

Box Sampler Tutorial

Using Excel, cells A1 through A1000 each have the formula `=RANDBETWEEN(0,9)`. In cell C1, the formula `=VARP(A1:A1000)` returns the value 8.395775 as shown in Figure 1-1.

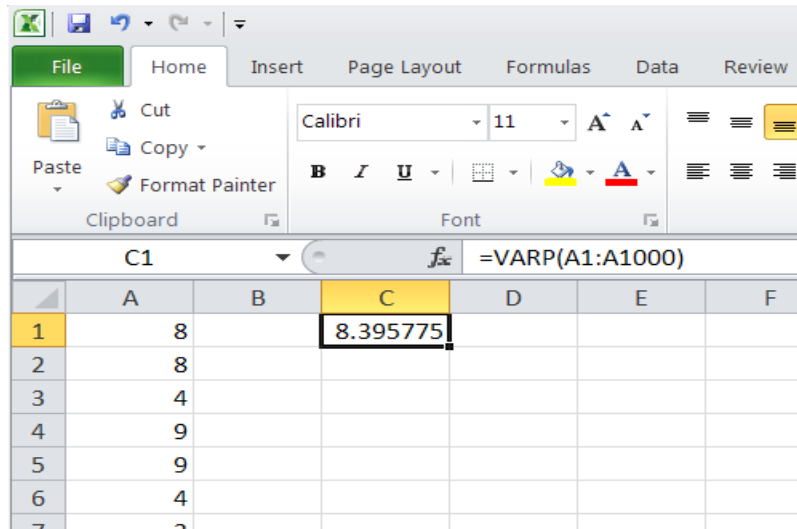


Figure 1-1: Population Variance Using 1000 Values

Now we will use Box Sampler to repeatedly draw 100 resamples of size 10. The population variance of each resample will be calculated as well as the mean of all 100 resamples. Begin with a new Excel workbook, load the Box Sampler add-in, and select the “Add-ins” menu as shown in Figure 1-2.

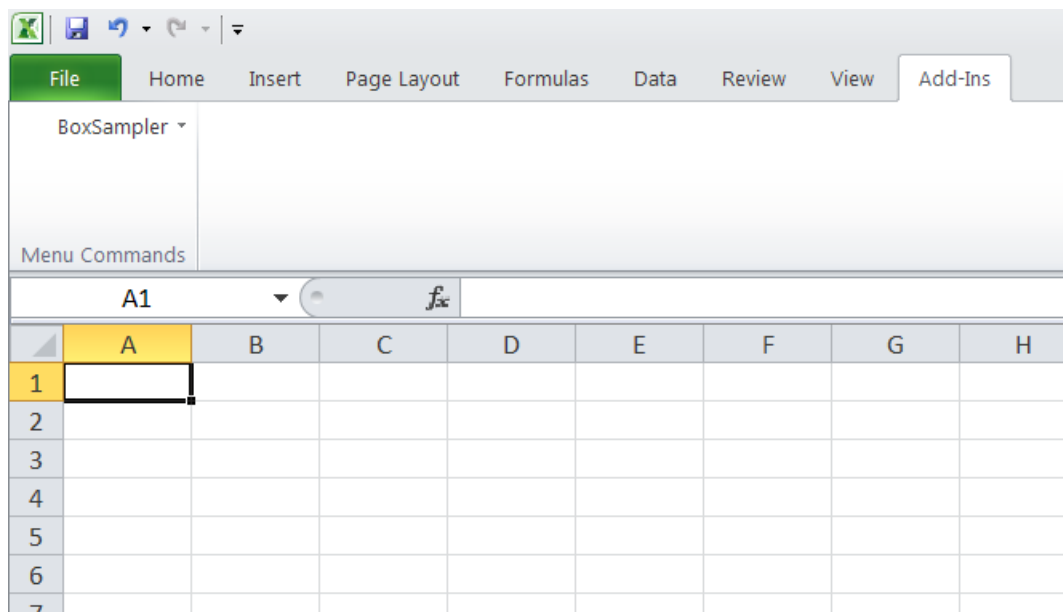


Figure 1-2: The BoxSampler Add-ins Menu

Click on the BoxSampler menu and choose “New Model” (Figure 1-3).

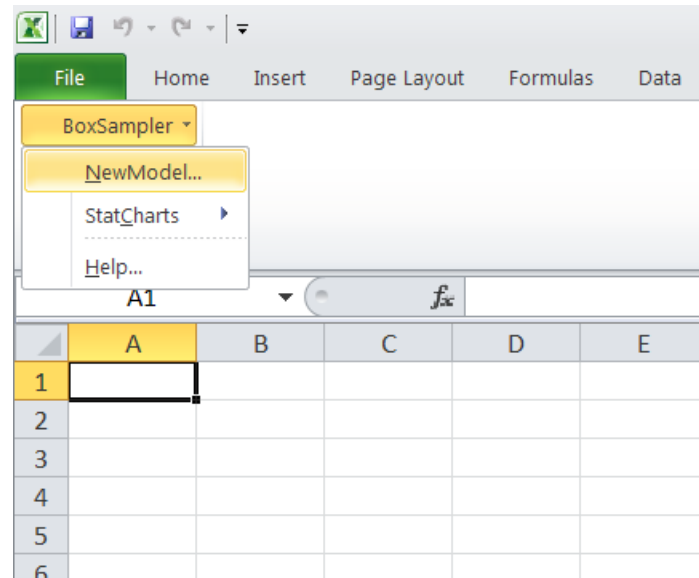


Figure 1-3: Create a New Model

The dialog in Figure 1-4 will appear. For this simulation, accept the default settings by clicking “OK”.

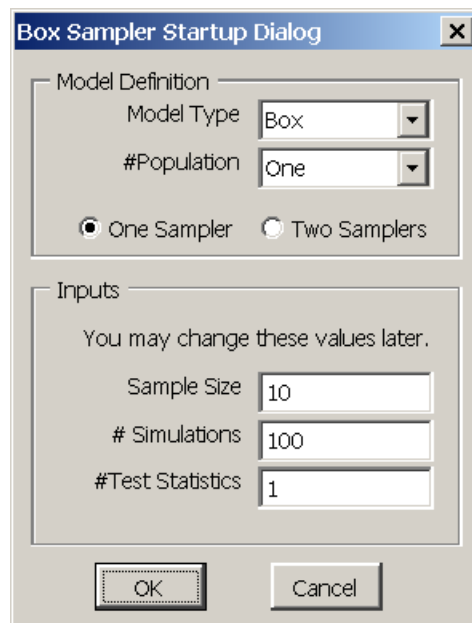


Figure 1-4: Box Sampler Startup Dialog

Box Sampler will create a preliminary model of the resampling simulation (see Figure 5-1). The next step will be to add values and formulas to calculate the population variance of each resample.

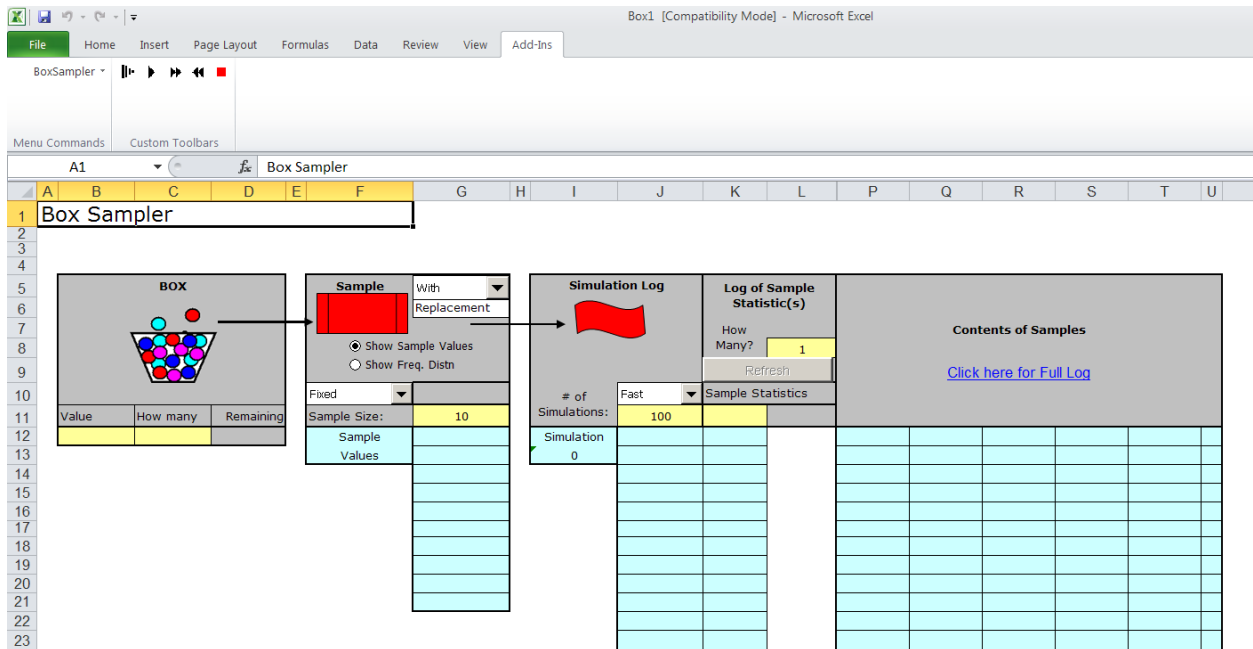


Figure 1-5: The Box Sampler Preliminary Setup

Here are the step-by-step instructions for completing and running the simulation:

- 1) As illustrated in Figure 1-6, populate the “Value” and “How many” columns. The numbers 0-9 in the “Values” column will be resampled with replacement “from the box”.

BOX		
Value	How many	Remaining
0	1	1
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
6	1	1
7	1	1
8	1	1
9	1	1

Figure 1-6: Resampled Values

2) The next step is to calculate the Sample Statistic for the resampled data. Type the formula “=VARP(Sample)” in the Sample Statistics cell, K11 (see Figure 1-7). The “DIV/0” error tells us that we do not yet have data available to calculate a population variance and the denominator, n, is zero. The “Sample” named range points to the result of a single resample and it is this Sample range that is used to calculate the resample “population” variance.

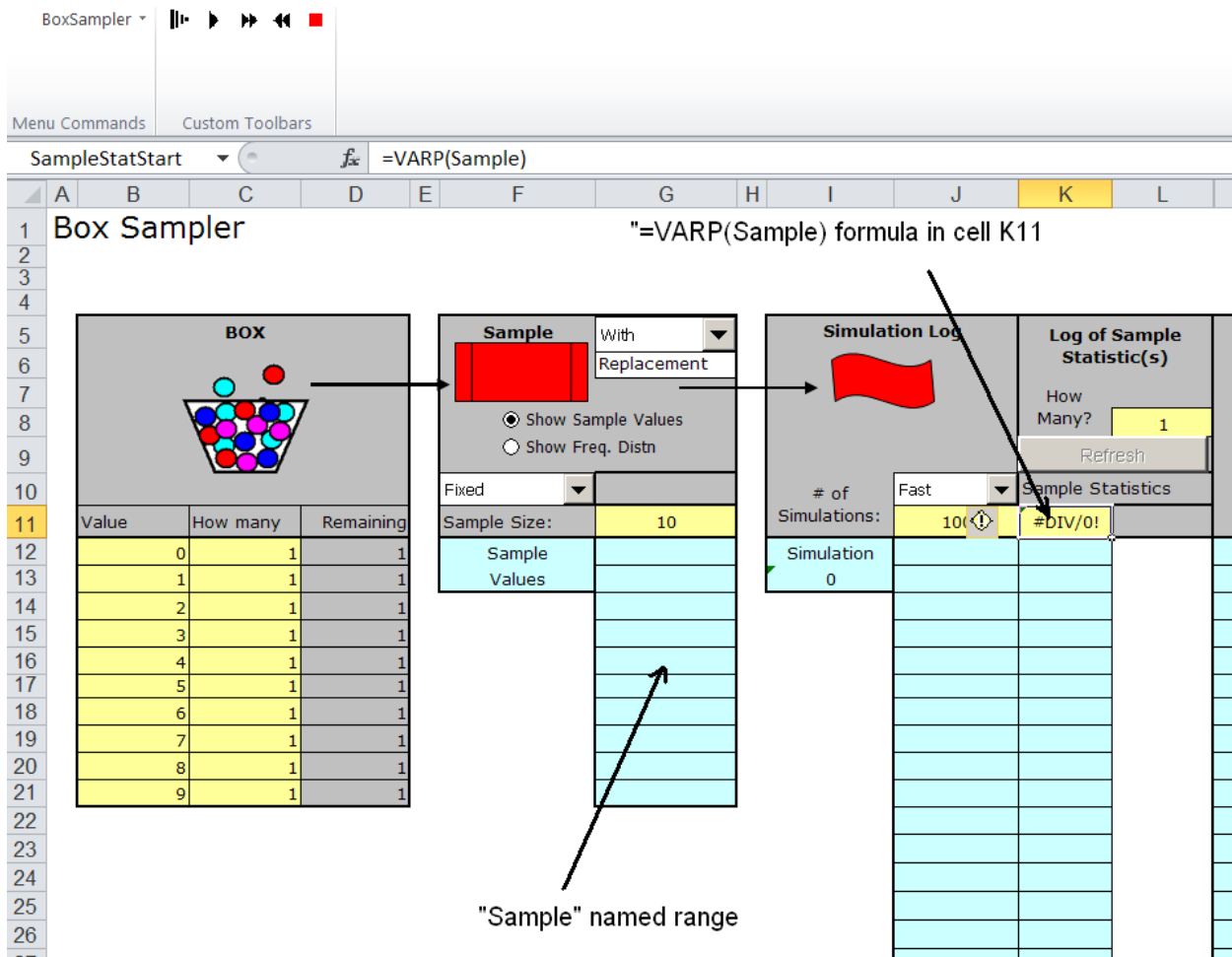


Figure 1-7: Text

3) The exercise asks for the average of all the resampled “population” variances. To calculate the average, one more formula needs to be added to the model. In Figure 1-8, the formula “=AVERAGE(Stat1)” has been entered in cell G24. “Stat1” represents the named 100 cell range containing the variances of each of the 100 resamples.

Figure 1-10 illustrates the results of a complete simulation of this model in Box Sampler.

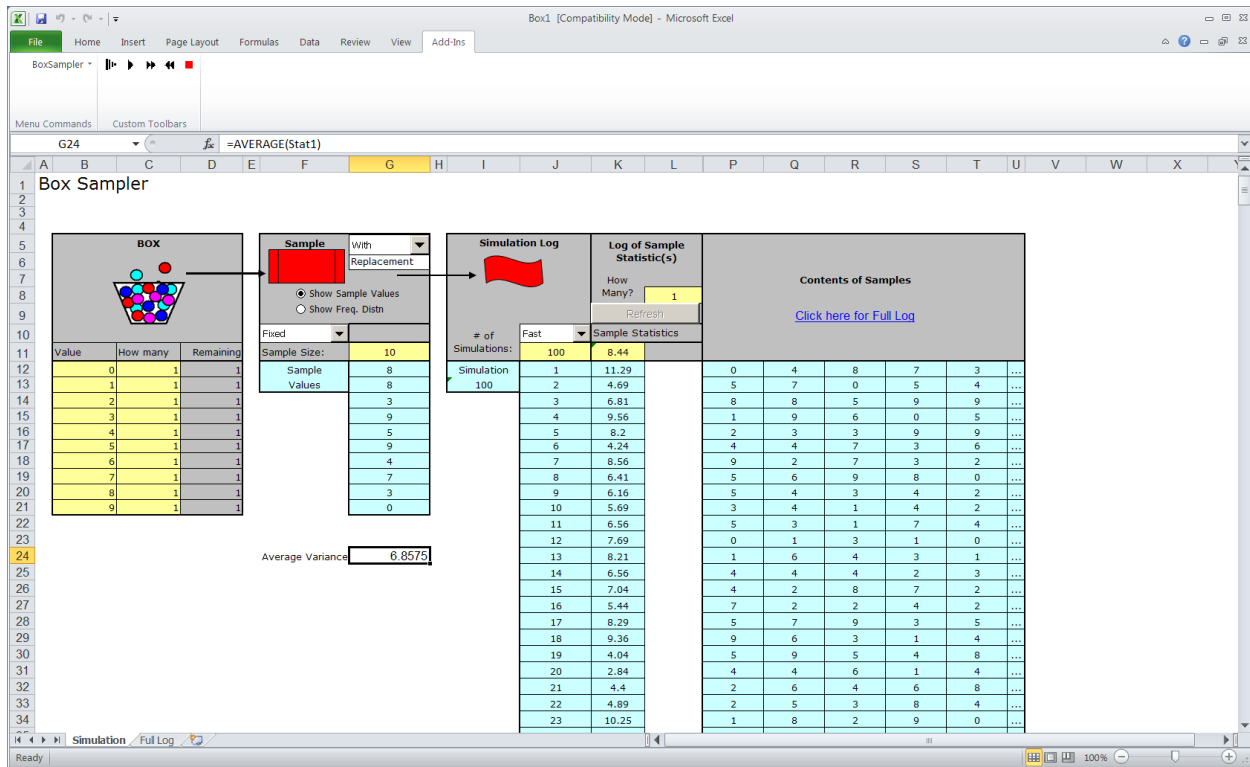


Figure 1-10: The Results of a Single Simulation in Box Sampler

A macro-enabled version of this Box Sampler model may be downloaded by clicking [here](#).