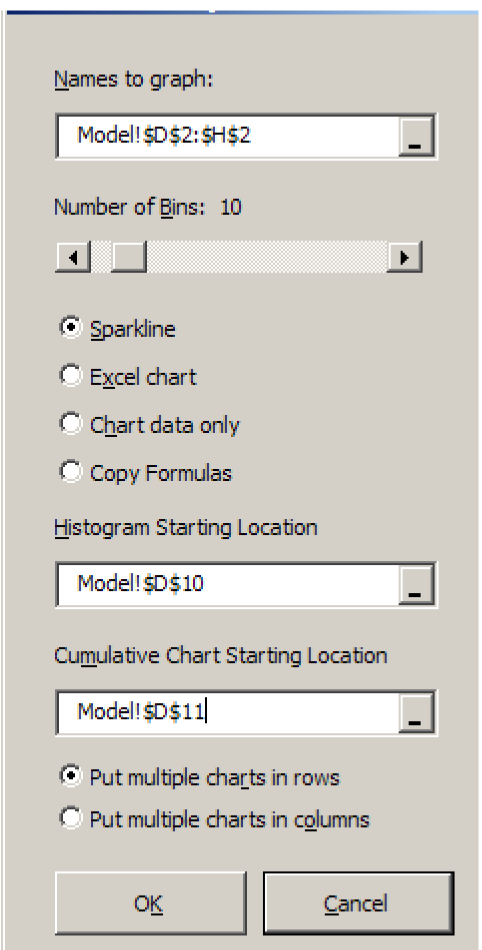
Port5 f_x {=TABLE(,A1)}

	A	B	C	D	E	F	G
1	1						
2		Index	Port1	Port2	Port3	Port4	Port5
3		Values	-0.07379	-0.05735	-0.08893	-0.10434	-0.09268
4		1	-0.07379	-0.05735	-0.08893	-0.10434	-0.09268
5		2	0.038278	0.037581	0.043186	0.059786	0.02471
6		3	0.349472	0.327784	0.415995	0.50835	0.268577
7		4	-0.1275	-0.10667	-0.14948	-0.18272	-0.14043
8		5	-0.05074	-0.03902	-0.05988	-0.07538	-0.06446
9		6	0.015329	0.021423	0.016696	0.019105	-0.00621
10		7	-0.01371	-0.00778	-0.01801	-0.01673	-0.02568
11		8	0.142096	0.107794	0.156957	0.177487	0.212447
12		9	0.125267	0.114789	0.135246	0.195759	0.111477
13		10	0.237369	0.205247	0.281652	0.315139	0.250571
14		11	0.126233	0.097299	0.147836	0.138211	0.18544
15		12	-0.02955	-0.02793	-0.03036	-0.06386	-0.01647

5 uncertain variables, 5000 numbers

The blue columns are the result distributions. Now you have to turn all these numbers into something that's easily grasped and can inform a decision.

4.6 PRESENT THE RESULTS

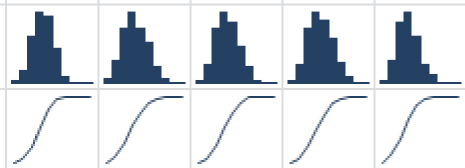


Note, when you refer to an output distribution, you must refer to its name, not the output cell in the model that generated the results in the data table. Now you want to present the results of the simulation. Switch back to the Model sheet to do this. Start with plotting sparklines for each distribution.

For that you'll use the **Stats & Graphs** tool on the **SIPmath** ribbon. Select the names of the distributions to plot – "Port1" etc. in D2:H2, and click on the **Stats & Graphs** tool. In the dialog, select the *Sparkline* option. This will extend the dialog box so that you can say where you want the sparklines. Set the Histogram field to D10 and the Cumulative field to D11 and click **OK**.

SIPmath will plot the histograms and cumulative probability charts under the corresponding portfolio columns. They'll be a bit squashed, so expand heights row 10 and row 11 to about the size of three rows so they look right. With a little decorating, you should have something like this:

	A	B	C	D	E	F	G	H
1								
2				Port1	Port2	Port3	Port4	Port5
3		Cash	0.01995	1	0.6	0.3	0	0
4		Large_Cap	-0.1444	0	0.2	0.2	0.4	0
5		Multi_Cap	-0.1315	0	0.2	0.3	0.4	0
6		Small_Cap	-0.1328	0	0	0.2	0.2	1
7				0.02	-0.043	-0.089	-0.137	-0.133
8								
9								
10								
11								



Now if you make changes to the portfolio weights, you'll see the effects on the sparklines.

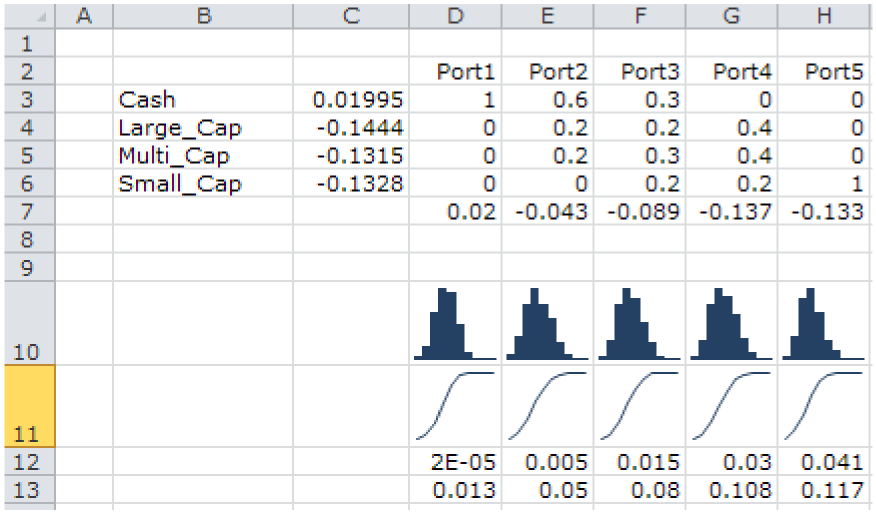
The objective is to compare portfolios, so you'll need some more numbers – the average and variance of each return distribution. As with the model, you'll set up the formulas for one portfolio and then extend it across to all of them.

In cell D12, put =*VAR(Port1)*

In cell D13, put =*AVERAGE(Port1)*

That sets up Port1.

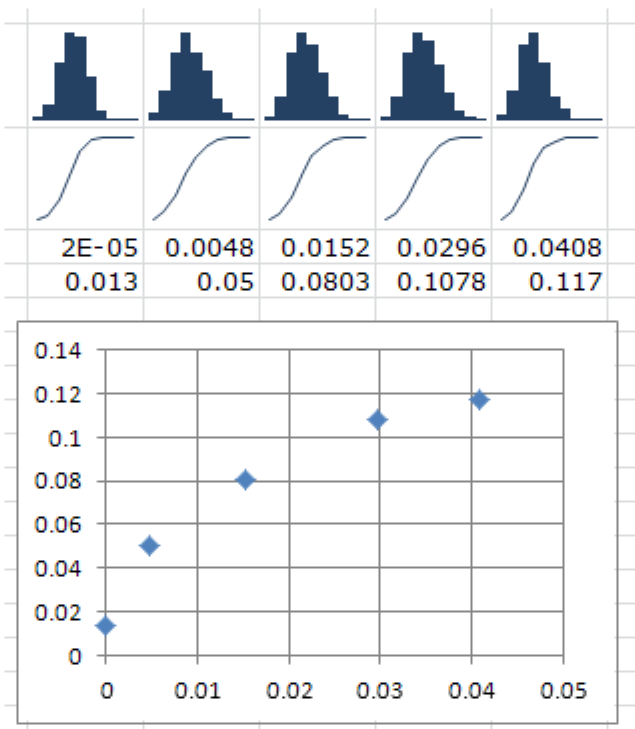
Select D2:H2 (the names of the portfolios) and click on Stats & Graphs. This time select *Copy Formulas.* In *Formulas to Copy* select D12:D13. When you click OK, Stats & Graphs will copy the formulas across to column H, changing "Port1" to the correct name reference for each column. You should now have something like this:



	A	B	C	D	E	F	G	H
1								
2				Port1	Port2	Port3	Port4	Port5
3		Cash	0.01995	1	0.6	0.3	0	0
4		Large_Cap	-0.1444	0	0.2	0.2	0.4	0
5		Multi_Cap	-0.1315	0	0.2	0.3	0.4	0
6		Small_Cap	-0.1328	0	0	0.2	0.2	1
7				0.02	-0.043	-0.089	-0.137	-0.133
8								
9								
10								
11								
12				2E-05	0.005	0.015	0.03	0.041
13				0.013	0.05	0.08	0.108	0.117

Finally, use the averages and variances in a scatterplot so that they can be compared more easily.

Select D12:H13, the variances and averages. On the Excel ribbon, click *Insert* | *Charts* | *Scatter* and choose the style with no lines. With a little bit of decorating, you should end up with something like this:



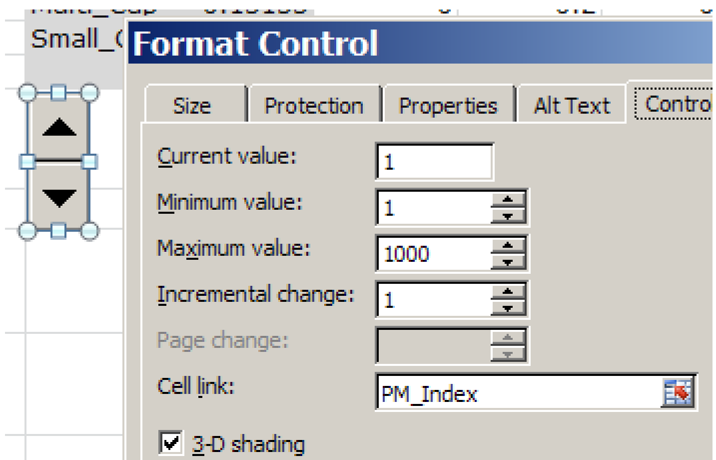
The scatterplot interpretation is that the vertical axis is return and the horizontal axis is risk. Up and to the left is better: higher return, lower risk. As you can see as you change portfolios, 'better' has a lot to do with how big a return you want and how much risk you can tolerate.

4.7 TRIAL BROWSING SPINNER

To give your users the ability to browse through the trials and see the different possible outcomes, you'll need to add a spinner control that steps PM_Index through the trials.

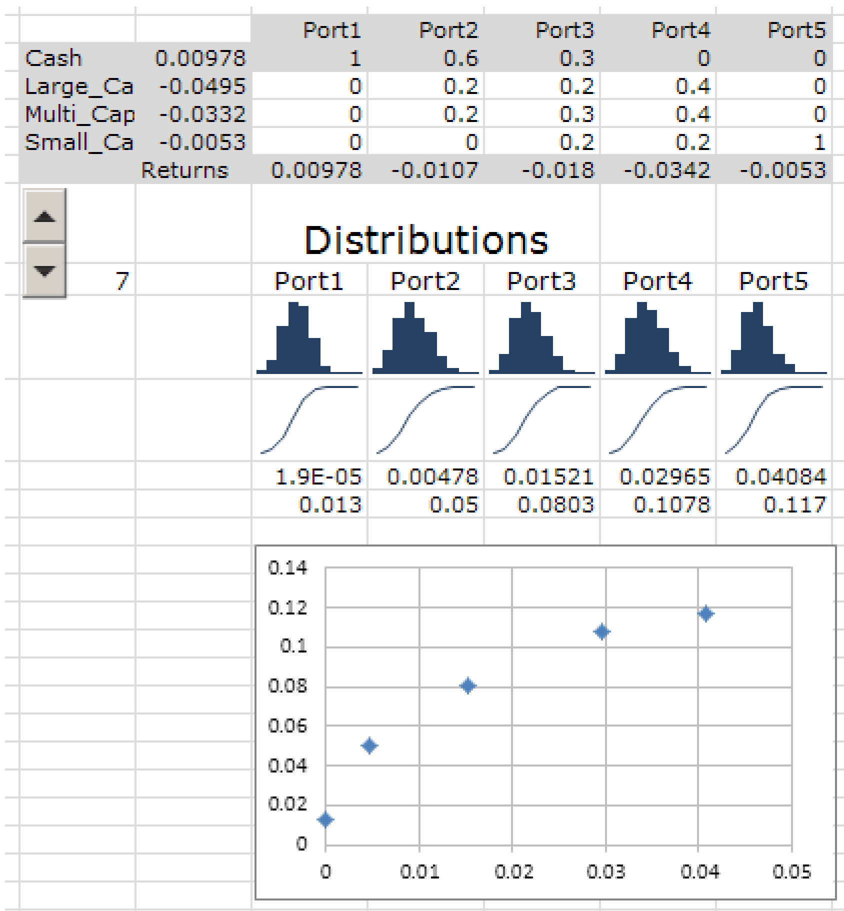
Go to the *Developer* tab, select *Insert* and click on the Form Controls Spin Button. Drag the cursor to draw the control in the open area next to B8 and B9. When you let go, you'll be in design mode. Right-click on the control and select *Format Control*.

On the *Control* tab, set *Current Value* and *Minimum value* to 1, *Maximum value* to 1000, and *Cell link* to *PM_Index*. You should be seeing something like this:



Click *OK* then click in any cell to turn off *Design Mode*.

The last thing is to put the trial number where it can be seen. In B9, put the formula =*PM_Index*. Click the button a few times to make sure things are working. Here's the whole picture:



		Port1	Port2	Port3	Port4	Port5
Cash	0.00978	1	0.6	0.3	0	0
Large_Ca	-0.0495	0	0.2	0.2	0.4	0
Multi_Cap	-0.0332	0	0.2	0.3	0.4	0
Small_Ca	-0.0053	0	0	0.2	0.2	1
	Returns	0.00978	-0.0107	-0.018	-0.0342	-0.0053

		Distributions				
7		Port1	Port2	Port3	Port4	Port5
		1.9E-05	0.00478	0.01521	0.02965	0.04084
		0.013	0.05	0.0803	0.1078	0.117

There's a finished version of this in file *SIPmathExercise.xlsx* that's included in the **SIPmath** distribution. Make sure to open the *Assets Library* workbook first.

4.8 RUNNING WITHOUT SIPMATH

One of the important design parameters for the **SIPmath** Modeling Tools is that the application must work without the **SIPmath** add-in installed. If you used a library in the current workbook, all you need to do is open your application workbook.

If you used an external library you **must** open the library workbook first and then your application workbook.

---

[i] Probability Management, Sam Savage, Stefan Scholtes and Daniel Zweidler, OR/MS Today, February 2006, Volume 33 Number 1
[ii] Sam L. Savage, The Flaw of Averages, Why we Underestimate Risk in the Face of Uncertainty, John Wiley 2009, 2012