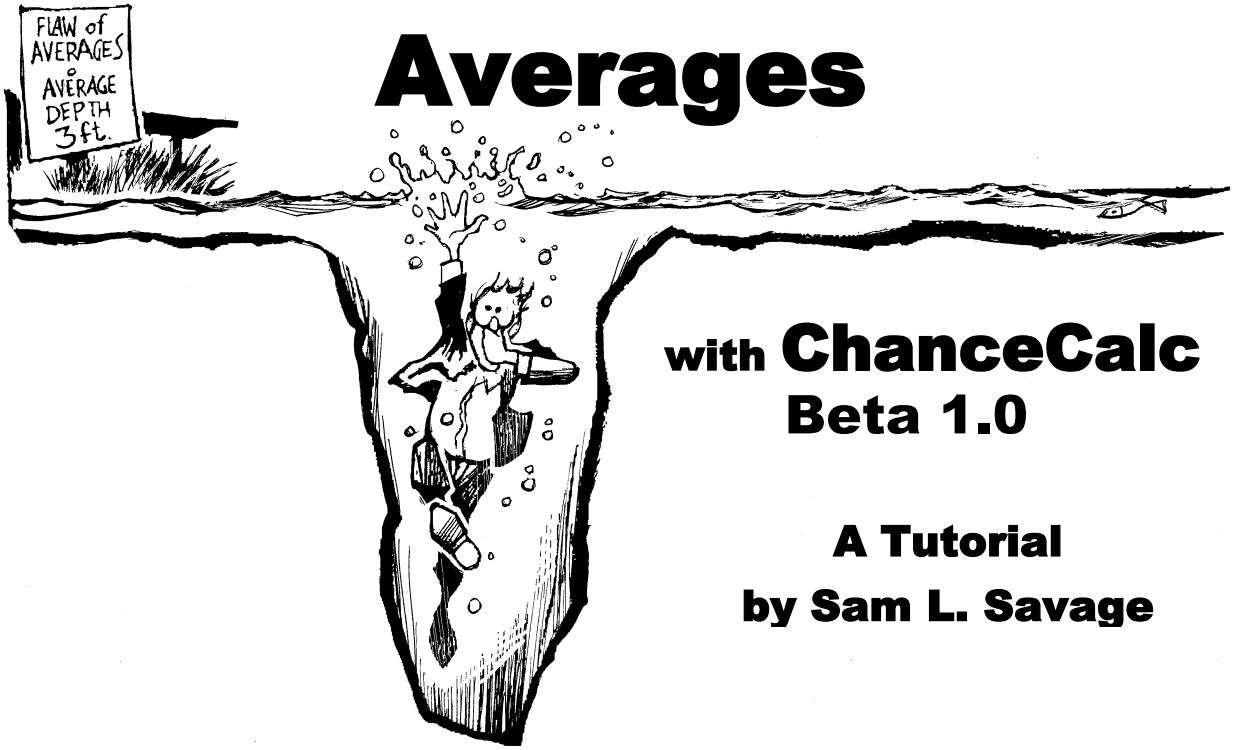
Probability Management

Curing the Flaw of Averages



with ChanceCalc Beta 1.0

A Tutorial

by Sam L. Savage

Built on three complementary breakthroughs in simulation technology

3.0 SIPmath Standard
Powering the arithmetic of uncertainty
From Sam Savage of ProbabilityManagement.org

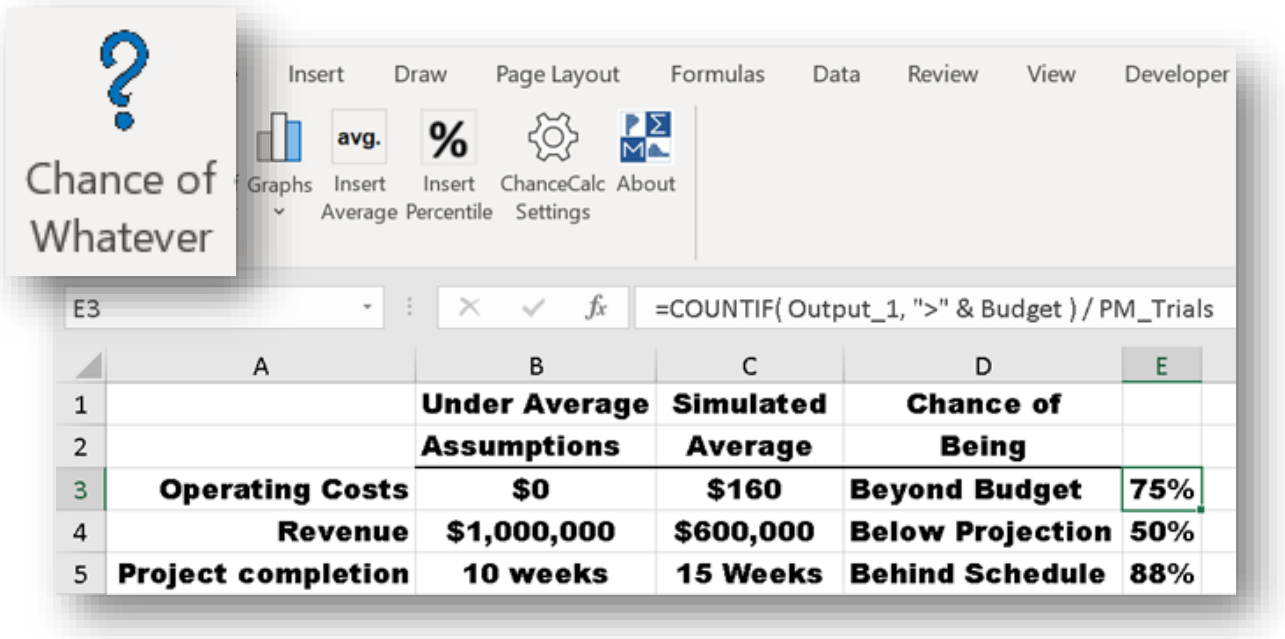
HDR Generator
Ensuring coherence
From Doug Hubbard

Metalog System
Quantifying uncertainty
From Tom Keelin

May 7, 2021

Acknowledgments

I want to thank Doug Hubbard and Tom Keelin, without whose complementary technologies the 3.0 PM Standard and ChanceCalc would not exist. Dave Empey, who has done the bulk of the computer programming that got ProbabilityManagement.org off the ground, developed the code for ChanceCalc. Daniel O'Neil and Aaron Brown assisted in the architecture and library generation. Bridget Cash, Melissa Kirmse, and Mary Claire Meijer have all influenced my approach to expressing probabilistic concepts in a way that does not trigger post traumatic statistics disorder (PTSD).



Contents

Introduction 1

Lesson 1a – Demonstrating the Flaw of Averages 3

Beyond Budget in Operating Costs 3

Nonlinearity 3

Lesson 1b – A Simple Simulation Based on Historical Results 5

SIPmath 5

Chance of Exceedance 5

Chance-Informed Stocking Decision 6

Lesson 2a – Using ChanceCalc and a SIP Library 7

A Capital Investment Problem 7

Loading Chancecalc 7

Loading a SIP from the Cloud 7

Finding the Simulated Average 9

The Chance of Whatever 10

Percentiles 10

Creating Graphs 11

Lesson 2b –A Chance-Informed Decision with Multiple Experiments 14

Performing Multiple Experiments 14

Creating a Chance-Informed Graph 16

Lesson 3 – Using a 3.0 Library 18

A Scheduling Model 18

Virtual SIPs 18

Chance-Informed Schedule 20

Lesson 4 – Failure Analysis, Multiple Identical Variables 21

Predicting the Failure of Aircraft Parts 21

Linking Simulations to Serial Numbers 21

Creating SIP Libraries 24

Examples 24

CDC COVID-19 Hospitalizations 24

The PRECISE Uncertainty Project 25

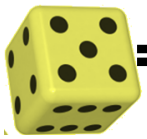
Lesson 5 – Creating a SIP Library of Profit from Lesson-2a 26

Tools for Generating SIP Libraries 27

Introduction

Have you heard about the statistician who drowned in the river that was on average three feet deep? This is just one example of what can go wrong when an average is substituted for an uncertainty.

Two decades ago, I coined the term Flaw of Averages to describe a whole family of such systematic errors, which imply that plans based on average assumptions are wrong on average. It explains, for example, why so many projects are beyond budget, below projection and behind schedule. I refer to these as the Flaw of Averages Trilogy. We will meet all three below. Today, new standards from nonprofit ProbabilityManagement.org are curing this problem by conveying uncertainty as auditable data called SIPs. Instead of containing a single number, SIPs may contain thousands of different numbers representing possible outcomes of an uncertainty. For example, the SIP of rolling a single die might be depicted as shown below.

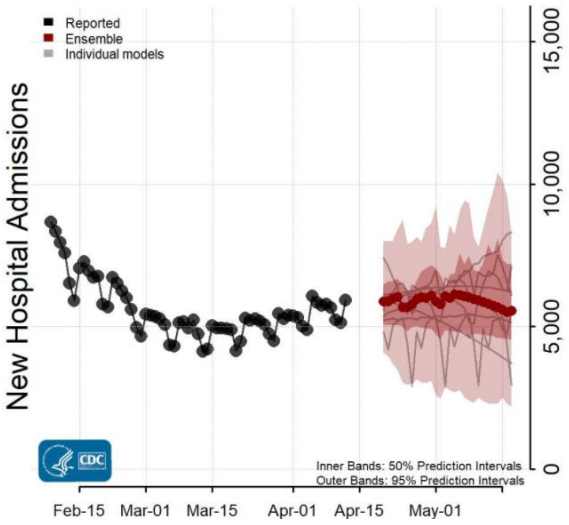


= 6, 1, 2, 3, 6, 2, 5, 5, 4, 5, 2, … 2,4, 1

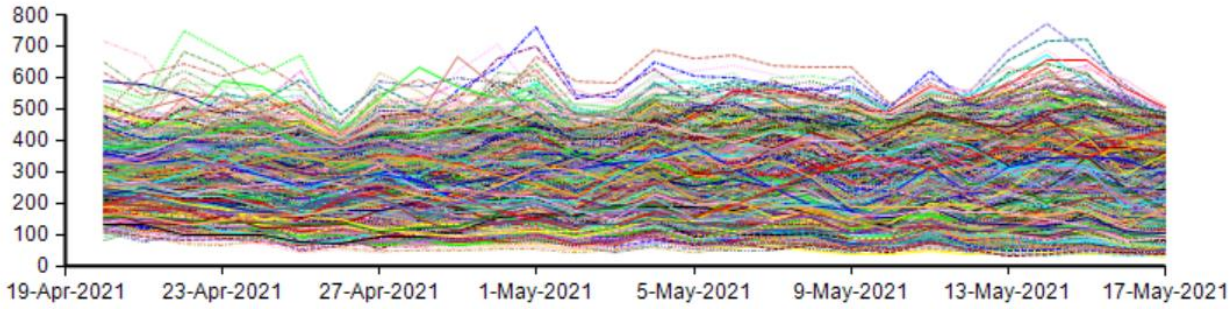
The SIP of a die roll

In general, SIPs require statistical expertise to generate. However, they may be used by non-statisticians in Excel, R, Python or virtually any other software environment. This tutorial will show you how to use SIP Libraries in Excel with a revolutionary add-in called ChanceCalc. You will see how "chance-informed" thinking leads to better decisions in the long run.

How is uncertainty captured?



Nearly all forecasting methods yield information on their own accuracy, but instead of being used this is often discarded. Probability management provides a way to get that "uncertainty" to the decision maker in an actionable form. As an important example, the CDC provides forecasts of COVID-19 hospitalizations that provide ranges of outcomes[i]. In the graph on the right, there is a 95% chance that future hospitalizations will lie within the light red band and a 50% chance it will lie within the darker red band. The darkest red line represents the median: that is, there is a 50/50 chance of hospitalizations above or below this level for each week of the forecast. These colored bands are not of much use to a decision maker, so people are likely to go with the single number median for each time period, which again leads to the Flaw of Averages. But the CDC also provides data that can be used by analysts to generate SIP Libraries through a process known as Monte Carlo simulation. In effect, this displays future hospitalizations in 1,000 parallel universes, in any one of which we might find ourselves, as shown below. This results in 28 SIPs of 1,000 trials, one for each day in the forecast.



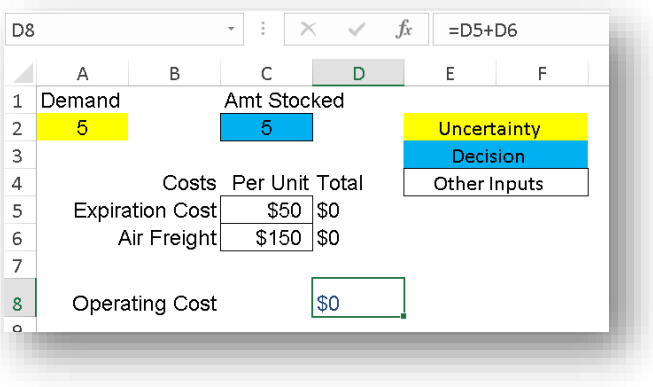
I will begin by demonstrating the systematic errors that result from using single numbers in place of uncertainties.

Lesson 1a – Demonstrating the Flaw of Averages

- Beyond Budget in Operating Costs
- Nonlinearity

Beyond Budget in Operating Costs

Open **Inventory.xlsx**. Here we model the quantity of a perishable drug to stock given uncertain monthly demand. We only consider the operating costs associated with the inventory, and not the cost of the drug itself. Demand is uncertain, but the month over month average is five units (A2) so obviously we will stock five units (C2). However, if monthly demand is less than the amount stocked, then any remaining units must be discarded with an expiration cost of $50 per unit as shown in the formula in cell D5.



D5: =IF(Amt_Stocked>Demand,(Amt_Stocked-Demand)*Exp_Cost,0)

If on the other hand, demand exceeds the amount stocked, the extra units must be airfreighted in at $150 per unit as shown in cell D6.

D6: =IF(Amt_Stocked<Demand,(Demand-Amt_Stocked)*F_Cost,0)

The Operating Cost in cell D8 is simply the sum of the two costs above, at most one of which will be grater than zero.

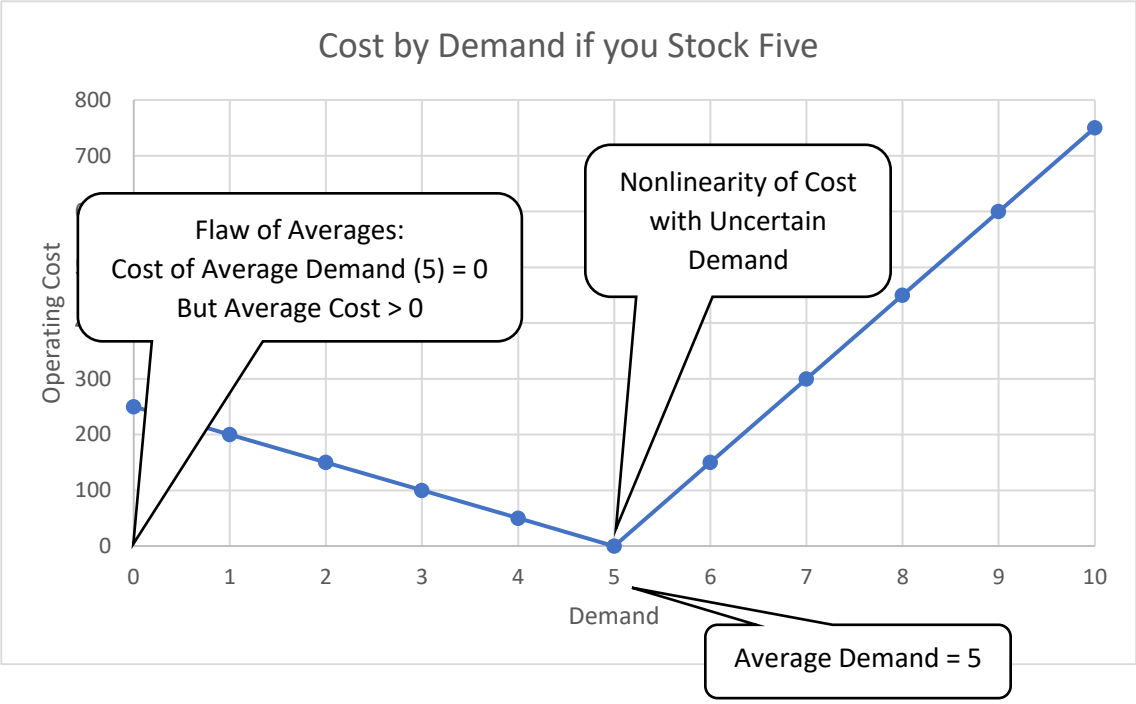
Experiment with the model

Change the demand in cell A2 to be both greater than and less than 5, and observe the formulas in cells D5, D6, and D8.

Suppose the boss comes in and asks what the operating cost will be next month. Many people will plug the average of five into cell A2 and estimate the cost to be zero.

Nonlinearity

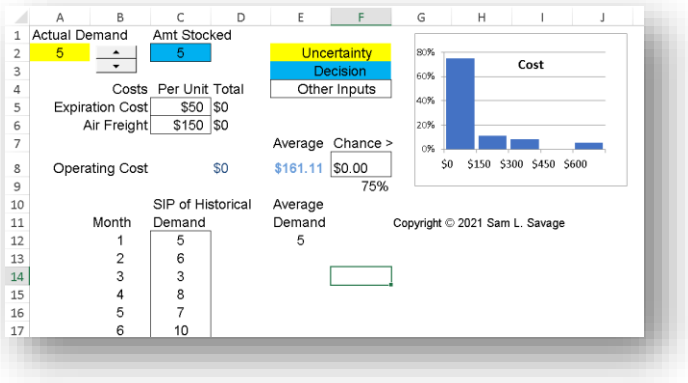
This is our first encounter with the Flaw of Averages. If demand deviates either direction from five, there will be a positive cost. This nonlinearity in the model triggers a mathematical concept called Jensen's Inequality, which results in a particular form of the Flaw of Averages. The cost of the average demand is zero, but the average cost is greater than zero.



The cost is said to be nonlinear with demand because the line above is not straight but bent. Whichever way demand goes from average, cost goes up. This demonstrates why averages can yield systematically bad results, but it does not tell us how much to budget for operating costs on average. We will answer this in the next lesson.

Lesson 1b – A Simple Simulation Based on Historical Results

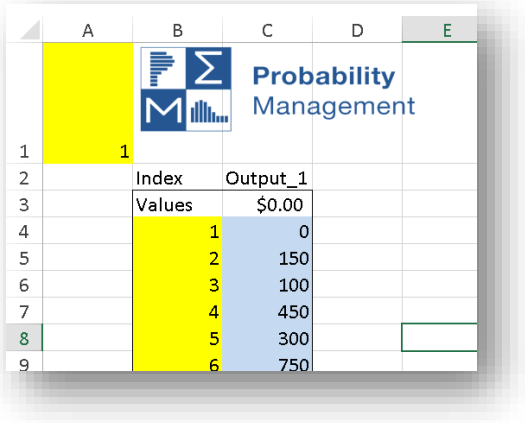
- SIPmath
- Simulated Average Cost
- Chance of Exceedance
- Chance-Informed Stocking Decision



Open **Beyond Budget.xlsx**. In this file we have explicitly modeled the uncertainty in demand with 36 months of historical data in cells C12:C47. This is a simple example of a SIP, or Stochastic Information Packet, analogous to the rolls of the die depicted above. Start by scrolling through a few of the possible demands using the control in column B.

SIPmath

We call this a SIPmath model because it does math with SIPs. That is, every calculation is done 36 times, or simulated, by running each of the historical demands through an INDEX formula in A2. The 36 resulting operating costs appear in a range called Output_1 on a sheet called PMTable. As with the hospitalizations discussed above, imagine observing the problem in 36 parallel universes to get a better idea of what will happen in the one universe in which we live. The Output_1 range was generated automatically by ChanceCalc when the model was created and makes use of the Excel Data Table command.



Simulated Average Cost

Cell E8 of the Model page contains the average of the operating cost simulated over all 36 months of $161.11 and is determined with the formula

E8: =AVERAGE(Output_1)

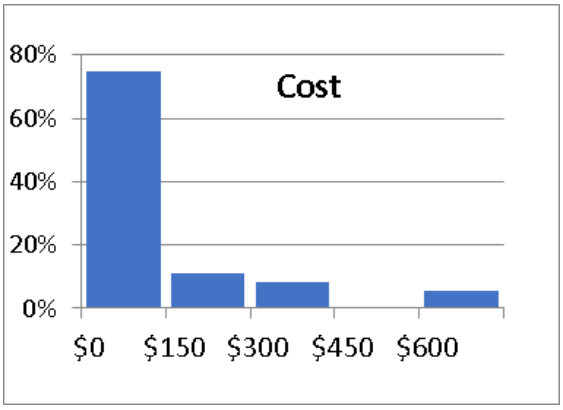
So, to review, if you had estimated operating cost based on the average demand of five, you would have budgeted $0, but on average you need a budget of over $160 per month. This example of the Flaw of Averages highlights the importance of explicitly modeling uncertainty.

Chance of Exceedance

The model also displays the chance of exceeding any specified cost. Cell F9 displays the chance of exceeding the costs entered in cell F8, and uses the formula:

F9: =COUNTIF(Output_1, ">" & F8) / 36

It counts all occurrences of the cost (Output_1) which exceed the number in cell F8, and then divides by 36 to turn it in to a probability. Try entering other costs in cell F8 to determine their chance of exceedance. NOTE: You do not need to remember this formula because ChanceCalc enters it for you.



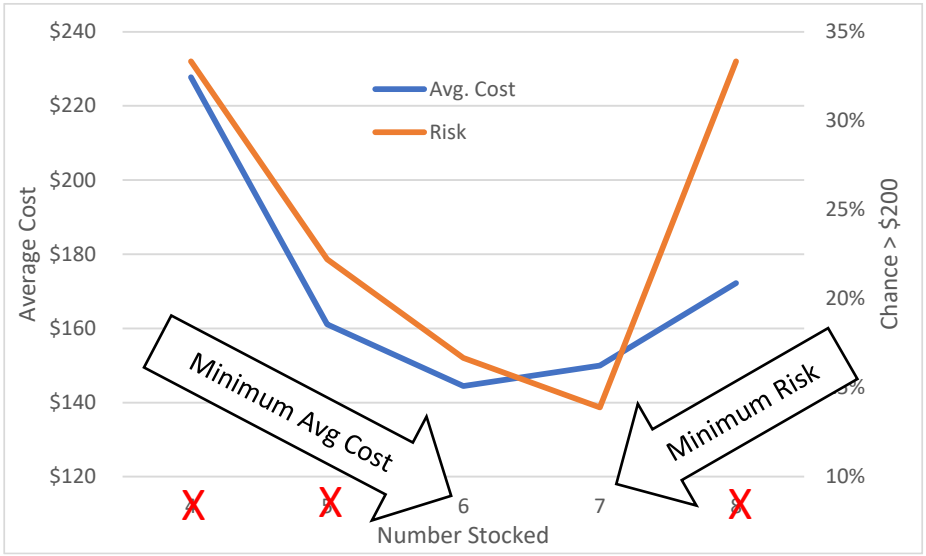
The Histogram
The histogram on the right shows the relative chances of various costs.

Chance-Informed Stocking Decision

In making decisions under uncertainty, we usually face tradeoffs between average outcomes and adverse outcomes. Let's imagine that in this case you were not so foolish as to budget $0 for operating costs but had budgeted $200. Now of course you want to minimize your average cost, but you are also worried about the risk of exceeding $200 on any month.

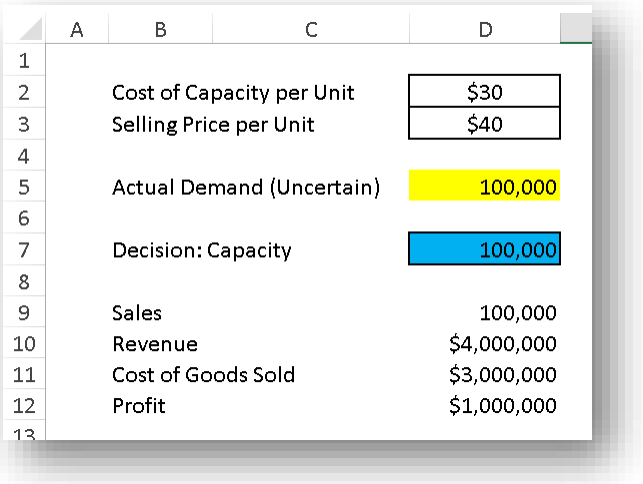
Set cell F8 to $200 and experiment with the quantity to stock in cell C2. Observe both the average cost, and the chance that cost is greater than $200. If you try all stocking levels between 4 and 7, you should observe the results in the table on the left. Remember that before modeling uncertainty we were simply stocking five units and expecting zero costs. The table shows that the minimum average cost occurs with six units, while the minimum risk of exceeding $200 occurs for seven units. Risk is in the eye of the beholder, and which of these two you would choose would depend on your risk attitude. But none of the other levels, including the formerly obvious quantity of five, make sense regardless of risk attitude. In describing this situation to decision makers, I find it useful to create graphs like the one below, which will be explained further on in this tutorial.

Number Stocked	Average Cost	Chance Cost > $200
4	$227.78	33%
5	$161.11	22%
6	$144.44	17%
7	$150.00	14%
8	$172.22	33%



Lesson 2a – Using ChanceCalc and a SIP Library

- A Capital Investment Problem
- Loading a SIP from the Cloud
- The Simulated Average
- The Chance of Whatever
- Percentiles
- Creating Graphs



	A	B	C	D
1				
2		Cost of Capacity per Unit		$30
3		Selling Price per Unit		$40
4				
5		Actual Demand (Uncertain)		100,000
6				
7		Decision: Capacity		100,000
8				
9		Sales		100,000
10		Revenue		$4,000,000
11		Cost of Goods Sold		$3,000,000
12		Profit		$1,000,000
13				

A Capital Investment Problem

Open **Below Projection Example.xlsx**. This model involves a capital investment in manufacturing capacity, given an uncertain demand. For simplicity we will assume that cost of production is zero once the capacity has been paid for up front. It is known that average demand will be 100,000 so the current plan is to purchase capacity of that amount at $30 per unit for a total investment of $3,000,000. If we sell the average demand for $40 per unit, we will gross $4,000,000 in revenue for a profit of $1,000,000. Does that sound familiar? This is usually accepted practice, but given that demand is uncertain, do you think we can really expect $1,000,000 on average? Again, think of parallel universes as we did in lesson 1, but this time there will be 1,000 of them instead of 36.

Explore the model

Start out by experimenting with the demand in cell D5. The nonlinearity in this model comes from the formula for Sales in cell D9.

D9: =MIN(Actual,Capacity)

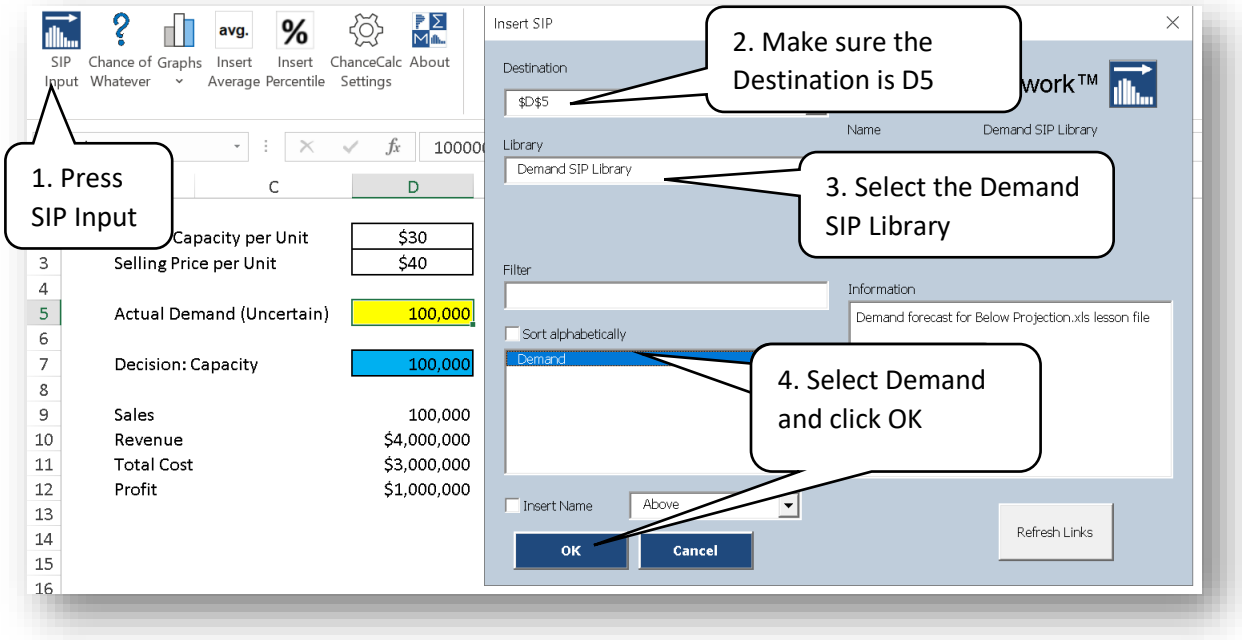
If you purchase capacity for 100,000, that's the *most* you can sell. There is no upside, so $1,000,000 can't be the average profit, but is the maximum possible or best-case scenario. To find the simulated profit we will use ChanceCalc.

Loading Chancecalc

To load Chancecalc, go to the Developer tab (if you do not see the Developer tab, click here). Find and click the "Excel Add-ins" button, then browse to the location of your downloaded copy of Chancecalc.xlam. Click Ok to return Excel. Chance Calc should have appeared as another ribbon option on Excel.

Loading a SIP from the Cloud

Imagine that your marketing department has forecast demand for the product and has captured the uncertainty in a SIP of 1,000 trials. Select ChanceCalc from the ribbon and place the cursor in cell D5, where the SIP will go. Press the SIP Input button, and after establishing a connection, the SIPmath.Network browser will appear.


SIP Input
Chance of Whatever
Graphs
Insert Average
Insert Percentile
ChanceCalc Settings
About
1. Press SIP Input
Capacity per Unit $30
Selling Price per Unit $40
Actual Demand (Uncertain) 100,000
Decision: Capacity 100,000
Sales 100,000
Revenue $4,000,000
Total Cost $3,000,000
Profit $1,000,000
Insert SIP
2. Make sure the Destination is D5
Destination
$D$5
Name
Demand SIP Library
Library
Demand SIP Library
3. Select the Demand SIP Library
Filter
Information
Demand forecast for Below Projection.xls lesson file
Sort alphabetically
Demand
4. Select Demand and click OK
Insert Name
Above
OK
Cancel
Refresh Links

The Result of Loading the SIP

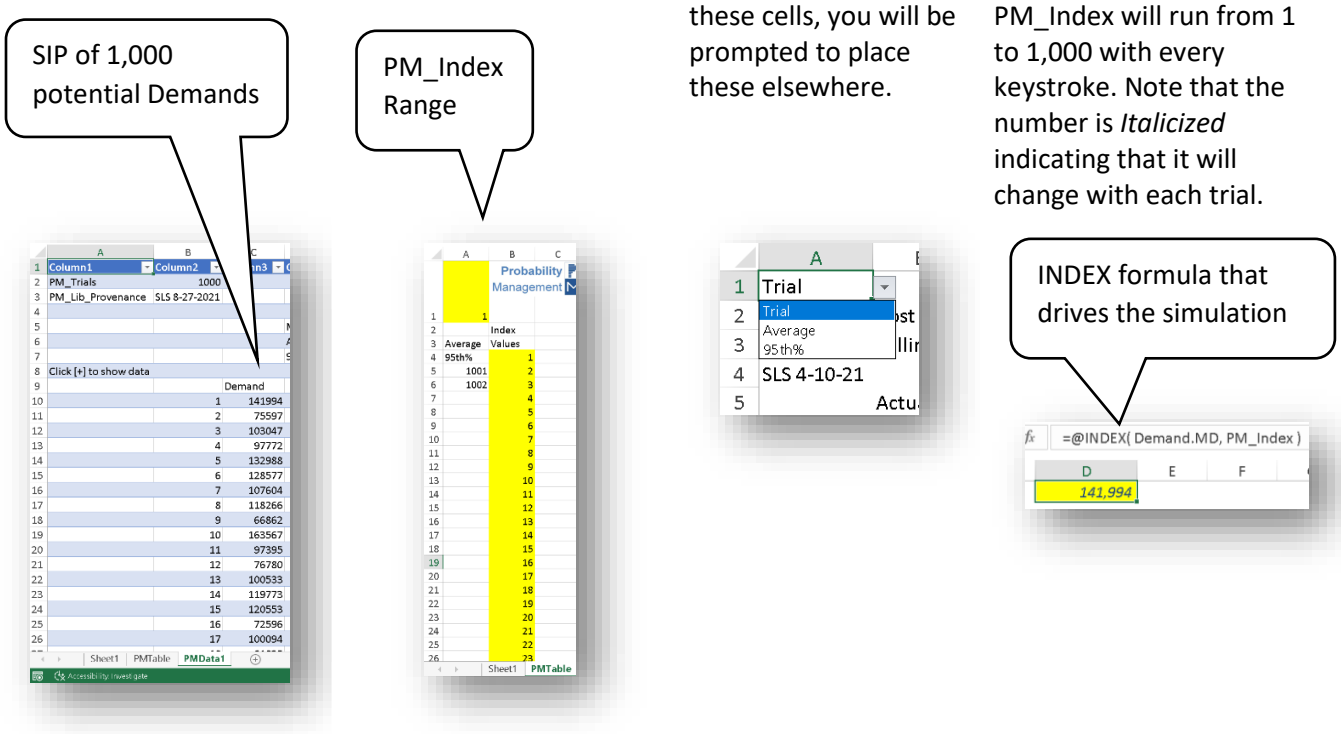
Four things happened when you clicked OK as shown in the table below.

1 A query read the SIP Library into a sheet named PMData1

2 A Sheet named PMTable was added, where the results will be simulated

3 A menu and other details appeared in cells A1:A4. If your own model has things in these cells, you will be prompted to place these elsewhere.

4 Cell D5 now has an INDEX formula that points to the Demand SIP on the Data page. The variable called PM_Index will run from 1 to 1,000 with every keystroke. Note that the number is *Italicized* indicating that it will change with each trial.

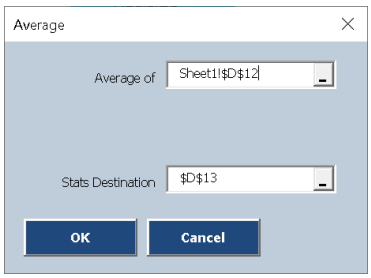


Experiment with the model

1. Look at various trials (parallel universes) by changing cell A2. Most people in the organization are imagining $1,000,000. In that light, trial 9 is sobering.
2. Select “Average” off the menu in cell A1 to observe the Profit given Average Demand.

Finding the Simulated Average

1. Type “Average Profit” in cell B13
2. Place the cursor in cell D13
3. Click the **avg.** button
4. Click in cell D12 while in the Average of field then click OK.

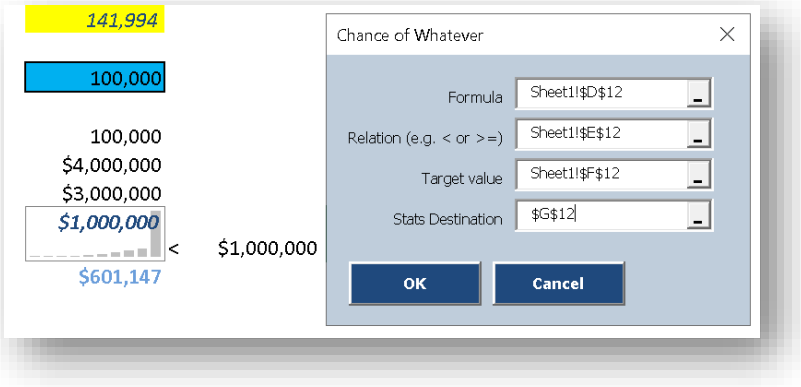


Cell D12 now has a Sparkline graph histogram showing that there is a downside but no upside. I like to increase the row height for these cells so you can see the histogram more clearly. And the Simulated Average Profit is only about $600,000, not the promised million. The Flaw of Averages strikes again!

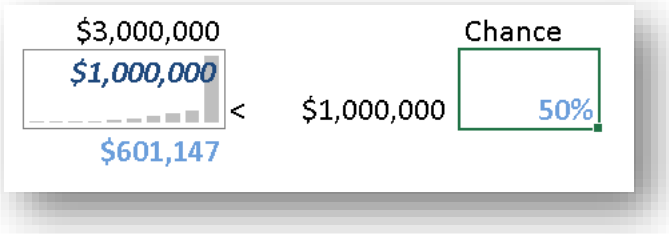
The Chance of Whatever

Let's determine the chance that profit is less than any particular amount.

1. Type "<" in cell E12, and $1,000,000 in cell F12
2. Put the cursor in cell G12 and click the Chance of Whatever button
3. Click in the specified cells in each field and click OK



You will see that there is a 50% chance of making less than $1 million. Next see what the chance is of losing money by typing $0 in cell F12.

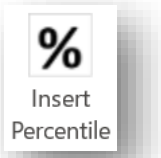


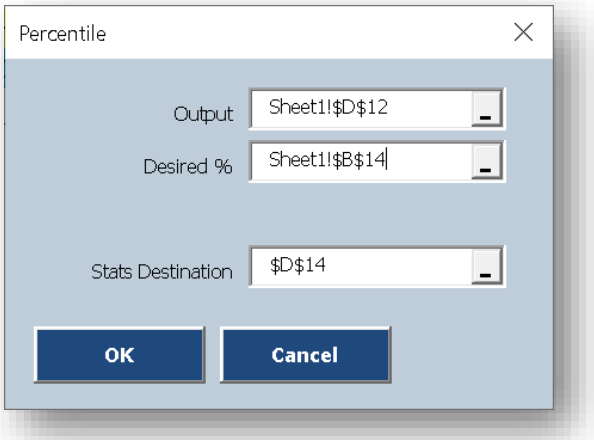
Percentiles

The percentile is the value, or less, that Profit will take on with a given probability. We will start by finding the 10th Percentile of Profit.

1. Enter "10%" in cell B14

2. Type “Percentile” in cell C14
3. Put the cursor in cell D14 and click the Insert Percentile button
4. Click in the specified cells in each field and click OK



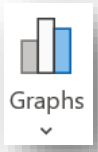


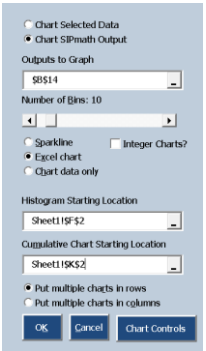
We see that there is a 10% chance of losing about $280,000 or more! Determine the 5th percentile by changing cell B14 to 5%.

I prefer using chances rather than percentiles as I find it a more natural way for most people to think. If you enter -$279,288 into cell F12 you will find that the chance of being less than this number is 10%, confirming what we just learned. Percentiles, however, were easier for mathematicians before computers, and are still in common use today. So, it is important to be able to work with both ways of expressing similar things.

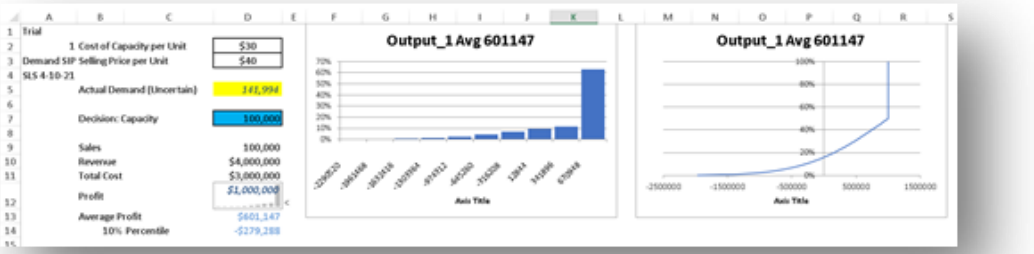
Creating Graphs

ChanceCalc supports Histograms, Cumulative, Exceedance and Scatter plots. Here we will discuss the first two.

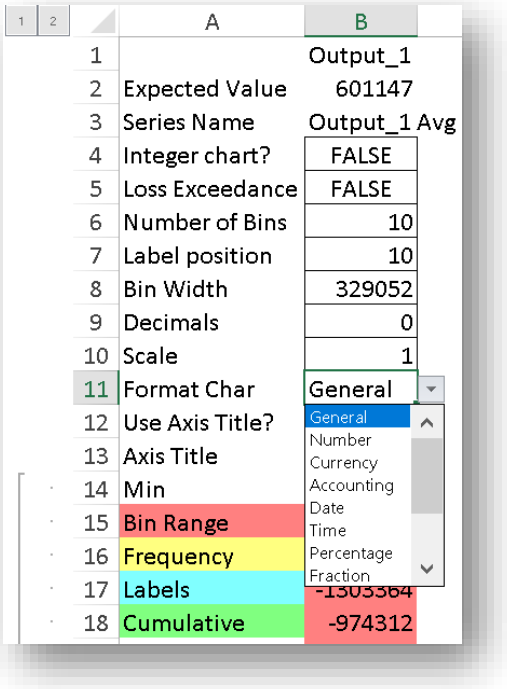




1. Place the cursor in the Profit cell, D12, and click the Graph button
2. Choose locations for the Histogram and/or Cumulative Graphs, then click OK. Note that the locations do not need to be empty cells as the Graphs will be pasted over and can be moved aside.



Many formatting options are available for the graphs on the SIPmath Chart Data Page

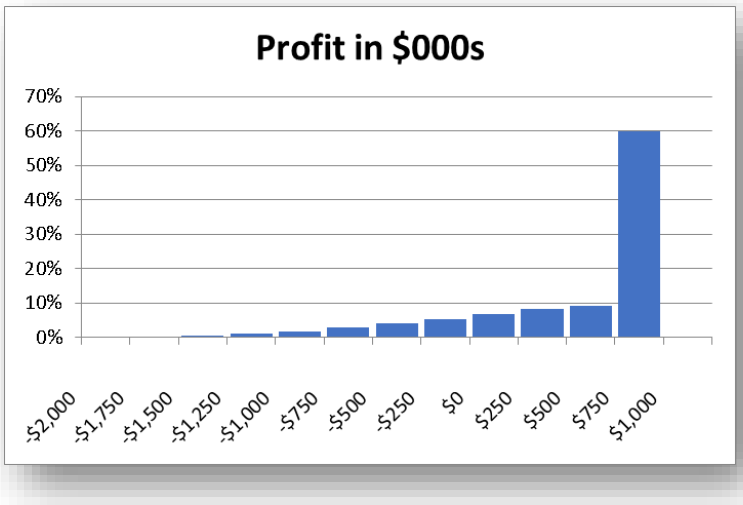


	A	B
1		Output_1
2	Expected Value	601147
3	Series Name	Output_1 Avg
4	Integer chart?	FALSE
5	Loss Exceedance	FALSE
6	Number of Bins	10
7	Label position	10
8	Bin Width	329052
9	Decimals	0
10	Scale	1
11	Format Char	General
12	Use Axis Title?	
13	Axis Title	
14	Min	
15	Bin Range	
16	Frequency	
17	Labels	-1303364
18	Cumulative	-974312

General
Number
Currency
Accounting
Date
Time
Percentage
Fraction

I made the following changes to get the histogram below.

1. Changed Format in row 11 to Currency
2. Removed decimals from Currency Format in row 25
3. Set minimum in row 14 to -2,000,000
4. Set bin width in row 8 to 250,000
5. Set number of bins in row 6 to 13
6. Set Scale in row 10 to 1,000
7. Set Series Name in row 3 to Profit in $000s



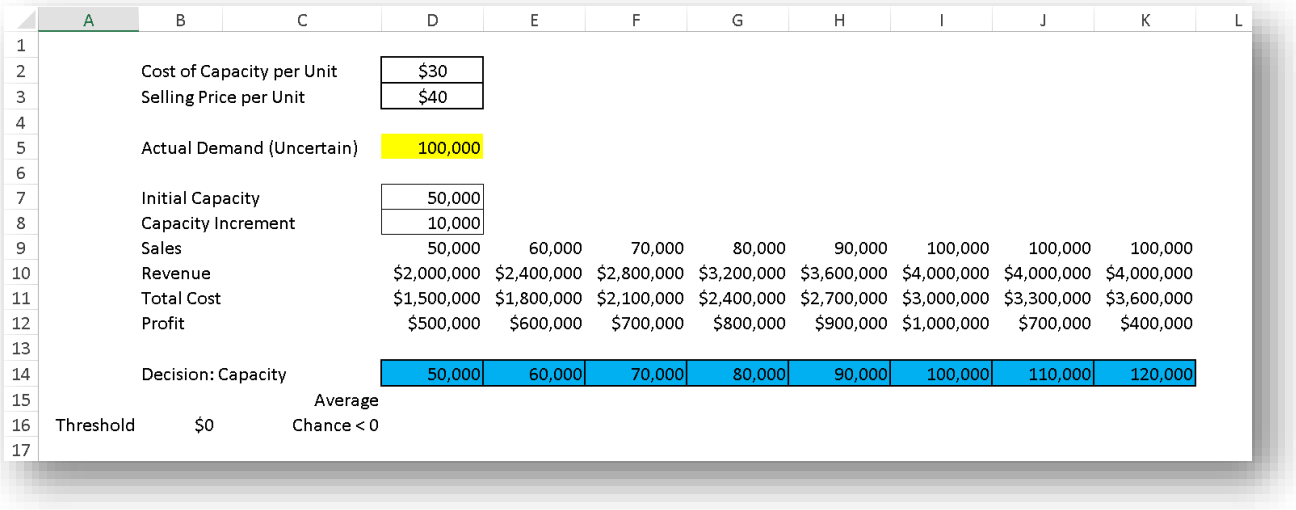
As with Lesson 1b, there are tradeoffs between profit and risk. If you set up your model correctly, you should have a Simulated Average Profit of $601,147 and a 16% chance of losing money. The spike in the histogram above indicates that profit is capped at $1 million because we only invested in a capacity of 100,000 units. It makes sense that if we invested more in capacity, we could take advantage of the possibility of higher demand and increase our average profit. It makes sense but it is dead wrong. Increase capacity to 110,000. Average profit drops to $469,651 and the chance of loss goes up to 24%, so that would be a terrible decision. The next lesson will provide the answers. Save this version of the file under a different name as we will use it again in Lesson 5 on creating SIP Libraries.

Lesson 2b –A Chance-Informed Decision with Multiple Experiments

- Performing multiple experiments
- Creating a Chance-Informed graph

Performing Multiple Experiments

Open **Below Projection Multi Experiment.xlsx** as shown below. This model will simultaneously analyze the average profit and chance of loss for each of eight capacity decisions.



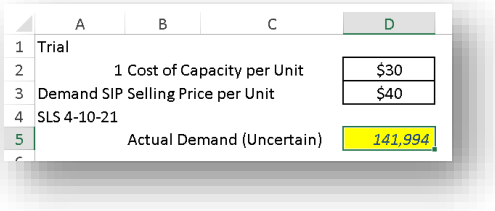
	A	B	C	D	E	F	G	H	I	J	K	L
1												
2		Cost of Capacity per Unit		$30								
3		Selling Price per Unit		$40								
4												
5		Actual Demand (Uncertain)		100,000								
6												
7		Initial Capacity		50,000								
8		Capacity Increment		10,000								
9		Sales		50,000	60,000	70,000	80,000	90,000	100,000	100,000	100,000	
10		Revenue		$2,000,000	$2,400,000	$2,800,000	$3,200,000	$3,600,000	$4,000,000	$4,000,000	$4,000,000	
11		Total Cost		$1,500,000	$1,800,000	$2,100,000	$2,400,000	$2,700,000	$3,000,000	$3,300,000	$3,600,000	
12		Profit		$500,000	$600,000	$700,000	$800,000	$900,000	$1,000,000	$700,000	$400,000	
13												
14		Decision: Capacity		50,000	60,000	70,000	80,000	90,000	100,000	110,000	120,000	
15			Average									
16	Threshold	$0	Chance < 0									
17												

Experiment with the model

1. Inspect the formulas in rows 9 through 12 to see how they depend on the capacity in row 14.
2. Try various Actual Demands in cell D5 to see how the Profit changes in row 12.

Load the Demand SIP

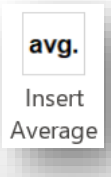
Use the same steps as in Lesson 2A load uncertain Demand into cell D5. The model should appear as shown on the right.



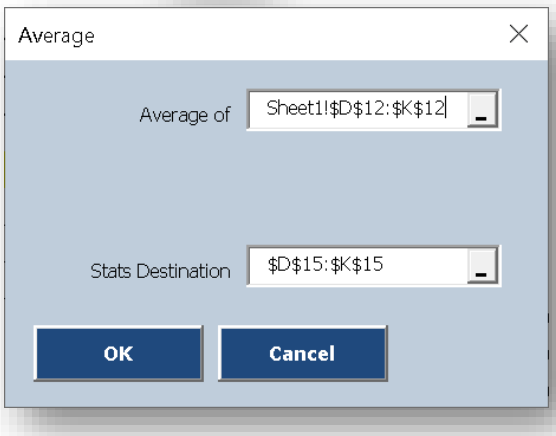
	A	B	C	D
1	Trial			
2		1	Cost of Capacity per Unit	$30
3	Demand SIP		Selling Price per Unit	$40
4	SLS 4-10-21			
5			Actual Demand (Uncertain)	141,994

Performing the Experiment

Find Simulated Average Profit for each capacity.



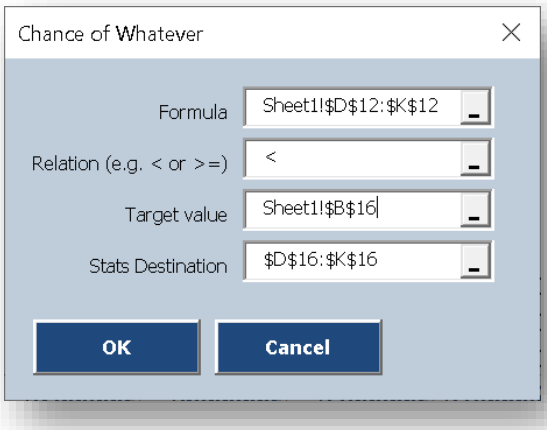
1. Select cells D15:K15
2. Click the Insert Average button and sweep out the Profit cells in the Average of field, then click OK



If you this did this correctly you should see our old friend $601,147 from Lesson 2a in cell I15.

Find the Chance of Loss for each capacity.

1. Select cells D16:K16
2. Click the Chance of Whatever button and sweep out the Profit cells in the Formula field
3. Type a < sign in the Relation field
4. Select cell B16 in the Target Value field
5. Click OK

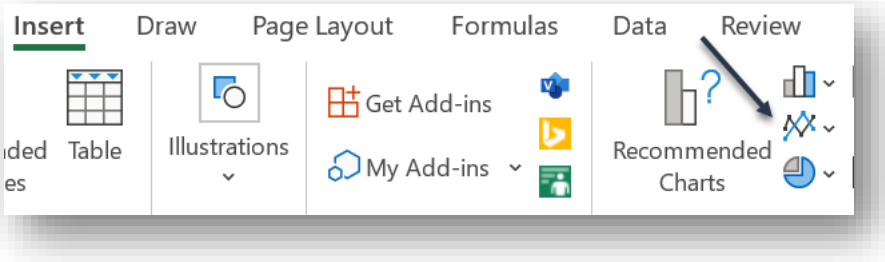


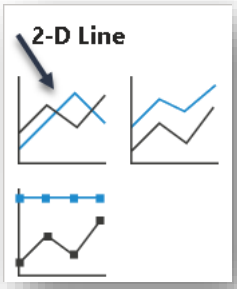
Cells C14:K16 should now contain the results below.

	50,000	60,000	70,000	80,000	90,000	100,000	110,000	120,000
Average	$491,598	$576,846	$643,986	$679,882	$669,651	$601,147	$469,651	$279,882
Chance < 0	1%	1%	3%	6%	10%	16%	24%	35%

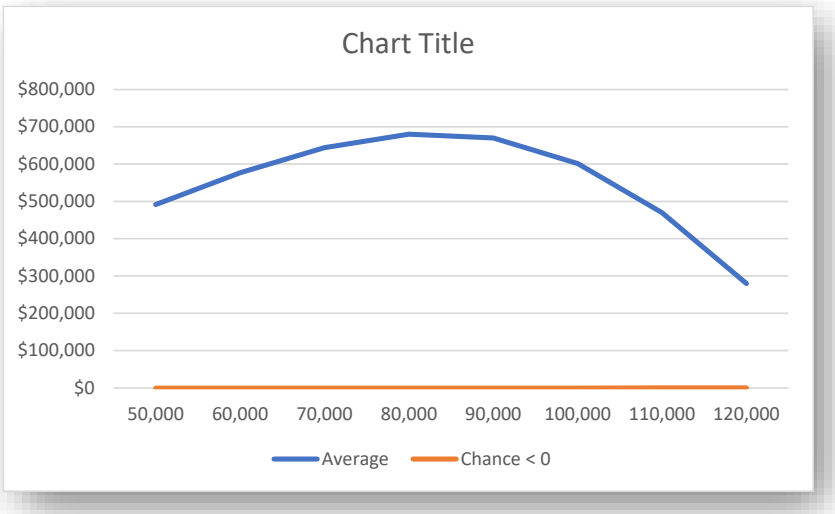
Creating a Chance-Informed Graph

1. Select cells C14:K16, then select the line graph from the Insert menu.

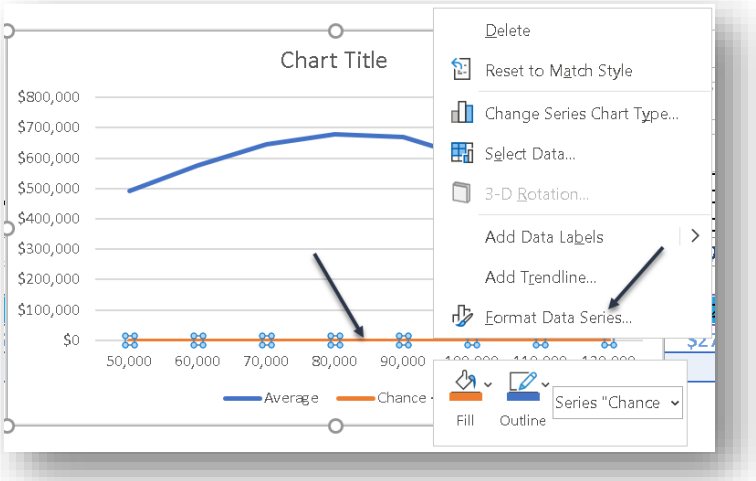


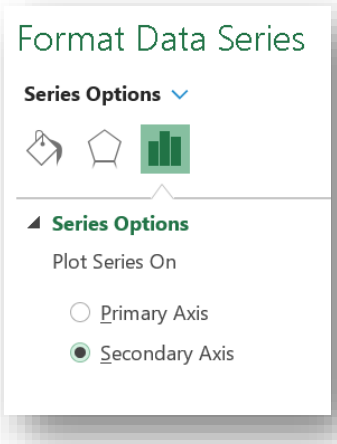


2. Your graph should look like this.

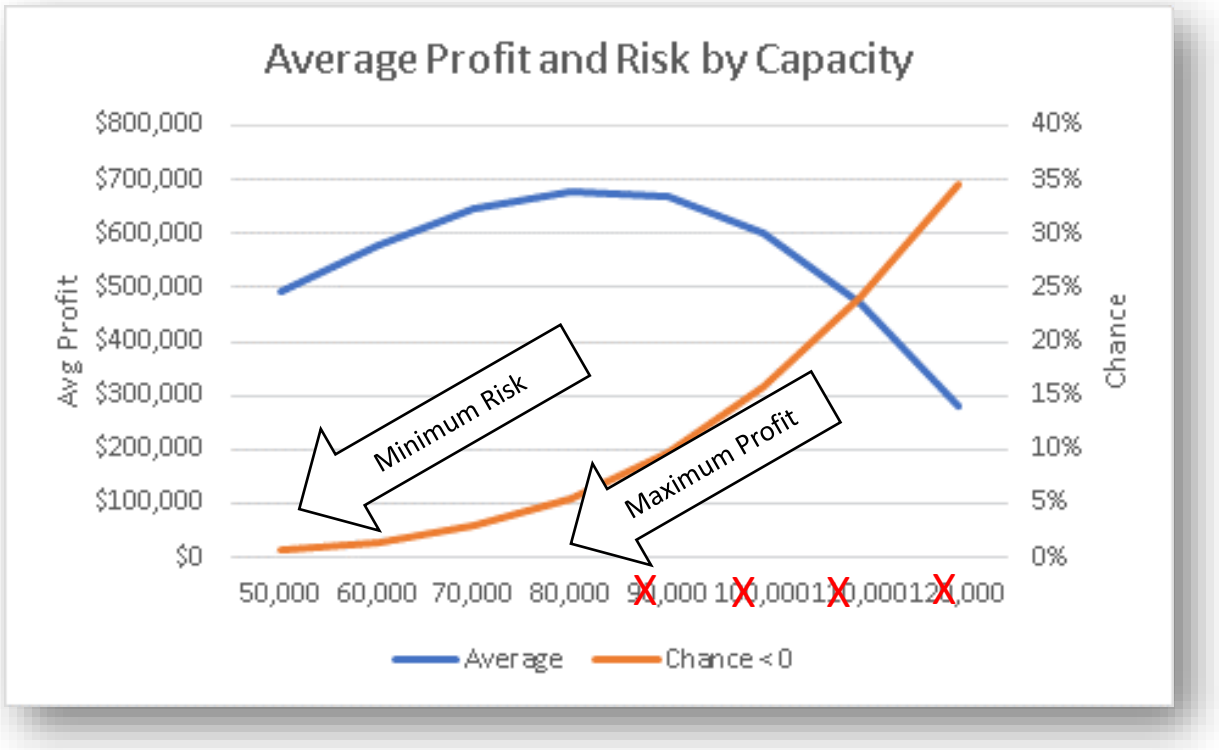


3. Because the chance of loss is so much smaller than the average profit, we will put them on separate vertical axes.
    a. Select the “Chance” line on the graph then Right Click.
    b. Select Format Data Series.
    c. Select Secondary Axis from the Format pane.





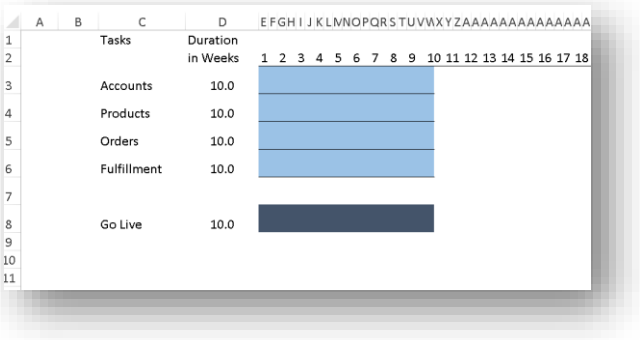
d. Add a title and axis Labels and your graph should now look like this:



We see that a capacity of 80,000 maximizes average profit. Anything above that, including the 100,000 we started with, makes no sense because it reduces average profit while increasing risk. For capacities below 80,000 there is a rational tradeoff between profit and risk, so again the choice will be guided by the risk attitude of the decision maker.

Lesson 3 – Using a 3.0 Library

- A Scheduling Model
- Virtual SIPs
- Why so many projects are behind schedule
- Chance Informed Schedule



A Scheduling Model

Open **Behind Schedule Example.xlsx** as shown on the right. This models the software development of the four major components of an online business: Customer Accounts, Product List, Orders, and Fulfillment. Each of these is being developed in parallel by separate teams, and from past data, it is known that each such development takes ten weeks on average. Of course, you don't go live until the last one is done. So, when the boss comes in and demands to know when the project will be done, what do you say?

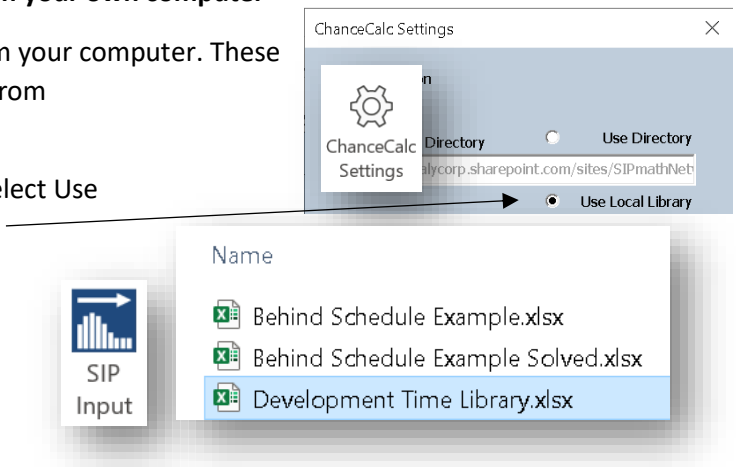
Virtual SIPs

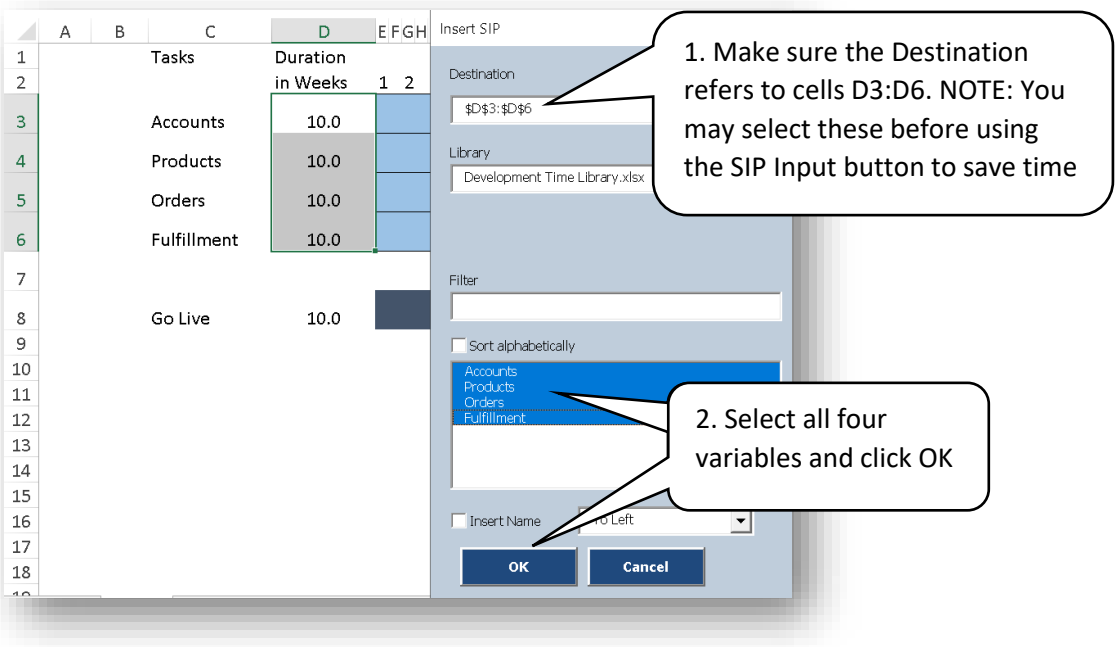
The SIPs in the earlier examples were analogous to collections of previously collected rolls of dice and are stored in the 2.0 PM SIPmath Standard[ii]. Virtual SIPs are more like the dice themselves. You can roll them as many times as you like, and they take up very little storage[iii]. An important difference from real dice is that they are controlled by the HDR random number generator from Doug Hubbard, which can produce repeatable, identical rolls across multiple platforms such as R, Python, or Excel. They are also based on the Metalog distribution of Tom Keelin, a new flexible way to quantify uncertainties with a single family of formulas.

Load the Development Time Library from your own computer

Often you may wish to load a library from your computer. These may be created with the SIPmath Tools from ProbabilityManagement.org.

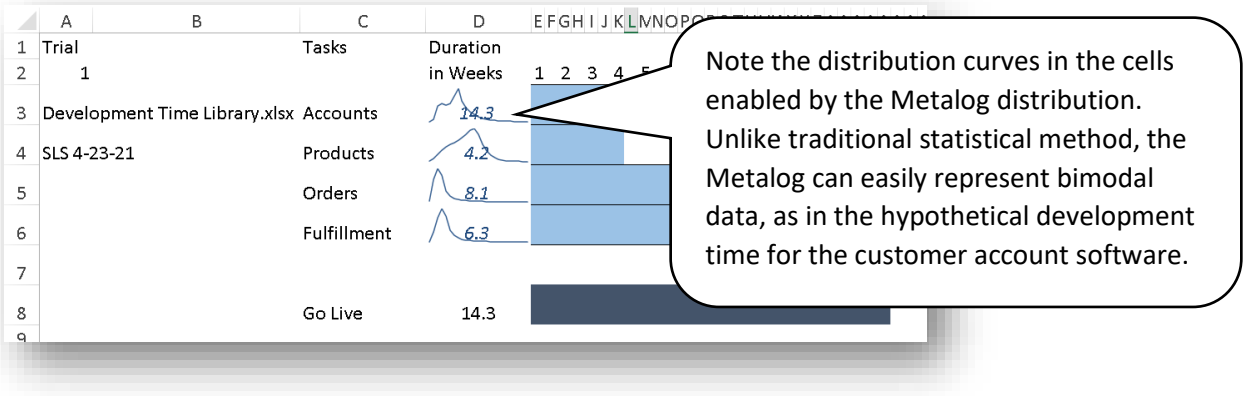
1. From the ChanceCalc settings, select Use Local Library
2. Click the SIP Input button, locate the Development Time Library, and click OK



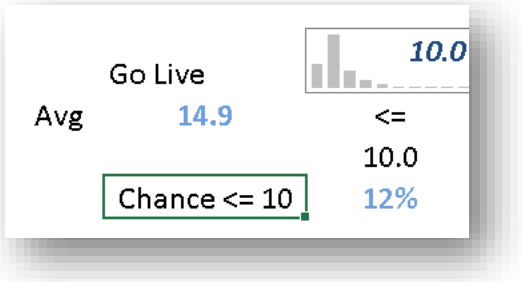

Tasks
Duration in Weeks
Accounts 10.0
Products 10.0
Orders 10.0
Fulfillment 10.0
Go Live 10.0
Insert SIP
Destination
$D$3:$D$6
Library
Development Time Library.xlsx
Filter
Sort alphabetically
Accounts
Products
Orders
Fulfillment
Insert Name
OK
Cancel
1. Make sure the Destination refers to cells D3:D6. NOTE: You may select these before using the SIP Input button to save time
2. Select all four variables and click OK

Explore the model

1. Run through a few trials to gauge the uncertainty.
2. Put the cursor in cell A1 and select the Average durations. What do you suppose the chance is of finishing in 10 or fewer weeks? When will we go live on average over 1,000 parallel universes? We will answer these questions next.



Chance-Informed Schedule



The steps are very similar to the ones in Lessons 2a and 2b, so we will merely show the results here for you to check your work against. You may be shocked that the chance of finishing in 10 weeks is so low, but here is the intuition. If the four distributions of development time had been independent and symmetric, then each task would have had a 50/50 chance of finishing over or under 10 weeks. The only way to finish in 10 weeks or less would be like flipping four heads in a row on a coin, for which the chance is $1/16^{th}$ or about 6%. So, the surprise, if any, is that the chance is as high as 12%. Also note that we finish in 10 weeks given the average inputs, but that the simulated average is about 15 weeks. Given that many contracts involve bonuses and penalties based on finish date, consistently having projects exceed your estimates by 50% does not lead to a sustainable business.

Lesson 4 – Failure Analysis, Multiple Identical Variables

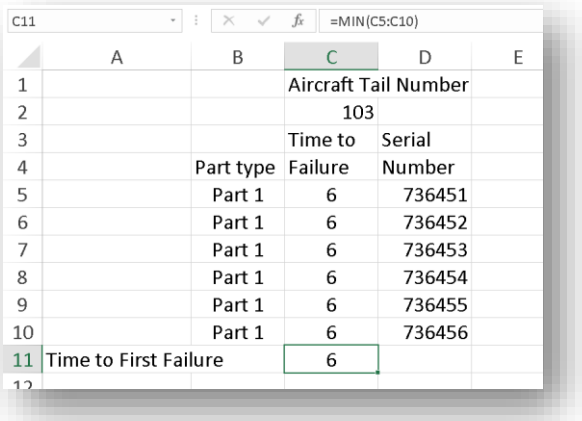
- Predicting the Failure of Aircraft Parts
- Linking Simulations to Serial Numbers

Predicting the Failure of Aircraft Parts

Consider an aircraft part, which on average fails and needs replacement every six flight hours. That is, it has a Mean Flight Hours Before Failure (MFHBF) of six. Now suppose there are six such parts on the plane, and if any one of them fails, the plane must be grounded and repaired. The expected time for this part to ground the plane is the minimum of six independent failure times. Replacing these with the mean (another word for average), we find that the minimum of 6,6,6,6,6 and 6 is 6 flight hours. But hopefully by now you're not going to fall for the Flaw of Averages, and you should be saying let's simulate the minimum of 6 failures.

Suppose we had a SIP of failure time for the part. That wouldn't work either. If all six parts refer to the same SIP, then they would all fail at the same instant, and although there would now be some variability, the average time to first failure would still be 6 flight hours. In the next example, we will use a Virtual SIP whose HDR seeds are tied to the serial numbers of the parts and tail number of the plane to estimate the chance that the time to the failure of the first part is at least 6 flight hours, and also what it is on average.

Open **First Part Failure.xlsx** as shown. Note that we are only modeling the time it would take for Part 1 to ground the plane. Clearly, such models could be generalized to include much larger numbers of both identical and different parts.



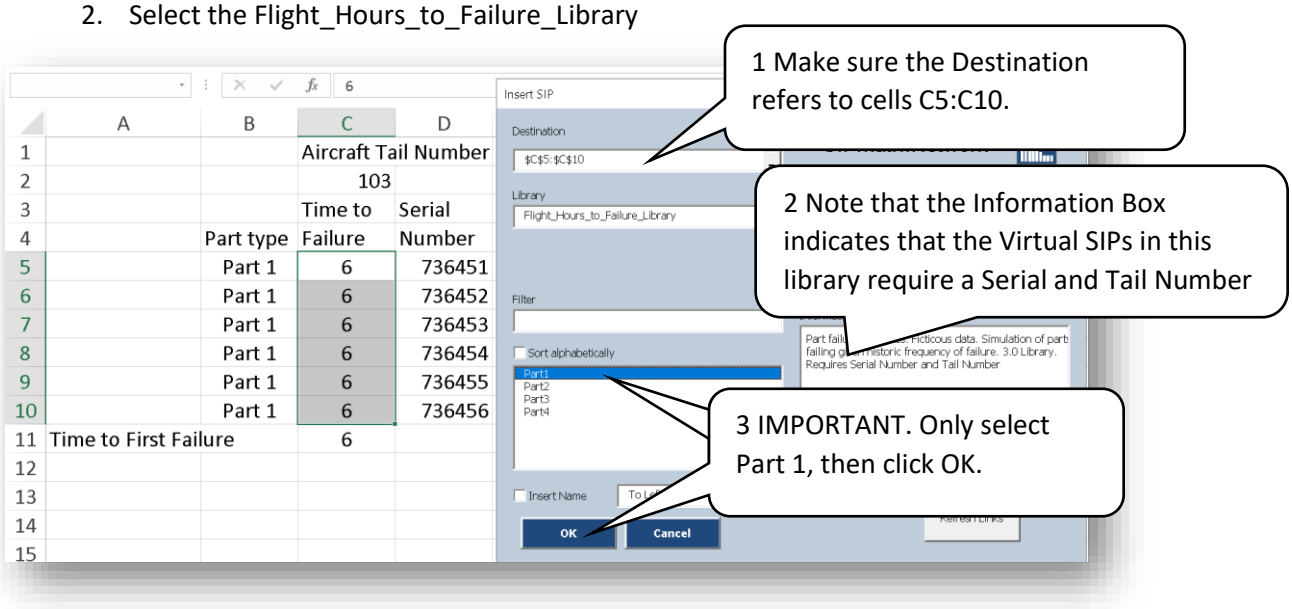
	A	B	C	D	E
1			Aircraft Tail Number		
2			103		
3			Time to	Serial	
4		Part type	Failure	Number	
5		Part 1	6	736451	
6		Part 1	6	736452	
7		Part 1	6	736453	
8		Part 1	6	736454	
9		Part 1	6	736455	
10		Part 1	6	736456	
11	Time to First Failure		6		

Linking Simulations to Serial Numbers

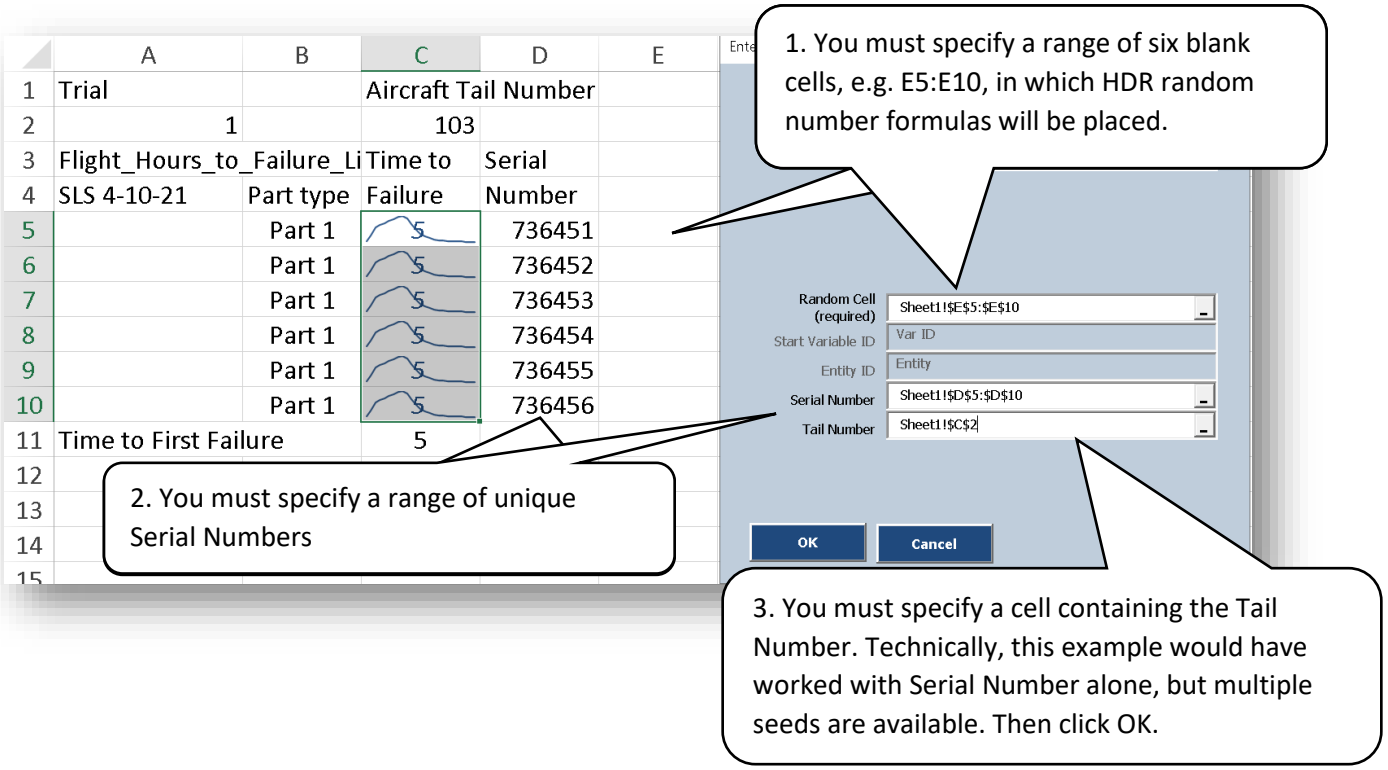
Load Flight Hours Library

1. Select cells C5:C10, then click the SIP Input button.

2. Select the Flight_Hours_to_Failure_Library



3. Fill out HDR seed fields



Chance-Informed Model

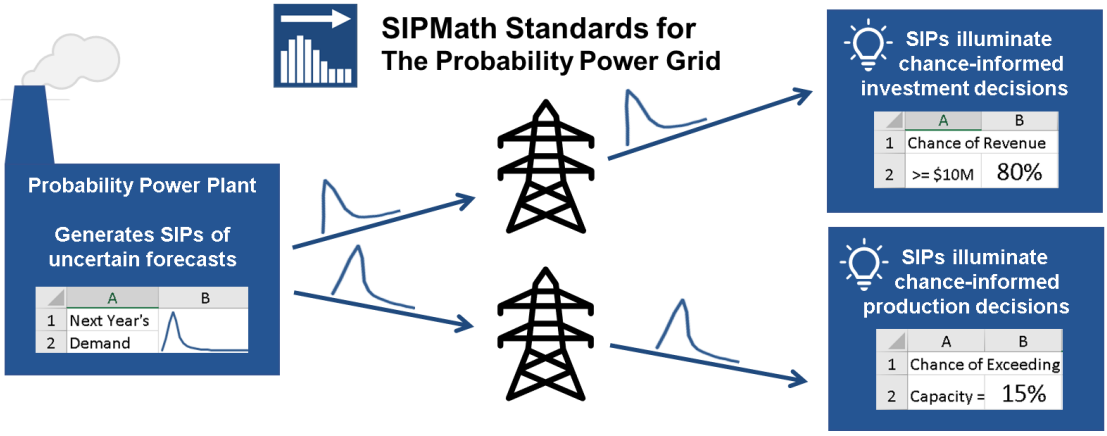
11	Time to First Failure	6	>=	6	0.8%
12	Average	3.143			

Using the steps demonstrated earlier, it is easy to demonstrate that there is virtually no chance of achieving 6 hours of flight time before failure of at least one of Part 1, and that on average the first failure will occur in just over three flight hours.

Creating SIP Libraries

ChanceCalc is the first application of its kind. Why? Because before the discipline of probability management, which established a standard for communicating uncertainty as actionable data, **Chance**Calc would have been as useless as a lightbulb before electricity. And the 2.0 and recent 3.0 PM SIPmath Standards play roles analogous to the DC and AC current standards that ushered in the electric age.

We refer to the process of replacing calculations using numbers with calculations using probabilities as chancification, and it has some parallels with electrification.



Beyond the PM SIPmath Standards, there are two other critical components:

Doug Hubbard's cross-platform HDR random number generator, which is analogous to the alternating electric current generators.

and

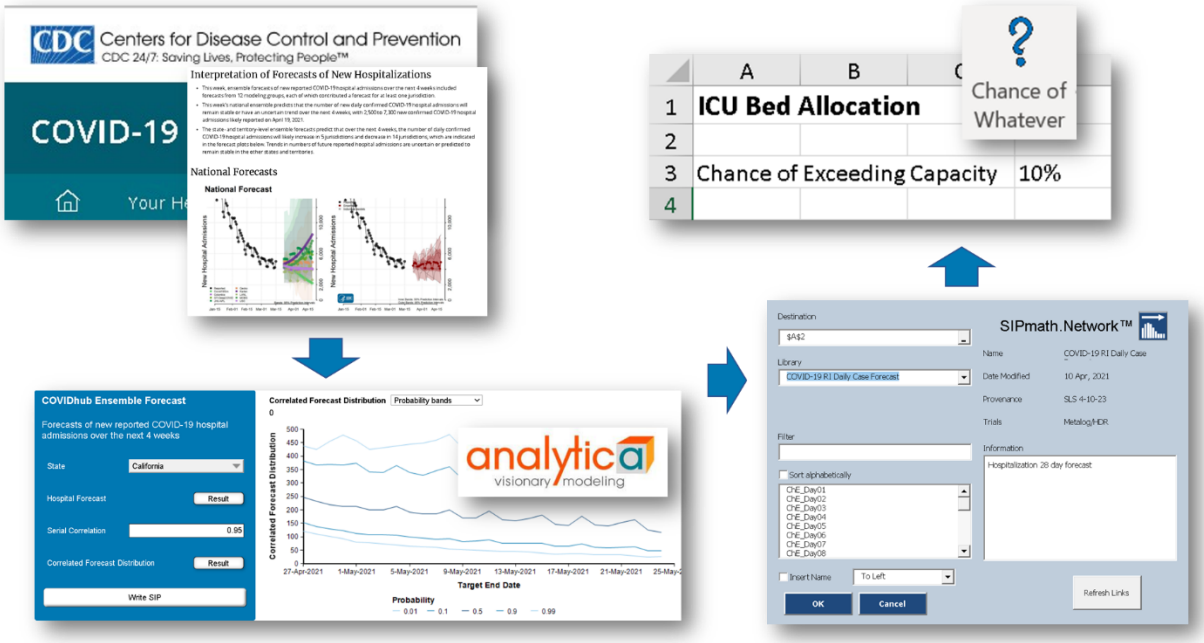
Tom Keelin's Metalog distributions that play the role of transformers, that convert current from the generators into numerous useful forms.

Just as electrification enables a small number of power engineers to supply electricity to millions of people for illuminating their homes, chancification allows a small number of data scientists and statisticians to supply SIP Libraries to a wide audience of managers for illuminating chance-informed decisions.

Examples

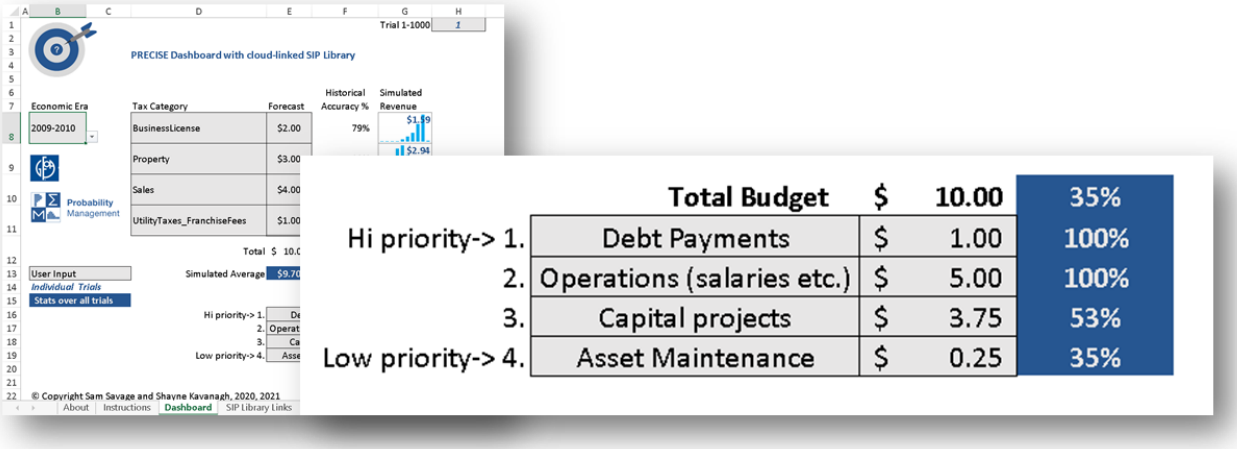
CDC COVID-19 Hospitalizations

Eng-wee Yeo, Principal, Technology Risk Modeling & Methodology at Kaiser Permanente, was concerned about the uncertainty in COVID-19 hospitalization rates. He created an application in Analytica, a powerful modeling system from Lumina Decision Systems, to read data directly from the CDC's web site and create SIP libraries of 28 day hospitalization forecasts, an example of which appears in the SIPmath.Network browser in **Chance**Calc. Such libraries can assist healthcare facilities make chance-informed decisions surrounding ICU bed allocation, the scheduling of elective surgeries and numerous other impacted areas of hospital management.



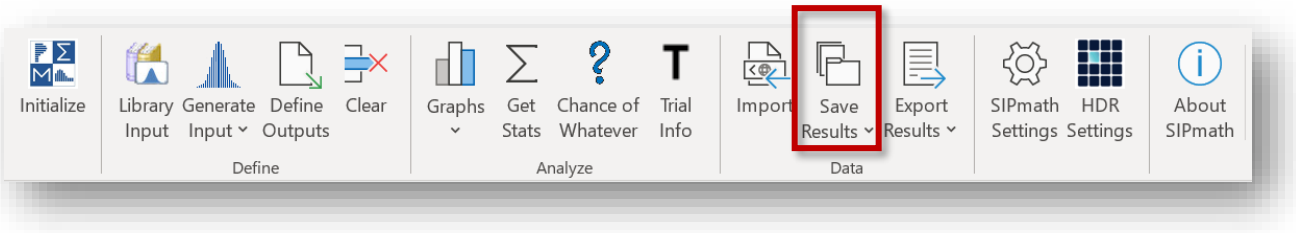
The PRECISE Uncertainty Project

The Government Finance Officers Association (GFOA), a professional organization of over 20,000 municipal CFOs and other financial managers, has assessed the accuracy of tax revenue forecasts by some of its members. ProbabilityManagement.org has created SIP Libraries of historical accuracy using R to assist CFOs in making chance-informed decisions around prioritized budgeting.

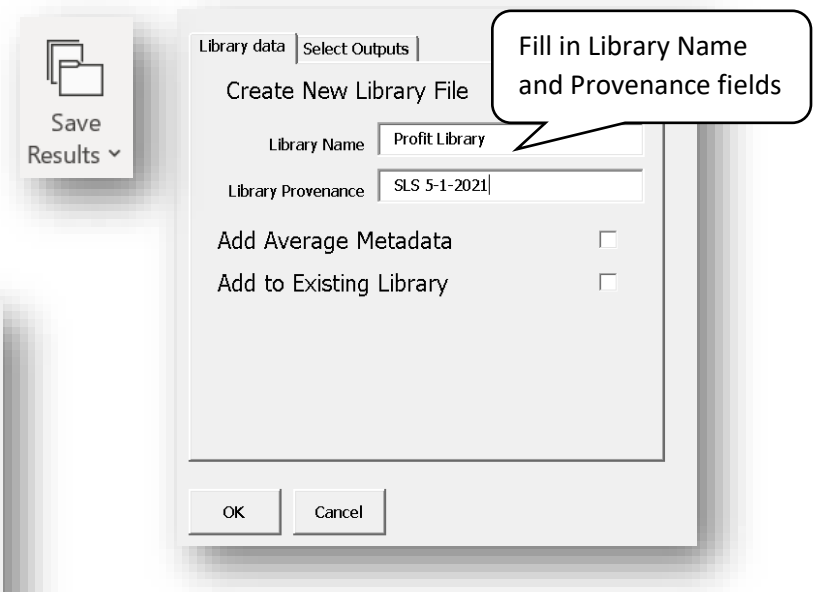


Lesson 5 – Creating a SIP Library of Profit from Lesson-2a

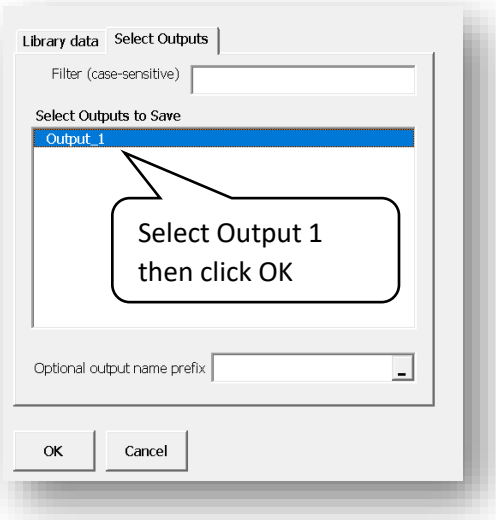
Although we will describe several packages that can create SIP Libraries below, here we will describe the process with either the free or Enterprise version of the PM SIPmath Modeler Tools, which are available for download from ProbabilityManagement.org. Whereas **Chance**Calc can only read SIPs from libraries, the Modeler Tools can generate random inputs on their own, and read from and save 2.0 SIP Libraries. The tools are more complex to learn than **Chance**Calc, but for this lesson you will only need to use the Save Results button.



1. Open the saved version of the Below Projection model saved in Lesson 2a and click the Save Results button.

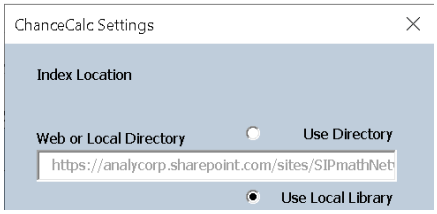


2. Click the Select Outputs tab

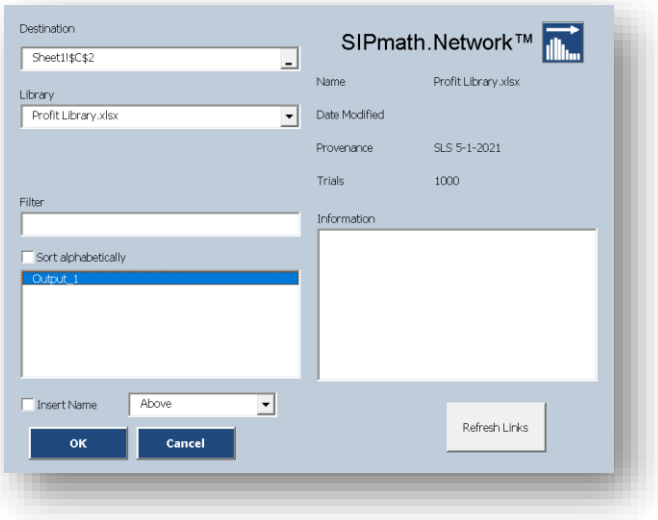


Then save the library in a folder that you can remember.

3. To use the Library in another model. Select Use Local Library from ChanceCalc settings.



4. You may now use Output 1 in a new model just as we have used the other SIPs.



Tools for Generating SIP Libraries

Below is a discussion of various packages that can generate SIP Libraries for **Chance**Calc, starting with those from ProbabilityManagement.org.

Package	**Organization**
SIPmath Modeler Tools for Excel	ProbabilityManagement.org

The SIPmath Modeler Tools create Monte Carlo simulations using Excel's Data Table. This means that:

1. Unlike other simulation packages, you can run the models in native Excel without having the tools installed.
2. The models are interactive, running thousands of trials per keystroke.

There is both a free and Enterprise version, and both can generate their own random numbers or read and write 2.0 SIPmath Libraries, which can be read by **Chance**Calc. They are a logical upgrade from **Chance**Calc for those who want to graduate beyond using SIP Libraries generated by others.

Package	Organization
RSIPlibrary	ProbabilityManagement.org

This is an open-source R package written by Aaron Brown, which takes Monte Carlo trials from any program, stored as columns in CSV format, and creates a 3.0 Virtual SIP Library, which may be read by **Chance**Calc.

Package	Organization
PySIPlibrary	ProbabilityManagement.org

This is a Python version of the R package above.

Package	Organization
Numerous analytical Excel Add-ins	Frontline Systems Solver.com

Frontline Systems is the company that created the Solver in Excel, and it offers a wide range of analytical tools that include simulation and optimization. Its packages were the first to recognize the SIP and SLURP data type. They have powerful solvers for optimizing models with uncertain inputs.

Package	Organization
	Lumina.com

Analytica is a powerful multi-dimensional modeling language, which can deliver a complete probability management solution, including simulation and optimization. It has recently been used to generate 2.0 SIP Libraries from hospitalization forecasts from the CDC.

Package	Organization
Autobox	Automatic Forecasting Systems Autobox.com

Autobox performs sophisticated time series analysis. Virtually all forecasting systems track estimates of their own accuracy. Autobox can output these automatically as a 2.0 SIP Library.

Package	Organization
	Oracle Corp. oracle.com/applications/crystalball/

Crystal Ball has been a widely used Monte Carlo simulation package since the 1980s. It has a large userbase and was acquired by Oracle in 2009. The Enterprise PM SIPmath Modeler Tools from ProbabilityManagement.org have a function to generate 2.0 SIP Libraries from Crystal Ball Models.

Package	Organization
	Palisade Corp. palisade.com

@RISK is another widely used Monte Carlo simulation package dating back to the 1980s. It has a large userbase and was developed Palisade Corp., which has other analytical packages beyond simulation. The PM Enterprise SIPmath Modeler Tools from ProbabilityManagement.org have a function to generate 2.0 SIP Libraries from @RISK Models.

[i] https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/hospitalizations-forecasts.html
[ii] LINK to 2.0 PM SIPmath Standard
[iii] LINK to 3.0 PM SIPmath Standard