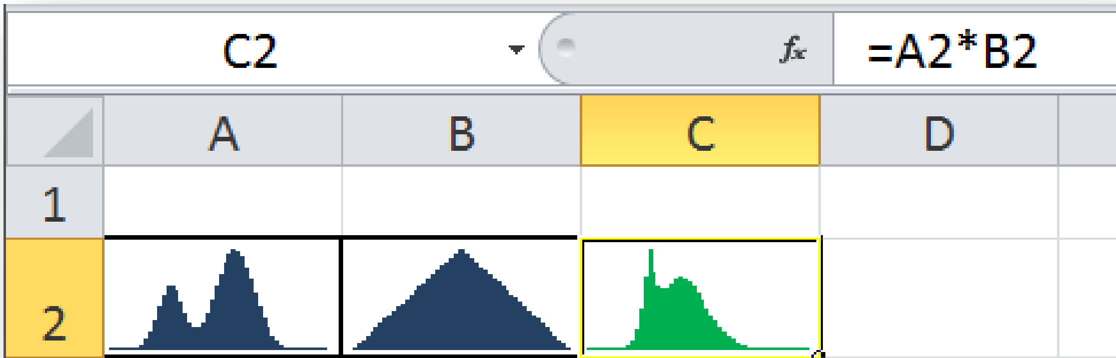
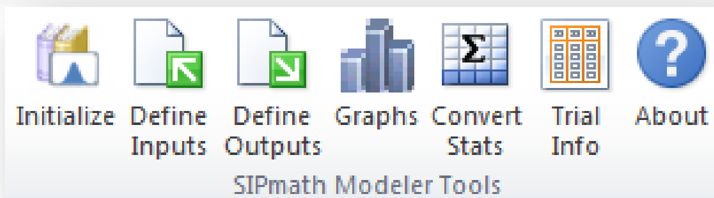




SIPmath™ Modeler Tools for Excel

REFERENCE MANUAL

Sam Savage (Sam@ProbabilityManagement.org)

Marc Thibault (Marc@ProbabilityManagement.org)

Rev. April 9 2015

SIPmath Modeler Tools Version 2.4

Current Build: 2.4.0

CONTENTS

1 INTRODUCTION

The SIPmath Modeler Tools for Excel (the **Tools**) facilitate the development of stochastic spreadsheet applications that make use of the native Data Table function[i]. The models may be based on SIP Libraries to support the discipline of probability management, or dynamic (random mode) simulations based on the Rand() formula. Once created, such applications do not require the **Tools** add-in to run. The **Tools** work with major versions of Microsoft Excel software on Windows and Mac OS. Although any of the functions of the **Tools** may be performed manually in Excel, they greatly facilitate the model development process.

Probability Management is a decision support framework for dealing with uncertainties much in the same way that traditional data management deals with regular spreadsheet numbers[ii],[iii].It is based on four principles[1]:

1 **Communicating Distributions as Data**

Normally, spreadsheets allow each cell to store only a single value. Thus, when estimating uncertainties, spreadsheets show just one number, typically a single "average" future scenario. This leads to a set of systematic errors known collectively as the Flaw of Averages[iv] which explains why so many projects are behind schedule, beyond budget, and below projection.

The SIPmath Modeler Tools allow Excel to simultaneously model thousands of potential futures, allowing us to make better decisions in the face of uncertain future prices, demand, project duration, etc.

These potential futures are stored in "Stochastic Information Packets" or SIPs, which contain hundreds or thousands of formula-driven (or historical) outcomes of the uncertainty in question. Conceptually, SIPs

---

[1]See more at www.ProbabilityManagement.org

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

are just arrays filled with realizations and optional metadata concerning the uncertainty in question.A central goal of ProbabilityManagement.org is to promote cross platform standards for formatting libraries of SIPs (SIP Libraries).

A SIP contains realizations of a distribution

2 Interactive Simulation

Microsoft Excel has become so powerful that today, worksheets may be vectorized, that is formulas are applied to whole arrays instead of individual numbers. Expressions based on the Rand() formula or pre-compiled distributions (stored as SIPs) may be used in interactive simulations in an approach we call SIPmath. It can yield output distributions of tens of thousands of trials before your finger leaves the <Enter> key, that can also be stored as SIPs.

SIPmath works in Microsoft Excel using no other computational functions but the standard `INDEX()` formula and the Data Table function. Therefore, models created with SIPMath Modeler Tools can be shared with people who do not have the Tools themselves.

3 Credibility of Uncertain Estimates

Many organizations do not grant the permission to be uncertain. As a managerial cure to this problem the discipline of probability management suggests a Chief Probability Officer (CPO) who takes responsibility for the accuracy and provenance of the SIP Libraries used by an organization. This provides access to centralized, cross platform, “standardized” representations of uncertainty that yield repeatable and identical results for all authorized model users.

4 Coherence of relevant distributions

When there are multiple uncertainties relevant to each other, it is important to make sure they stay coherent. For example, consider two SIPs of 1,000 trials, one representing the value of your house, and the other representing the payout 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 value SIP is zero (total loss), while the other 999 take on its actual value. However, a SIP representing the fire insurance payout is clearly relevant to the house value SIP: it must have one positive element in which it pays up to the coverage terms, while the other 999 entries contain the negative value (your insurance premium). These two

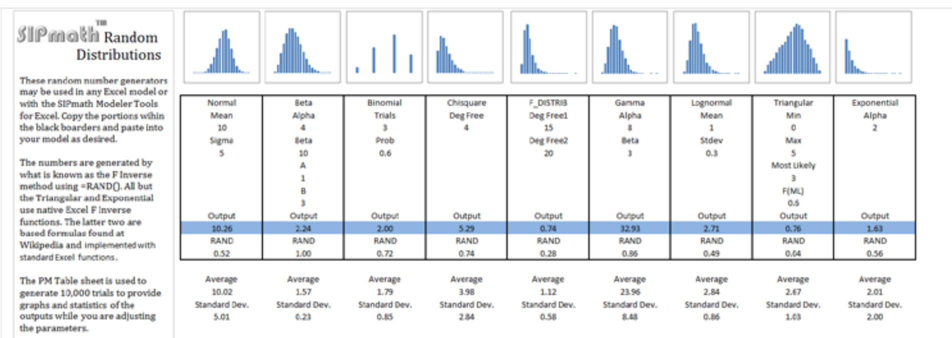
SIPs are said to be coherent if the realization of the value loss (fire) on the value SIP lines up with the realization of payment in the insurance payout SIP. A coherent set of SIPs is also known as a Stochastic Library Unit with Relationships Preserved or SLURP.

If the house and fire insurance SIPs were not coherent, it would be like buying insurance on someone else's house. There would be an event when the house has burned down (zero value) but insurance has not paid off, or the insurance pays off but the house doesn't burn down.

2 SIPMATH OPERATING MODES

The SIPMath Modeler Tools consist of an Excel add-in that can be configured to run in one of the two major modes:

Random Mode: Random mode is much like traditional Monte Carlo simulation except that the computations take place interactively with up to tens of thousands of trials taking place before your finger leaves the <Enter> key. The **Tools** make use of the built-in Excel random number generator function (`RAND ()`) to drive model inputs. Today Excel offers numerous distributions built off this function, as provided in Random Generators.xlsx.



In addition to the distributions contained in the file above, historical data may be resampled with a formula of the type `=INDEX(Data,RANDBETWEEN(1,N))`,where N is the number of data elements.

Random Mode is quick way to get a "sense" of a stochastic process and to generate SIPs for use in SIP Library mode. Models in this mode are not auditable because it is difficult to control the order of the random trials.

SIP Library Mode: In this mode, the inputs are drawn from a SIP Library, which may be either stored within the model, or linked to as an external file. This is the mode that we refer to as SIPmath, and it is required for the discipline of probability management. From the perspective of the Excel user it adds a third dimension to your worksheet representing thousands of potential outcomes. You may scroll through this

See Distribution Processing and the Arithmetic of Uncertainty at www.analytics-magazine.org third dimension either with the View Trials tool or by changing the value of PM_Index on the PMTable sheet.

This is fundamentally different from earlier forms of Monte Carlo simulation in that SIPmath draws its values from the pre-compiled libraries of Stochastic Information Packets (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 and thus are suitable for powering the robust business decisions.

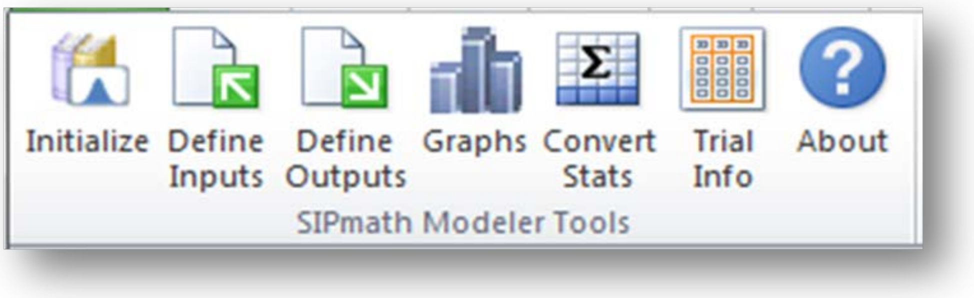
These tools facilitate the creation of applications, but are not required to run them.

As a result, if various SIPmath applications are based on the same set of coherent input SIPs the results may be networked together. Thus, the results of complex stochastic models may be aggregated from sub models.

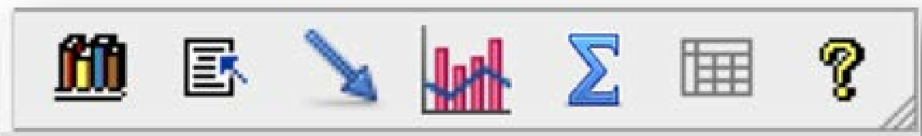
3 SIPMATHMODELER TOOLOVERVIEW

To start working with the SIPMath package, open the `SIPMath.xlam`. A ribbon with six tool buttons will appear.

Windows



Mac



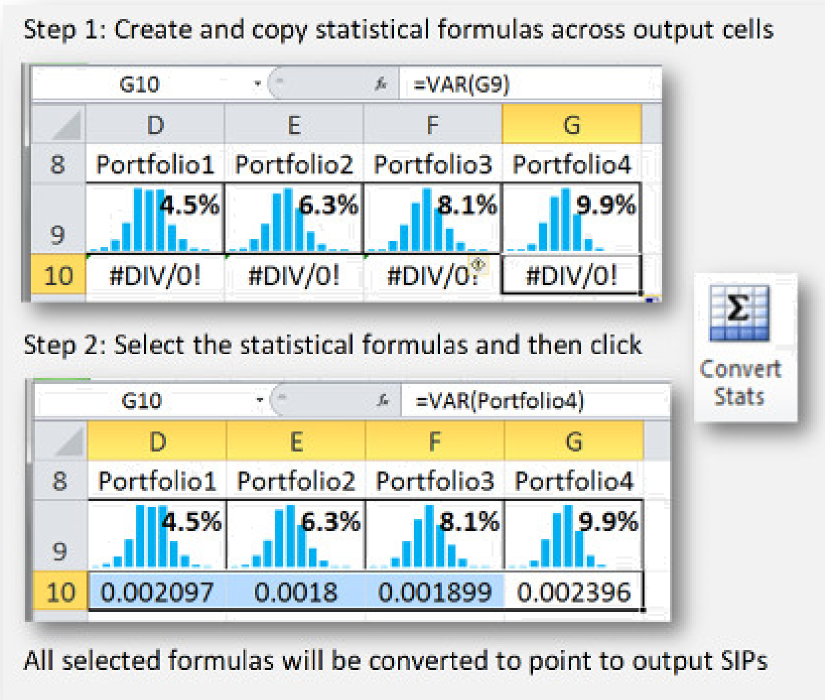
The leftmost *Initialize* tool identifies the Stochastic Library containing the Input SIPs in SIP Library mode or specifies the number of trials to run in the Random Mode. This button will also reinitialize existing models or allow the number of trials to be changed in Random Mode.

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

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

The *Graphs*tool quickly crates histograms cumulative graphs for the output cells once a simulation has been created. It may also create such graphs from pre-existing arrays of data in Excel.

The *Convert Stats*tool is used to apply statistical functions such as `=AVERAGE` or `=VAR` to the outputs of a simulation. The function is applied to the output cell itself. When the Convert Stats button is clicked, the argument of the formula is replaced with the output SIP as shown below.



The *Trial Info*tool works only in SIP Library modes to step through the input SIP values one trial at a time, or displays optional metadata such as the average or specified percentile as stored in the SIP.

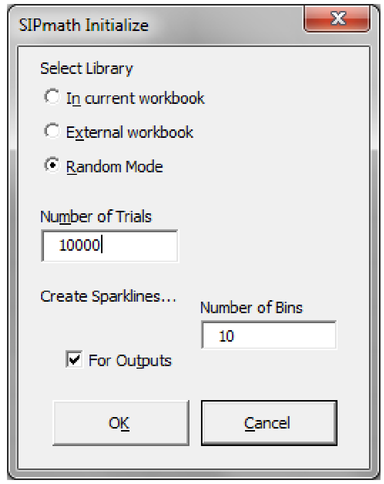
The ribbon can be permanently integrated into Excel as a plug-in via the *Tools / Add-ins* menu.

Note: It is a good idea to save your work before initializing SIPMath.

Due to the differences in versions and platforms of Excel, some action sequences may make Excel unstable; if this happens, please make sure the problem is repeatable and submit a bug report with a list software versions and user actions to Support@ProbabilityManagement.org.

4 EXERCISE: MODEL IN RANDOM MODE

SIPMath operates in two main modes: “Random” and “SIP Library”. The main difference is that the former uses internal Excel random function generators to produce distributions; hence the simulation may yield slightly different results every time it is run.

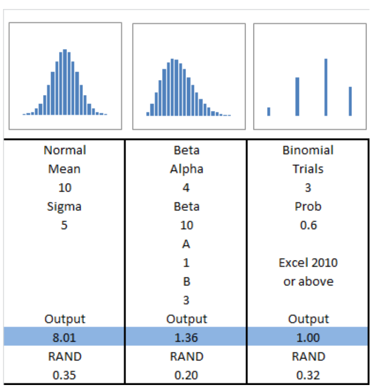


Let’s say we are interested in finding the distribution of a minimum of two normal random variables. Such a distribution is hard to deduce analytically, but easy to simulate with the SIPmath Modeler Tools in just a few steps:

Step 1. Initialize in Random Mode

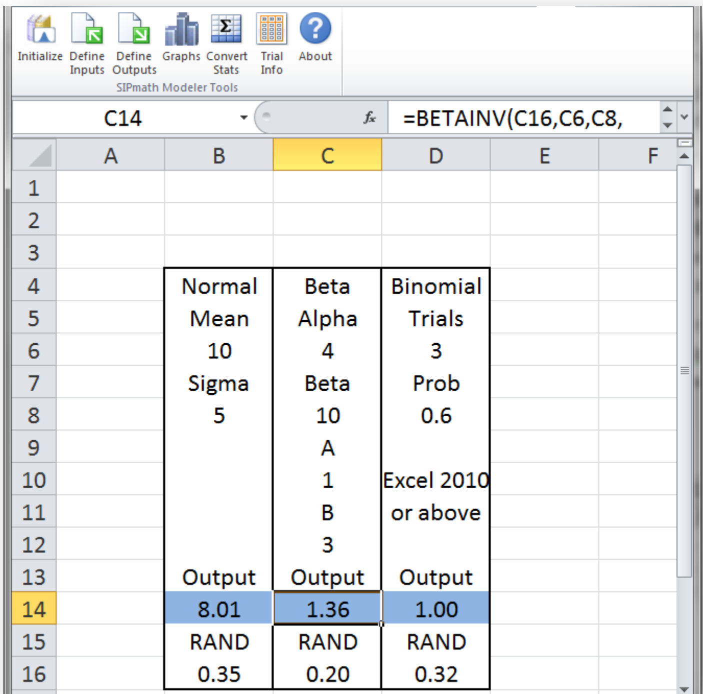
Press the *Initialize* button, choose “Random Mode”, type the desired number of trials and hit OK: This will set up the PMTable and SIPmath Chart Data sheets in your workbook

Step 2.Create random inputs

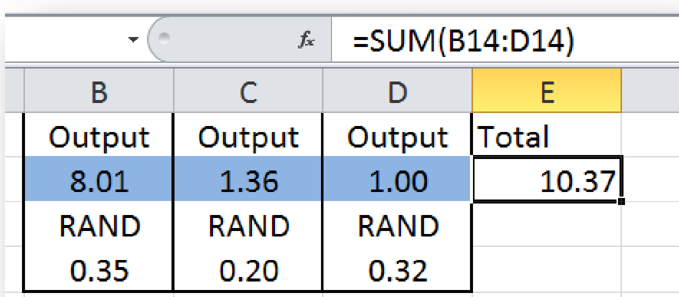


The simplest approach is to cut and paste the formulas from the Random Generators.xlsx workbook, a portion of which is shown on the left.

Copy the rectangular regions for the desired distributions and paste them into your worksheet (do not copy the graphs).**NOTE** Close the Random Generators workbook before running your model, as a second sheet with RAND() formulas in it will slow down the simulation.

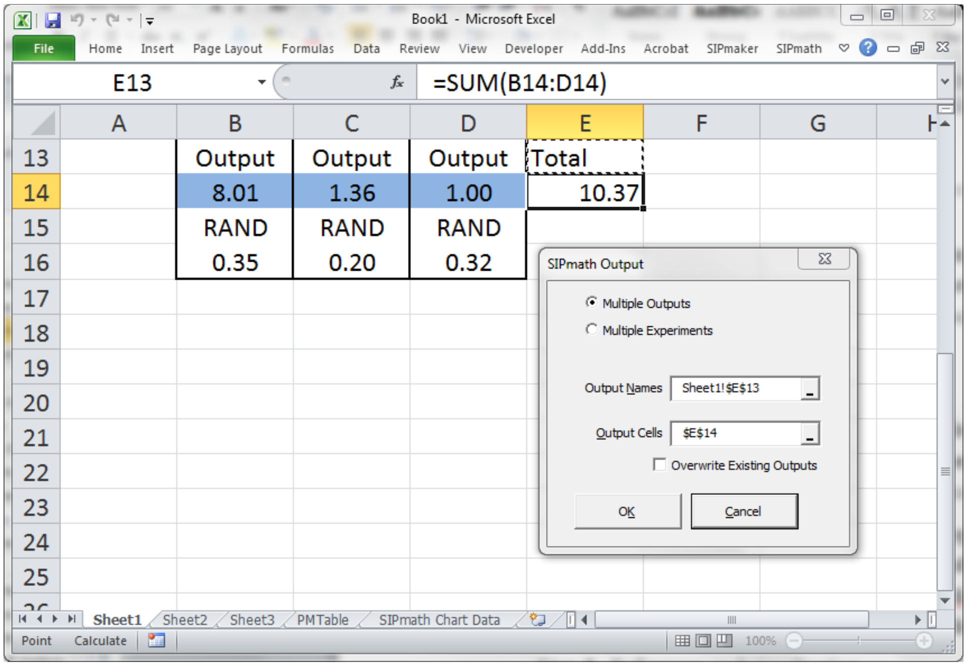


Step 3. You may adjust the parameters of the input distributions and create the cell you wish to output. In this case we will sum the three random variables and give it the name “Total”.

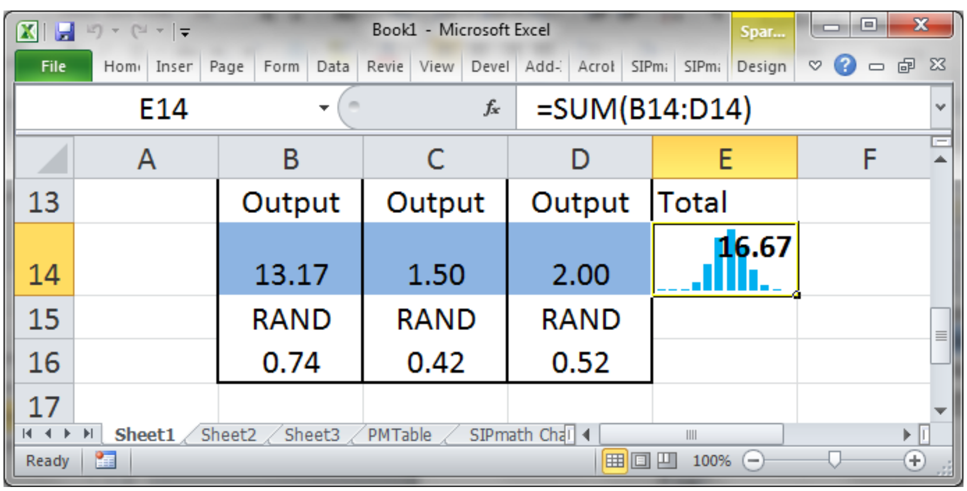


Step 4. With the cursor in output cell (E14) press *Define Outputs* button on the ribbon and specify the name (cell E13). If there are multiple output

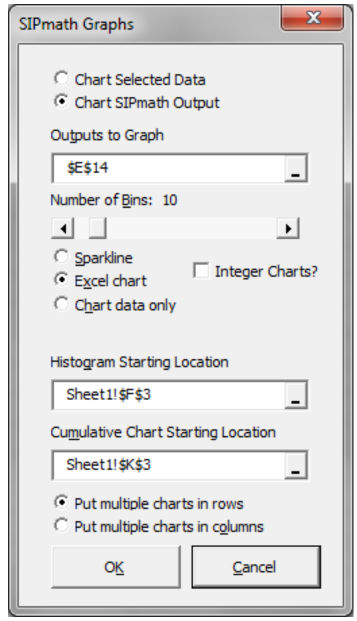
cells (several random variables to study), they can be entered together in the normal Excel fashion (as continuous ranges or lists). Names should also follow Excel conventions and include only letters, underscores and numbers and not overlap with built-in identifiers.



After hitting OK, SIPMath will convert the text found in cell E13 into a range name and use cell E14to feed multiple consecutive trials into a table on the PMTable sheet. It also automatically places a Sparkline graph in cell E14 if this option was specified during initialization.



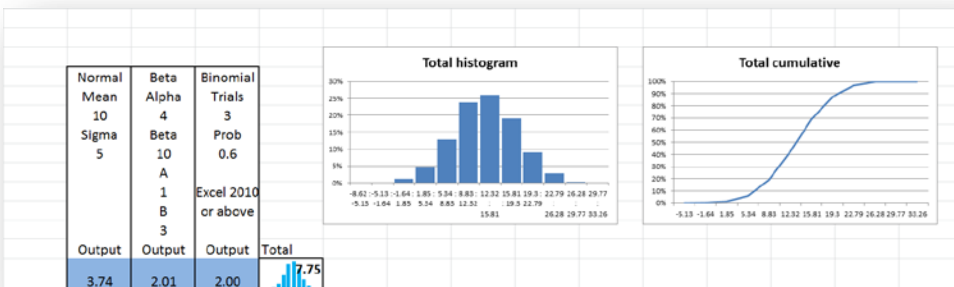
At this point, the simulation is complete and the spreadsheet now has 10,000 samples of the sum of the three random variables stored as a range named “Total” on the PMTable sheet. Now try changing the parameters of the distribution or the formula in cell E14 to instantly run a new simulation.



Step 5. Graphs

To create histograms or cumulative graphs as either sparklines or regular Excel graphs, select the simulation output cell (E14) and click the Graphs button.

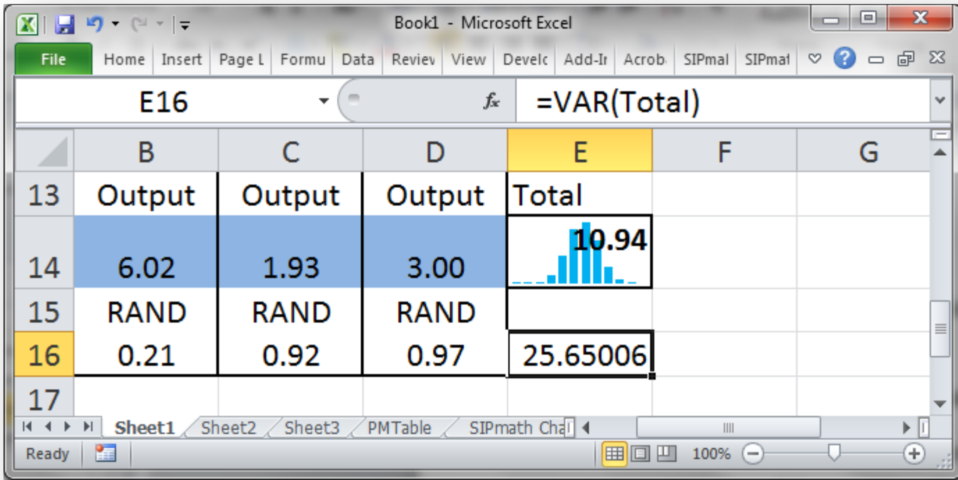
In this example we will specify both an Excel histogram and cumulative graph in cells F3 and K3 Respectively.



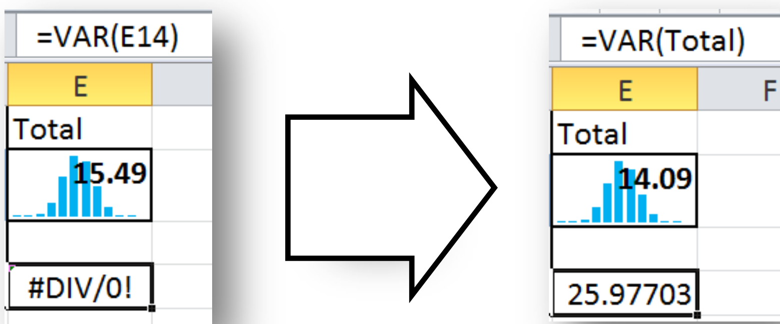
The number of decimal places displayed in the graph as well as the offset of the horizontal axis may be specified in the SIPmath Chart Data sheet.

Step 6. Statistics

We can use the built-in Excel functions like `Var()` to estimate various statistics of the resulting distribution by referring to its range name:



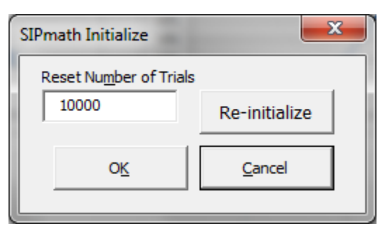
Alternately you may apply the statistical formula to the output cell itself and then press the Convert Stats button.



This is particularly useful when copying statistical formulas across a range of output cells.

Step7. Re-initializing

If you want to run a larger number of trials, press the *Initialize* button again. Re-initializing also provides the option to change to SIP Library mode and connect to an external or internal library.



5 EXERCISE: SIP LIBRARY MODEL

5.1 THE OBJECTIVE

We will now demonstrate how to build a SIP Library application using Excel and the *SIPmath Modeler Tools*. The application can be run without the use of VBA 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.

A financial portfolio assessment application

We are going to build a financial portfolio simulation so that we can compare the expected risks and returns for a variety of investment portfolio choices. We will present these as a graph plotting the mean against the variance of the return for each portfolio.

In financial circles, variance is a proxy for risk, so better portfolios have higher return and lower risk and variance. Risk and return are trade-offs, so the best portfolio depends on your risk attitude. That’s why you need a picture.

We will simulate five different portfolios and be able to compare their risk/return characteristics while making changes to their mix of assets.

Typically we would 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 we will 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.

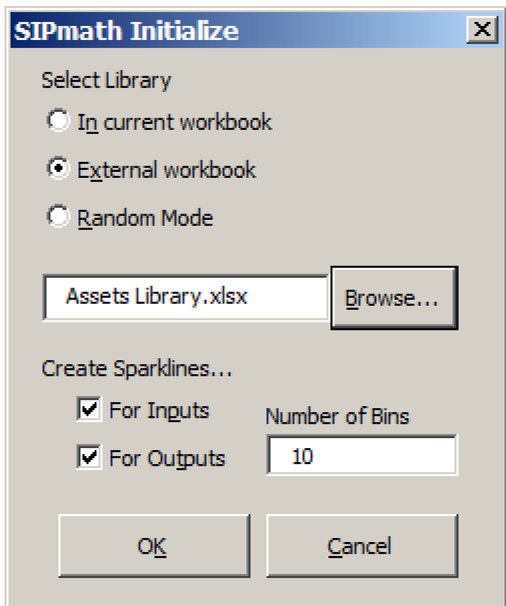
5.2 INPUT DISTRIBUTIONS

The first thing we need is a source of data about assets we 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, we will use the Assets Library.xlsx workbook that comes with the SIPmath Modeler Tools.

To get a stochastic library, use the *Initialize* tool on the ribbon. You are then 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 **Tools**.

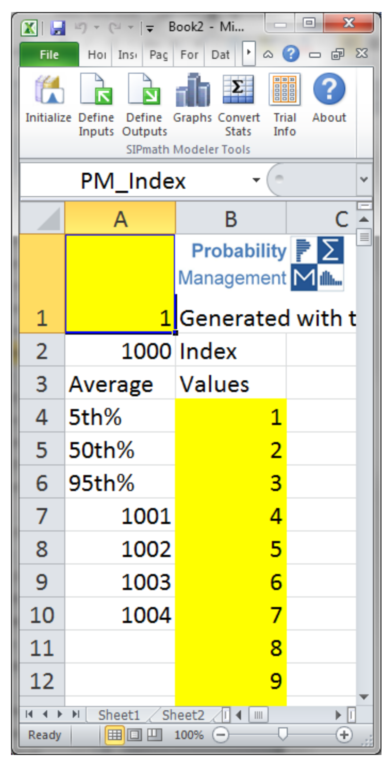
When you click OK, the *Initialize* function will

1 Open the library file.

2 Create a "PMTable" sheet.

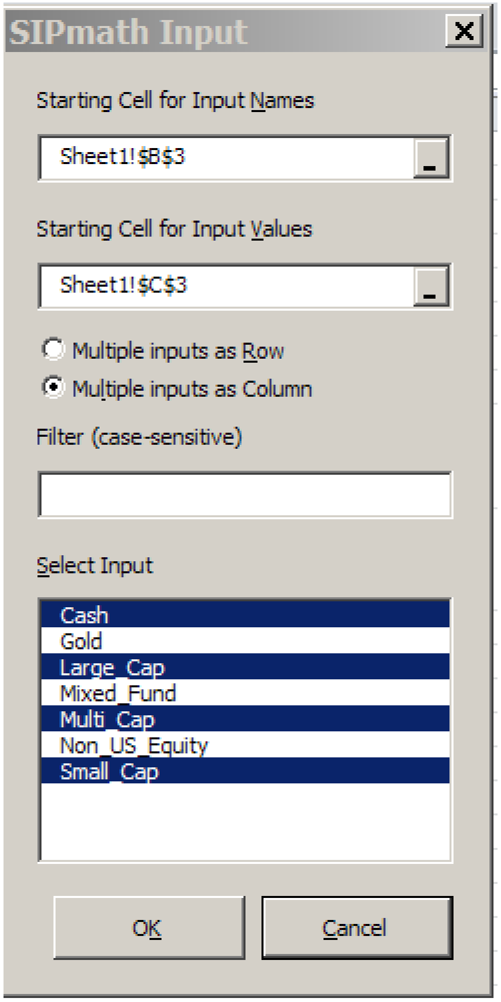
3 Name cell A1 of the PMTable sheet `PM_Index`, and add links to Meta Data such as Average, 5^{th} Percentile, etc., which is stored in the SIP Library.

4 Create a SIPmath Chart Data sheet.



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

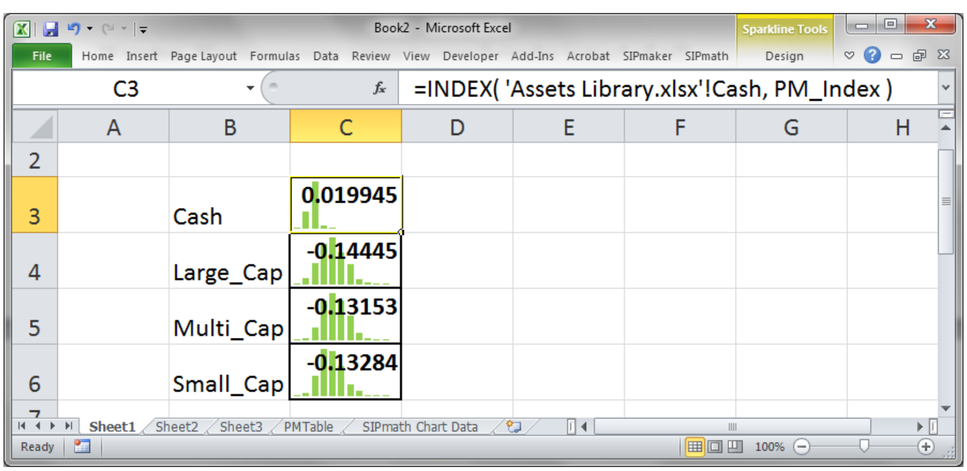
The next task is to set up the input variables, which get linked to the distributions from the library and PM_Index cell.



For this we will use the *Define Inputs* tool. We need to specify where to put the input cells for the distributions we'll be using – in this case we'll put the names (optionally) 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" field, select cell B3. Make sure the "Multiple inputs as Column" radio button 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 have chosen Cash, Large_Cap, Multi_Cap&Small_Cap. Because we opted to create Sparkline graphs for Inputs and Outputs, they are also created at this time.

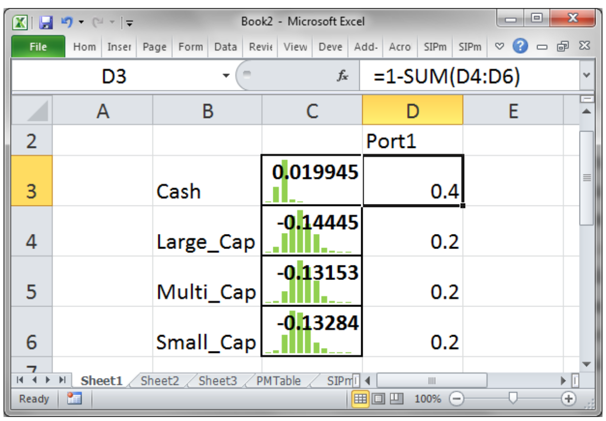


When you click OK, *Define Inputs* will put the names of the assets you chose into a column below B3 and the index formulas in a column below

C3. The numbers in column C are the first trial of each asset distribution. Because we opted to create Sparkline graphs for Inputs and Outputs, they are also created at this time.

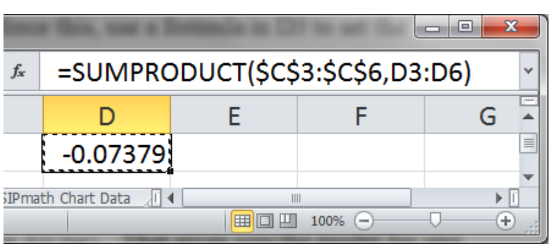
Warning If you use range names in your formulas instead of cell addresses, you'll need to have different names for the model input cells versus the corresponding input SIPs. This is a limitation of a name space in Excel.

5.3 THE MODEL



The model will consist of a few portfolios and their calculated returns. 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 achieves it in one operation.



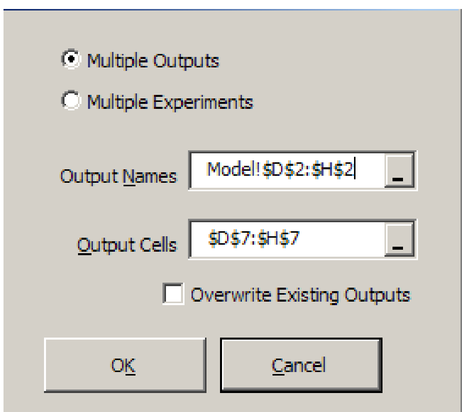
That gives us the model for one portfolio. Note that $'s appear on C3:C6 in the formula but not on D3:D6). This will ensure that when we create multiple portfolios, they will all be pointing at the same inputs. We will make four more portfolios now. Select D2 through to D7 and drag/copy them over to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

		Port1	Port2	Port3	Port4	Port5
Cash	0.019945	1	0.6	0.3	0	0
Large_Cap	-0.14445	0	0.2	0.2	0.4	0
Multi_Cap	-0.13153	0	0.2	0.3	0.4	0
Small_Cap	-0.13284	0	0	0.2	0.2	1
		0.019945	-0.04323	-0.08893	-0.13696	-0.13284

A variety of portfolio choices

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 to create the interactive simulation.

5.4 THE OUTPUT DISTRIBUTIONS

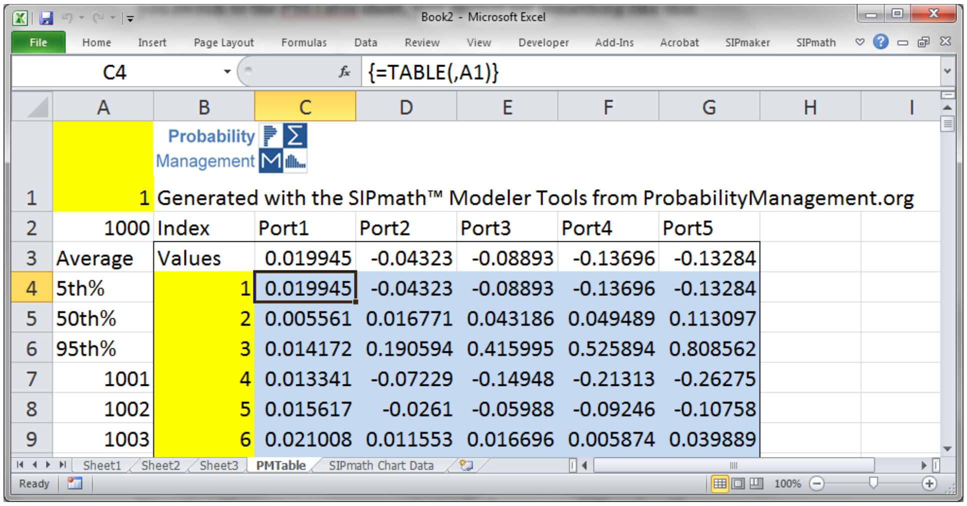


For this we'll use the *Define Outputs* tool on the SIPmath ribbon. Select the cells with the model results, D7:H7 and click *Define Outputs*. In the dialog box that comes up, set the Output Names field to D2:H2 by selecting the cells containing "Port1" to "Port5" while the cursor is in the Names field.

When you click OK, the simulation is set up and run, with Sparklines appearing in the output cells as opted for during initialization. At this point if you change any of the weights of the portfolios, the model will instantly run another 1,000 trials.

		Port1	Port2	Port3	Port4	Port5
Cash	0.019945	1	0.6	0.3	0	0
Large_Cap	-0.14445	0	0.2	0.2	0.4	0
Multi_Cap	-0.13153	0	0.2	0.3	0.4	0
Small_Cap	-0.13284	0	0	0.2	0.2	1
		0.019945	-0.04323	-0.08893	-0.13696	-0.13284

The *Define Outputs* command creates named ranges 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 a 1,000 trial simulation. If you switch to the PMTable sheet, you should see something like this:



	A	B	C	D	E	F	G
1	1	Generated with the SIPmath™ Modeler Tools from ProbabilityManagement.org					
2	1000	Index	Port1	Port2	Port3	Port4	Port5
3	Average	Values	0.019945	-0.04323	-0.08893	-0.13696	-0.13284
4	5th%	1	0.019945	-0.04323	-0.08893	-0.13696	-0.13284
5	50th%	2	0.005561	0.016771	0.043186	0.049489	0.113097
6	95th%	3	0.014172	0.190594	0.415995	0.525894	0.808562
7	1001	4	0.013341	-0.07229	-0.14948	-0.21313	-0.26275
8	1002	5	0.015617	-0.0261	-0.05988	-0.09246	-0.10758
9	1003	6	0.021008	0.011553	0.016696	0.005874	0.039889

The Data Table does all the work

5 uncertain variables, 5000 numbers

Presenting the results

As we showed in the Random Mode exercise, you can you can create histograms and cumulative graphs of the output cells with the Graphs button. In this exercise we will focus on calculating and comparing the variance and average of each portfolio.

We could proceed by entering the following formulas for portfolio 1.

In cell D9, put `=VAR(Port1)`

In cell D10, put `=AVERAGE(Port1)`

But that would just set up portfolio 1, and the results could not be copied to the other portfolios because the range names would not translate.

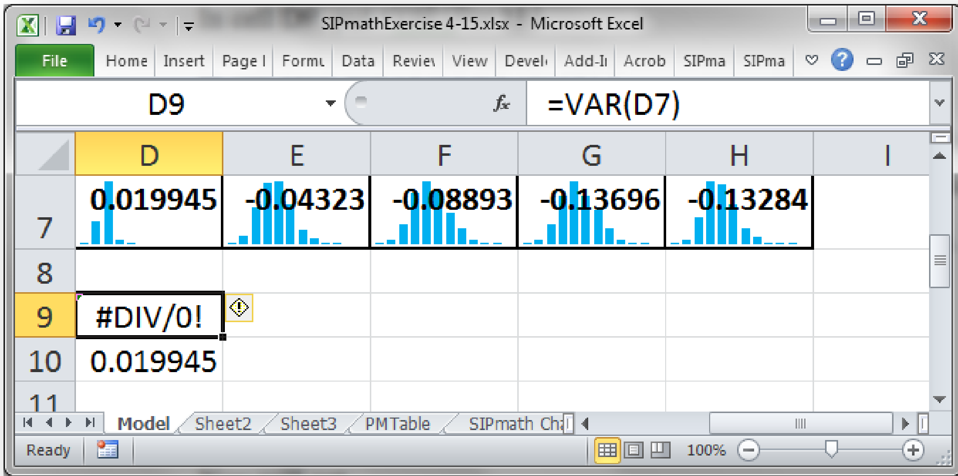
We solve this problem with the Convert Stats button as follows.

First

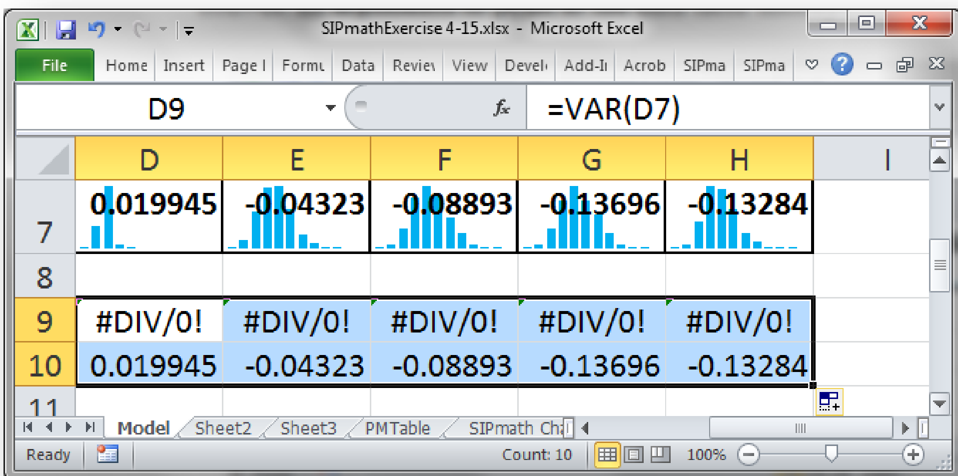
In cell D9, put `=VAR(D7)`

In cell D10, put `=AVERAGE(D7)`

You will get #DIV/0! In cell D9, but don't worry about it, as we will convert all arguments to point to the data table shortly.



Next, copy cells D9 and D10 to E9 through H10.



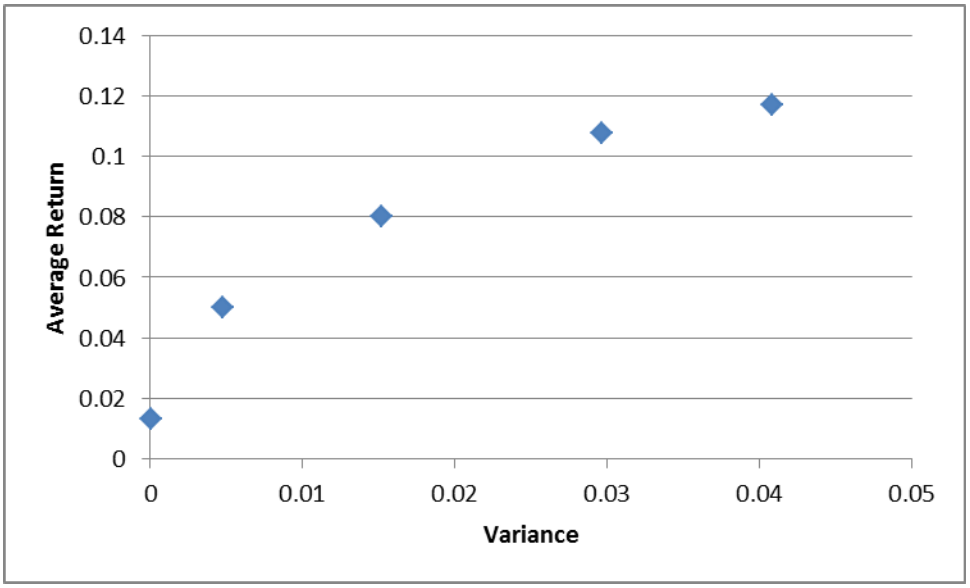
Finally with all ten of the formulas selected, click on the Convert Stats button, and they will be linked to the ranges on the PMTable sheet as shown.

NOTE: DO NOT USE =INDIRECT TO SOLVE THIS PROBLEM

In theory, the Indirect() formula could help in copying formulas, but it seriously degrades the speed of the Data Table.

Now we will use the averages and variances in a scatter plot so that they can be compared more easily.

Select D9:H10, 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 a similar result:



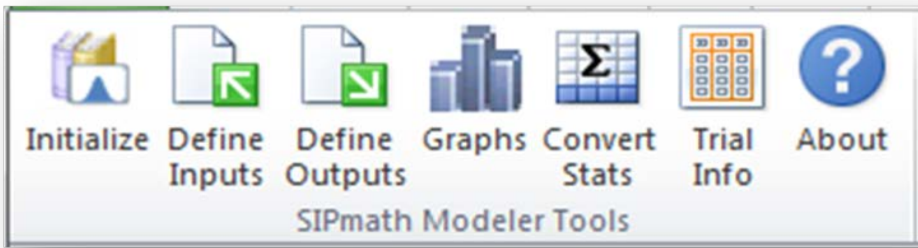
The scatterplot interpretation is that the vertical axis is return and the horizontal axis is risk. Up and to the left is better: higher return, lower risk. As you can see as you change portfolios, 'better' has a lot to do with how big a return you want and how much risk you can tolerate. Now as you change the portfolio weights, you will see the dots move around in the graph to reflect the risk and return.

5.5 RUNNING YOUR MODEL WITHOUT THE TOOLS

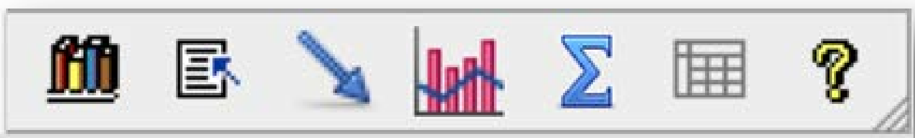
One of the important design parameters for the SIPmath Modeler Tools is that the application must work without the SIPmath add-in installed. If you used a library in the current workbook (or created a model in random mode), all you need to do is open your application. Thus, SIPMath-generated spreadsheets are easy to share.

6 THE SIPMATHRIBBON DETAILS

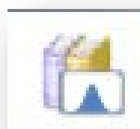
Windows

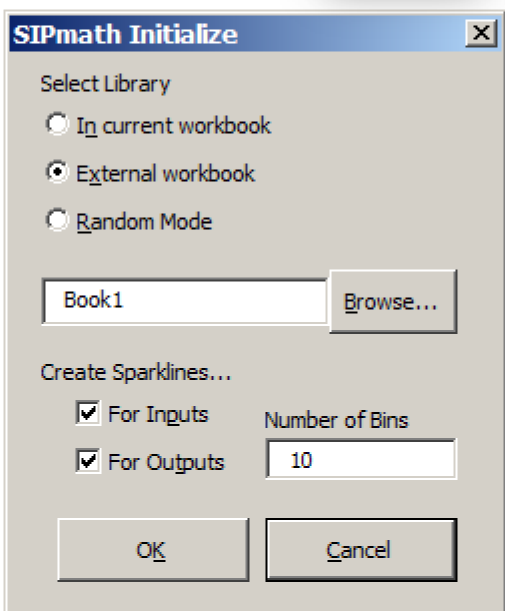


Mac



6.1 INITIALIZE





The *Initialize* tool must be used to create a model in either SIP Library or Random mode.

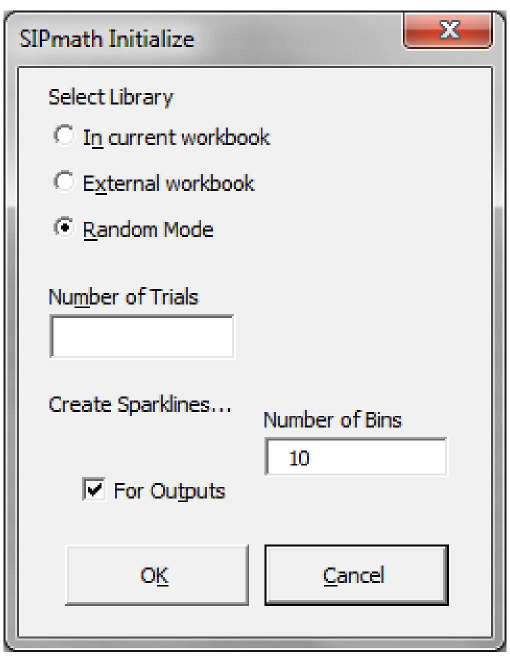
In Random mode the initialization queries the user as to the number of trials desired. The varaible `PM_Trials`, containing the number of trials, is then either read from the SIP Library, or in Random mode, created as a defined name.

In SIP Library mode input distributions are assumed to be named ranges in an Excel Workbook. The *Initialize* tool tells the SIPmath Modeler Tools where to find the distributions which will be used for input.

They can be in the model workbook or an external workbook. To specify an external workbook, click “Browse”and use the file dialog to select the library file required.

The model will be linked to the library file at this stage through the use of INDEX() formulas, the second argument of which is PM_Index, a cell on the PMTable sheet that it erates the simulation. 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.

The Initialize tool also allows you to specify whether the SIPmath Modeler Tools will automatically create sparkline histograms for the input and output distributions created in the model. If the corresponding boxes are checked, the input or output sparklines will be created in the cells containing the INDEX() formula (for inputs) or the formula that computes the output. Input sparklines default to light green; output sparklines default to light blue. The cells will be bordered in black. The “Number of Bins” field allows you to specify how many bins will be used to created the sparkline histograms.



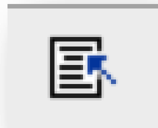
NOTE When Initializing a model in Random Mode, you will be asked only for the Number of Trials, the Number of Bins, and whether to create sparklines for outputs.

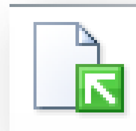
6.1.1 Reinitializing

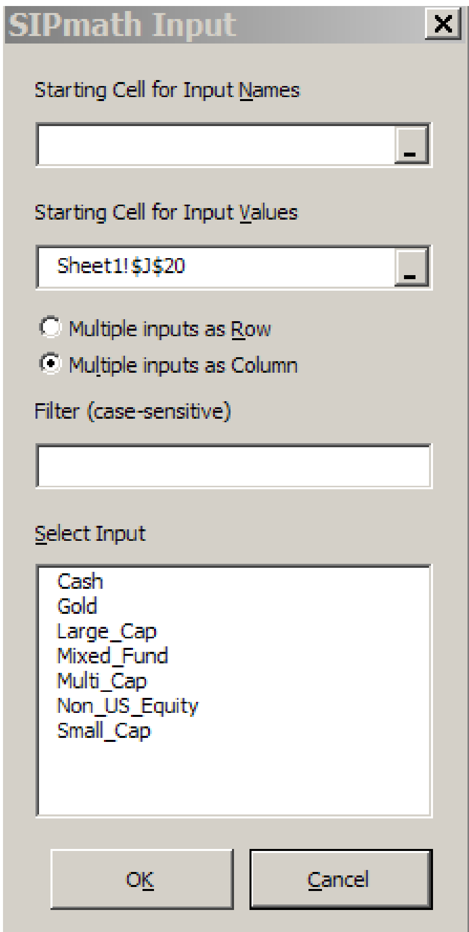
Clicking the Intialize button on an existing model will give you the option to reinitialize existing models or allow the number of trials to be changed in Random Mode.

6.2 DEFINE INPUTS

For use in the SIP Library mode only





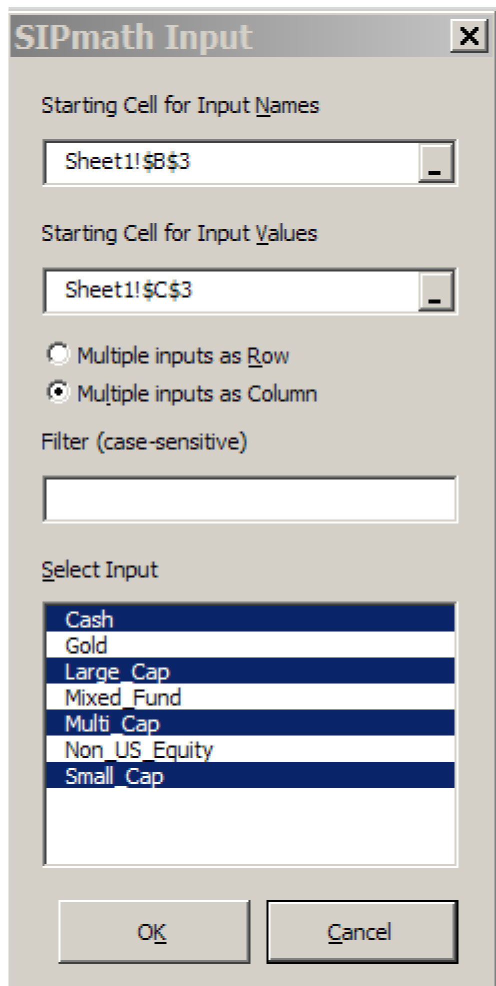


Use the *Define Inputs* tool to tell the SIPmath Modeler Tools which distributions to use from the library chosen during initialization, and which cells the model will use to get input trial values.

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

The Filter box can facilitate your search when you have a large number of variables.

Use the shift and control keys in the usual way to select multiple distributions. Click OK when you're done. You may optionally specify



cells into which to print the variable names.

The values in the cells are from the SIPs in the library, usually the first trial. The `PM_Index` default value is 1, but you may view any value with the Trial Info button.

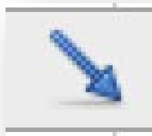
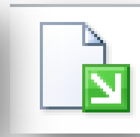
If you look at the formulas under the values, you'll see that the *DefineInputs* tool has set them to something like

```
=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.

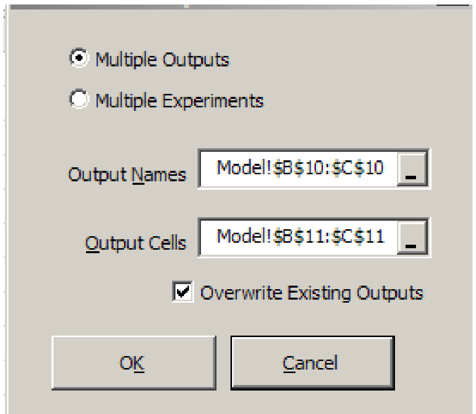
If you checked the Create Sparklines for Inputs box in the Initialize dialog, you will see that a light green sparkline histogram has been placed in the cell with the `INDEX()` formula, and the cell has a black border drawn around it. An example is shown below.

Cash	Gold	Large_Cap	Mixed_Fund	Multi_Cap	Non_US_Equity	Small_Cap
0.019945362	-0.001127084	-0.144446604	-0.144579575	-0.131532497	-0.157626654	-0.132836872

Define Outputs

Define Outputs is actually two different dialogs: Multiple Outputs, which



is displayed in the examples, or Multiple Experiments, for "What If" analysis. The dialogs work the same way in both SIP Library and Random mode, but the Experiments setting is less informative in Random mode.

Link the model outputs to the data table

Multiple Outputs runs a single set of trials through one or more output cells, each of which will correspond to a column in the PMTable sheet. This dialog lets you specify names for the output distributions and the cells containing the formulas to be simulated.

If you checked the "Create sparklines for Outputs" box in the *Initialize* dialog, then the cells with the output formulas will have a light blue sparkline histogram and a black border.

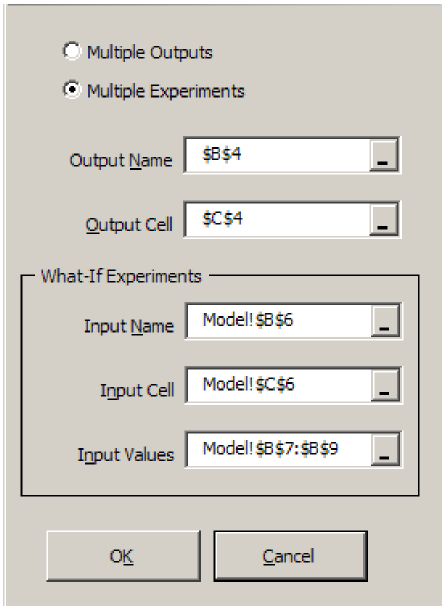
Define Outputs will put a named output column range in the data table for each cell and associated name that you specify. These names can then be used to refer to the output distributions for creating graphs and summary statistics using native Excel formulas anywhere in your model.

You may specify ranges of output cell. If you check "Overwrite Existing Outputs," a new table will be created; 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 shot. Also, the output cells don't have to be contiguous and can be entered as Excel lists (comma-separated).

You can add to the existing outputs.

6.2.1 Multiple Experiments: What-If Analysis

You get the second dialog by selecting "Multiple Experiments." This is for "What If" experiments on a ***single*** output cell, in which you run a set of experimental parameter values through a numeric cell in your model and record the output distribution for each value. For example you might

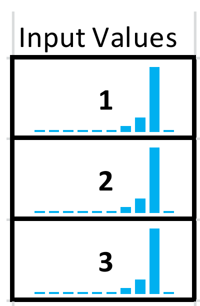


simulate the cost of managing inventory under uncertain demand for 5 different stocking levels.

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.

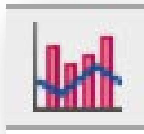
Note that the experiment input values are NOT the same as the input distributions you specified in the "Define Inputs" dialog, which will still drive each simulation. The "What-If" values will drive multiple simulations and the output simulation columns are the distributions of the result of each experiment.

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. For example if the output name was "Cost", the input name was "Stock", and the input stocking levels were 10,15 and 20, then the names in the PMTable sheet would be: Cost_Stock_10, Cost_Stock_15, and Cost_Stock_20.

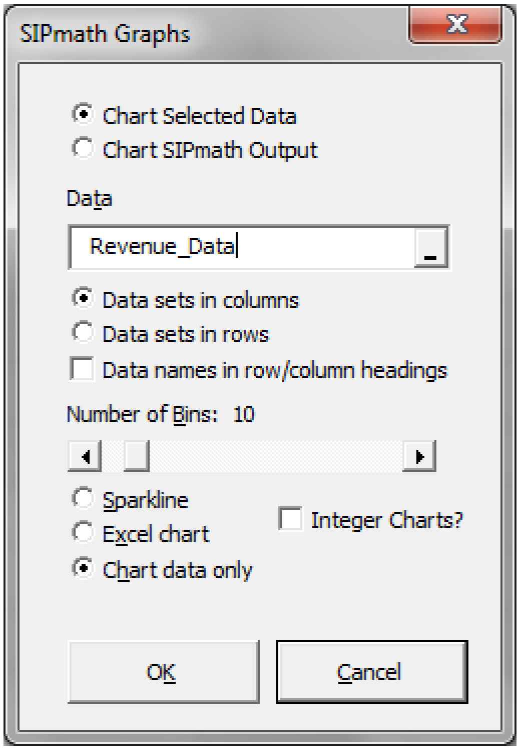


If you checked the "Create Sparklines for Outputs" box in the "Initialize" dialog, then a light blue sparkline with a black border will be created in each of the input value cells.

6.3 GRAPHS



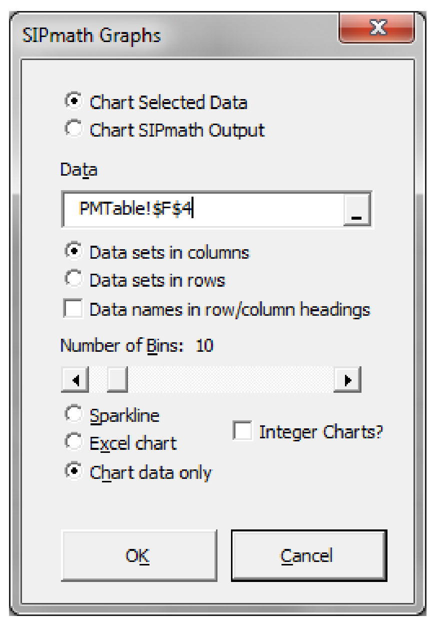
The *Graphs* button lets you create histograms and cumulative graphs from



either ***pre-existing data ranges*** or ***SIPmath Output Cells***. For pre-existing data, select a range of values to be charted. You may either: 1) select that range before pressing the *Graphs* button, 2) type the range specification or name of the set in the Data field manually, or 3) click in the Data field and select the range to be charted. For this example, we have chosen a range called "Revenue_Data."

Graphs present clear results

Graphs function makes a guess as to whether the data is grouped by columns or by rows, based on the shape of the selected region. If the guess is incorrect, you can correct it by clicking the appropriate radio button.



Graphs also guesses whether or not your data sets have names in a heading column or row. If you have data in headings, *Graphs* will use those names when creating range names to be used by charts. Otherwise, the range names will be based on the address of the selected data.

The "Number of Bins" slider sets the number of histogram bars or the number of points on the cumulative curve.

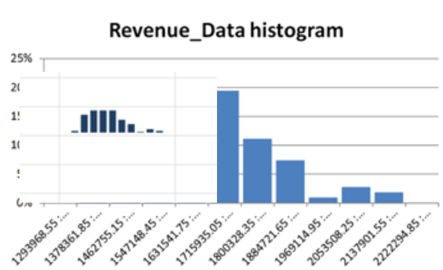
If you check the Integer Charts checkbox, *Graphs* will assume the data to be charted consists of integers (whole numbers only), such as 1, 2, 87, and −452. The number of bins will be automatically determined by *Graphs*, and the slider control will be greyed out when you check “Integer Charts.”The histogram will have one bar for each whole number between the smallest and largest whole number in the data. The bars will be labeled as whole numbers.

Note This option is not recommended if the range of data from smallest to largest value is more than 100.

If you select “Sparkline[2]” or “Excel Chart” the dialog will expand to ask you which graphs you want and where you’d like them put.

If you are charting more than one set of data, additional charts can be placed in a row to the right of the first chart, or in a column below the first chart. You can select which with the radio buttons.

Pressing OK will create the chart data and Sparkline or Excel chart, if any, in the specified location. Here are examples of what a Sparkline and chart might look like once created:

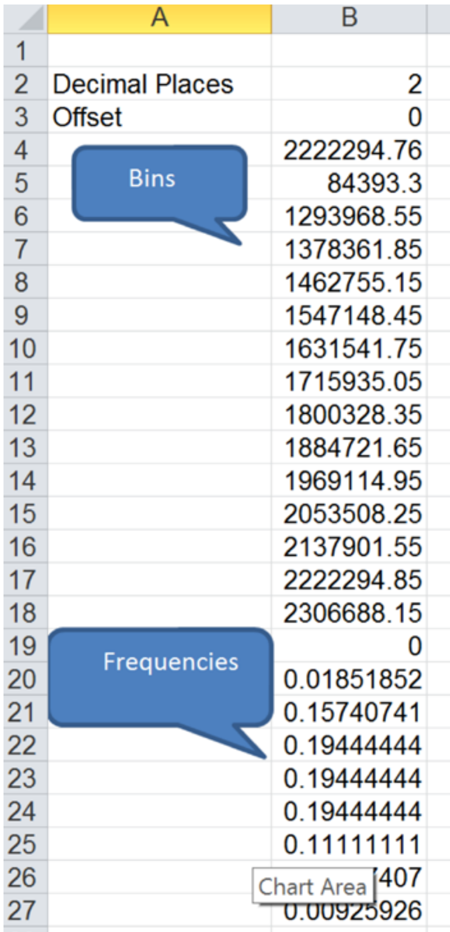


6.3.1 SIPmath Chart Data

When you initialize a model or click “Ok” to close the *Graphs* dialogue box under the Select Data option, a new sheet in the workbook will appear titled “SIPmath Chart Data.” Below is a screenshot of example chart data. Column A has been added to show what the numbers represent, the chart is created without labels.

---

[2]In Excel 2010 or later.



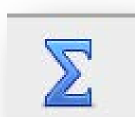
	A	B
1		
2	Decimal Places	2
3	Offset	0
4		2222294.76
5		84393.3
6		1293968.55
7		1378361.85
8		1462755.15
9		1547148.45
10		1631541.75
11		1715935.05
12		1800328.35
13		1884721.65
14		1969114.95
15		2053508.25
16		2137901.55
17		2222294.85
18		2306688.15
19		0
20		0.01851852
21		0.15740741
22		0.19444444
23		0.19444444
24		0.19444444
25		0.11111111
26		407
27		0.00925926

To make a chart, the *Graphs* tool must first set up a small table in the "SIPmath Chart Data" sheet with a row for each bin. The columns contain the bin and percentile values that, respectively, become the X and Y axis inputs to the histogram chart.

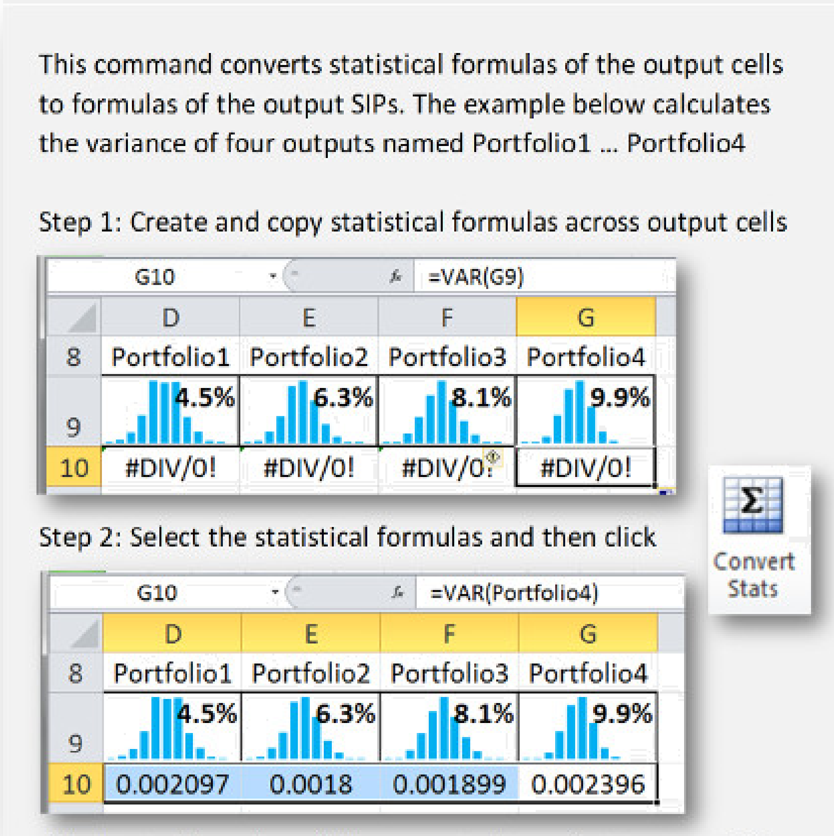
Data in this workbook can be manually edited to change the way your output histogram on the other sheet looks. Bins can be manually edited, an offset can be applied, the number of decimal places allowed for class widths can be changed, and the histogram bar sizes (frequencies) can be altered.

Selecting "Chart Data Only" in the *Graphs* box will make the table without creating a chart. You can then use this to create your own charts using Excel's *Insert Chart* tool. The ranges for the bins are suggestively named (e.g. Profit_bins, Profit_freq, assuming your data had a heading of "Profit".).

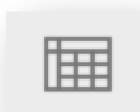
6.4 CONVERT STATS

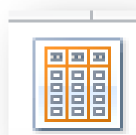


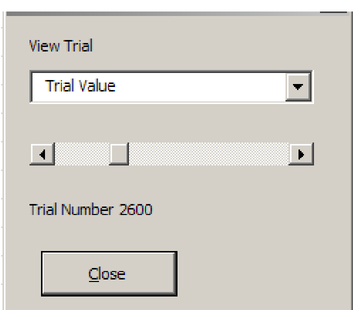
The *Stats* tool is designed to simplify the application of statistical formulas such as Var(), Percentile(), etc. to model outputs. Output variables are essentially SIPs (arrays of observed realizations), and must be referred to by the range names that were automatically created when they were defined (see *Define Outputs* above). See example below. If you forget how to use this button, click on any cell without a formula and click the button. The explanation below will appear.



6.5 TRIAL INFO

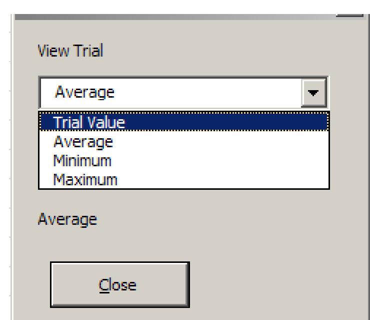






The *Trial Info* tool is only applicable in SIP Library mode. It lets you step through the simulation trials one at a time.

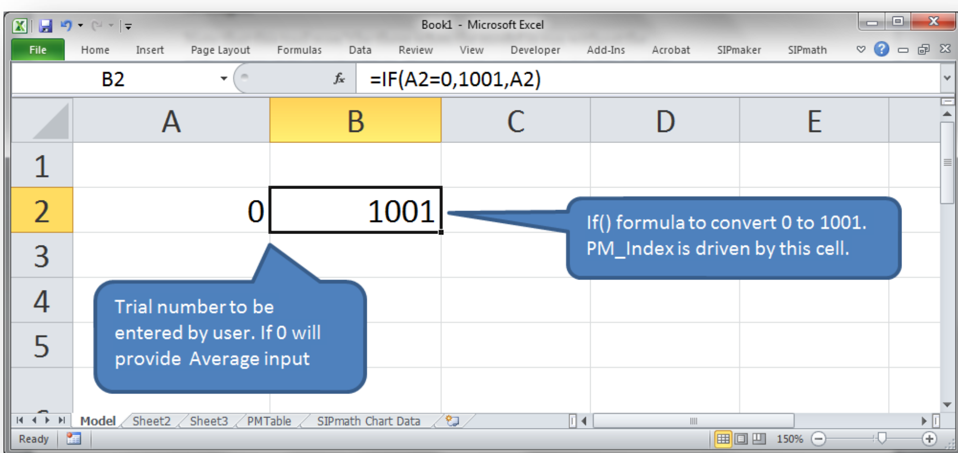
Wherever you set the trial number, you can see the corresponding values in the input cells and the corresponding result values in the output cells. *Trial Info* is actually incrementing the number in the PM_Index range on the PMTable sheet, and can be done manually.



If the SIP Library contains metadata extensions to the distributions, you can also use those to instantiate the model. 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, commonly called the Flaw of Averages.

Note 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 let them change the `PM_Index` range and include a control to access the metadata at the ends of the SIPs if desired. Suppose, for example that the “Average” input is stored at location 1001. Then it is convenient to create two cells as follows to drive `PM_Index`.



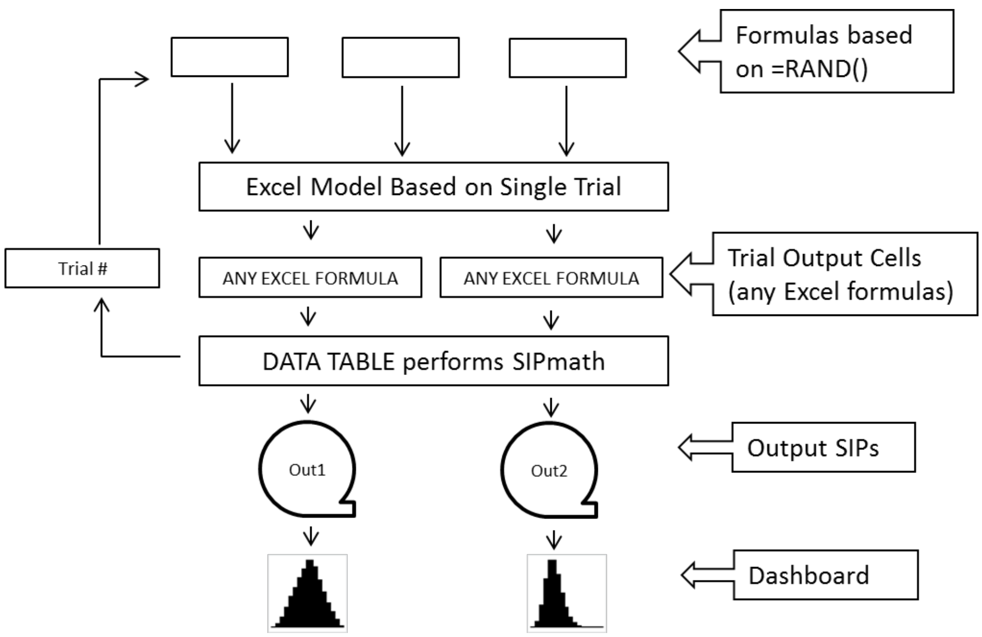
7 SIPMATH ARCHITECTURE

7.1 OVERVIEW

NOTE that for interactive models in either mode, Excel's calculation mode must be set to automatic. Also do NOT use =Indirect or =Offset in your models or they will be greatly slowed down.

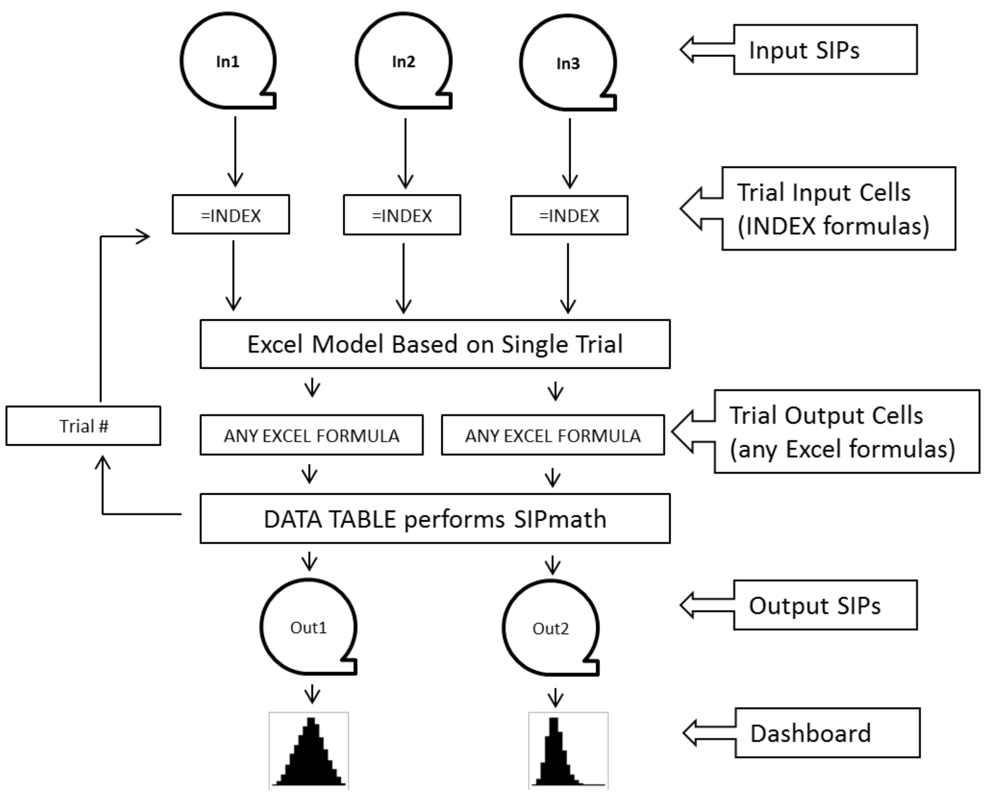
Random Mode

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



The Data Table function iterates the model through the trials of random inputs and collects the results in arrays in the PMTable sheet. This is accomplished entirely with built-in commands and without using Visual Basic. So once created, such models may be run by anyone in native Excel without the **Tools**. The SIPmath Modeler Tools are a set of macros that merely facilitate the development of such models 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 function causes the simulation to run the model for the number of iterations stored in the `PM_Trials` variable. **NOTE this should be the number of trials only and should not include the number of Meta Data elements**. Data Table computations are very efficient, and for medium sized models, thousands of calculations may be performed nearly instantaneously.

SIP Library Model



SIPs In

SIPs Out

In SIP Library Mode the input formulas driven by the `RAND()` formula are replaced with `INDEX()` formulas that step the model through the trials of input SIPs and collect the results into output SIPs.

7.2 STOCHASTIC LIBRARIES

When not using random mode, input distributions come from *Stochastic Libraries* – Excel worksheets containing SIPs and SLURPs (coherent sets

Input SIPs come from libraries

of SIPs).There are some specifics that must be met for the SIPmath Modeler Tools to use them:

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

2 Optionally, there can be metadata 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_Meta` containing the names of the metadata elements, and another named range `PM_Meta_Index` with the displacements to the metadata. These displacements must be common to all the SIPs in the library. For example, assuming a 1,000trials, the two ranges might appear as below:

PM_Meta_Index	PM_Meta
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_Meta` 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.

Important

1 Include the metadata elements in the defined ranges of your inputs.

2 Add the number of metadata elements to variable `PM_TRIALS`.

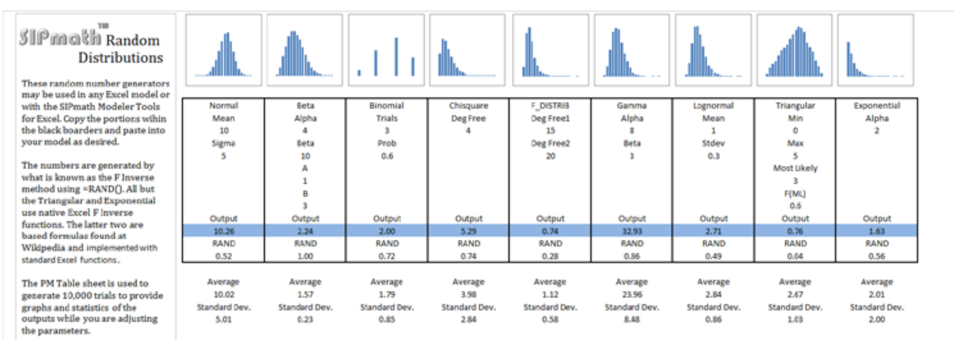
7.3 INPUT CELLS

Input cells interface between the input SIPs and the model

In SIP Library 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 have an 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.

7.4 PARAMETRIC DISTRIBUTION INPUTS

The *F Inverse 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.

7.5 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.

7.6 THE DATA TABLE

The data table indexes the input and records the output.

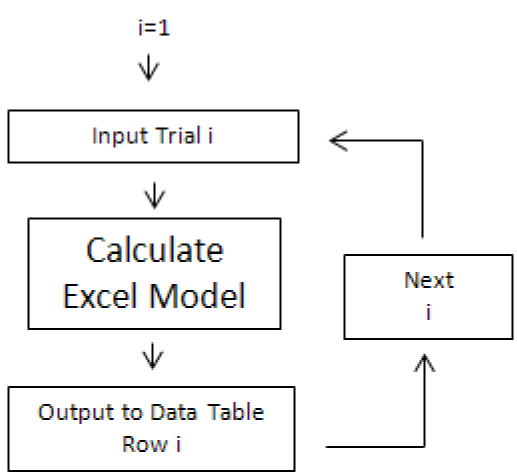
SIPmath sets up the data table on a worksheet called `PMTable` as part of the initialization. This includes defining `PM_Index` to be used as the trial number by the input cell formulas.

It also fills the input column of the data table with a count from 1 to a total number of 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.

To learn more about the Data Table, see Tutorial.xlsx on the SIPmath page of ProbabilityManagement.org

7.7 THE PROCESS



A sheet recalculation event causes the data table to sequentially copy all the numbers from 1 to the number of trials into `PM_Index`. This results in re-evaluating the model for each of the trials in the stochastic library.

The Simulation Loop

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 desired number of trials is reached.

[i] Distribution Processing and the Arithmetic of Uncertainty, Sam Savage, November/December 2012, Analytics Magazine
[ii] Probability Management, Sam Savage, Stefan Scholtes and Daniel Zweidler, OR/MS Today, February 2006, Volume 33 Number 1
[iii] Probability Management 2.0, by Sam Savage and Melissa Kirmse, ORMS Today, October 2014
[iv] Sam L. Savage, The Flaw of Averages, Why we Underestimate Risk in the Face of Uncertainty, John Wiley 2009, 2012