Raytheon

Customer Success Is Our Mission

Using JMP® to Assess Risk in Financial Predictions by Using Monte Carlo Simulations

Joan Maney
Statistician
18 September 2014

Estimate at Completion (EAC)

- The Estimate At Completion (EAC) is an estimation of the total costs required to fulfill contract requirements at the completion of the contract. An EAC is calculated by taking cumulative to-date actual costs and adding the forecasted costs of remaining, authorized work.
- The goals of an EAC include the following:
  - Early identification of contract cost-performance issues
  - Identification and assessment of contract risks and opportunities
  - Documentation of contract risk mitigation and opportunity realization plans
  - Timely consideration of contract risks and opportunities in booking rates for revenue recognition

Estimate at Completion (EAC)

EAC

=

Actuals(cum to date)

+

ETC(Estimate To Completion)

Accurate EAC relies on ETC estimation

Variance at Completion (VAC)

- BAC (Budget at Completion) is the budget for the program at program completion
- VAC (Variance at Completion) = BAC - EAC
  - If VAC > 0 $\Rightarrow$ possible underrun (Opportunity)
  - If VAC < 0 $\Rightarrow$ possible overrun (Risk)

Example

- A major program has a product that has four major subassemblies that have to be manufactured and get combined to create the final product.
- For this example, we have a hypothetical product “LOGO” that is made up of the 4 subassemblies (‘j’, ‘m’, ‘p’ and the dot over the ‘j’). We will use a very simplified (unrealistic) example where EAC is only based on manufacturing hours.

Calculation of ETC

- To date we have completed 50 LOGO units. There are 30 more to complete by the end of the contract.
- The Actuals(cum to date) for the 50 completed LOGO units is $393,764.
- Raw data for each of the four subassemblies of the 50 completed units is collected. The average hours for each subassembly are:

j	m	p	dot
108.58 hours	12.61 hours	35.83 hours	0.48 hours

- Therefore, the average hours to build a LOGO unit is 157.50 hours (sum of four subassemblies averages).
- Assuming the labor rate is $50/hour the cost to complete the remaining 30 units can be estimated by:
  ETC=157.50 hours/unit * $50/hour *30 units = $236,258.
- EAC = Actuals(cum to date) + ETC
  = $393,764+$236,258 = $630,022

Calculation of ETC

- The Budget at Completion (BAC) is $600K.
- Variance at Completion:
  VAC = BAC – EAC
  = $600,000 - $630,022 = ($30,022)

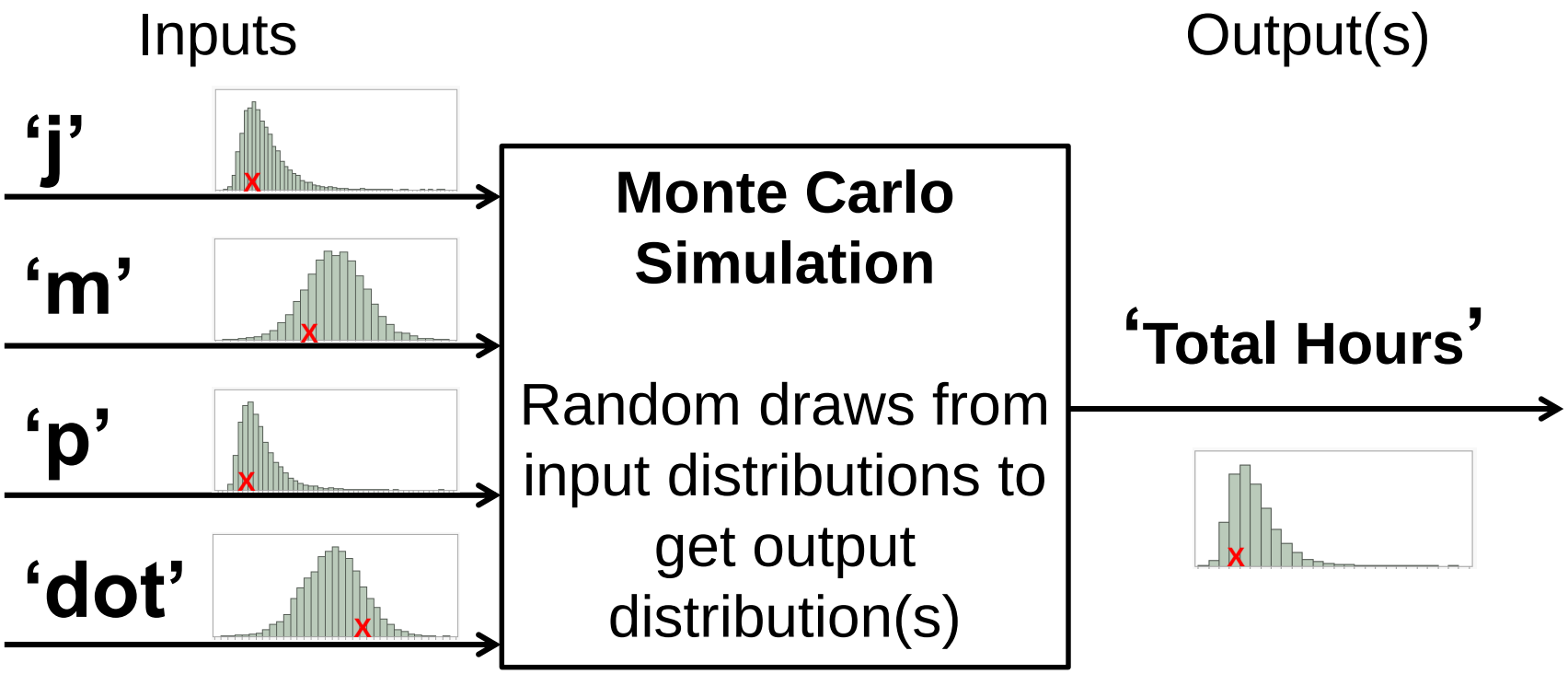
LOGO program has potential overrun of $30,022

Risk in ETC Estimate

- Basing ETC on Avg hours/unit does not include risk since it does not include any variability in the hours on each subassembly.
- Avg hours/unit is could be overestimating or underestimating ETC.
- To account for variability, a Monte Carlo Simulation can be used.
  - Determine probability of overrunning/underrunning
  - Determine range of overrun or underrun

Monte Carlo Simulation

- Monte Carlo Simulation is a technique that takes into account the uncertainty/variability in the inputs to make an estimate of the output(s). Instead of taking on a single value, the output will take on a range of values.

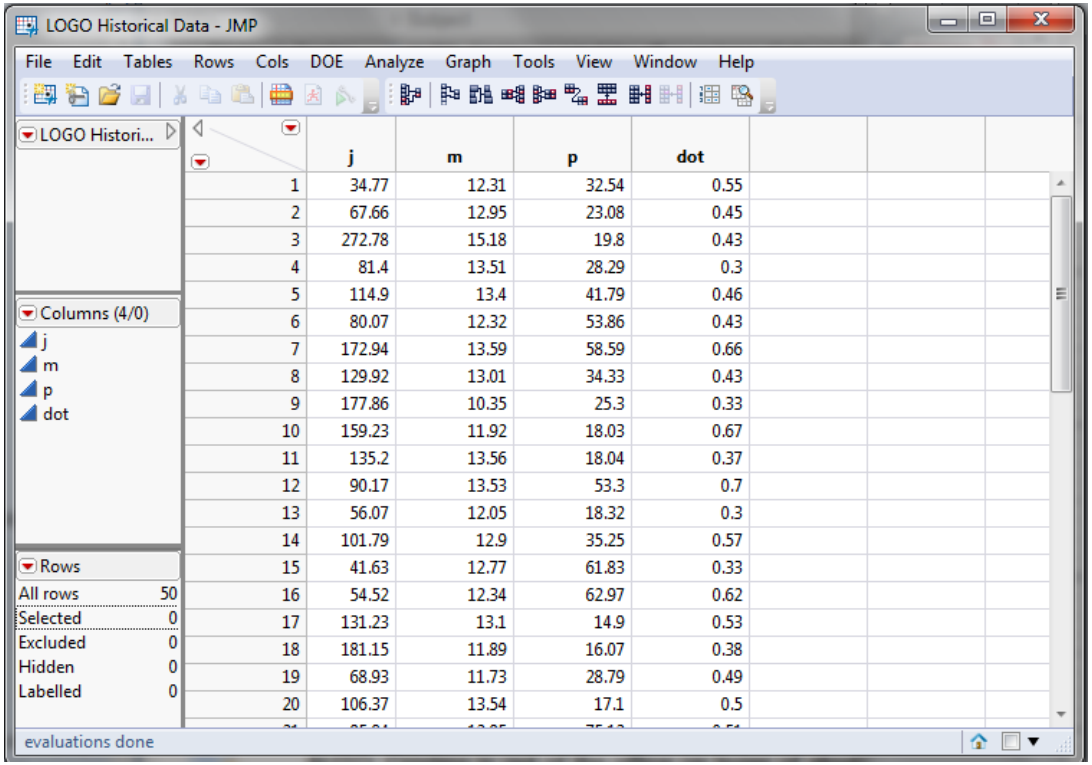


JMP Demo

- The following pages are steps to complete a Monte Carlo simulation in JMP.

Monte Carlo Steps

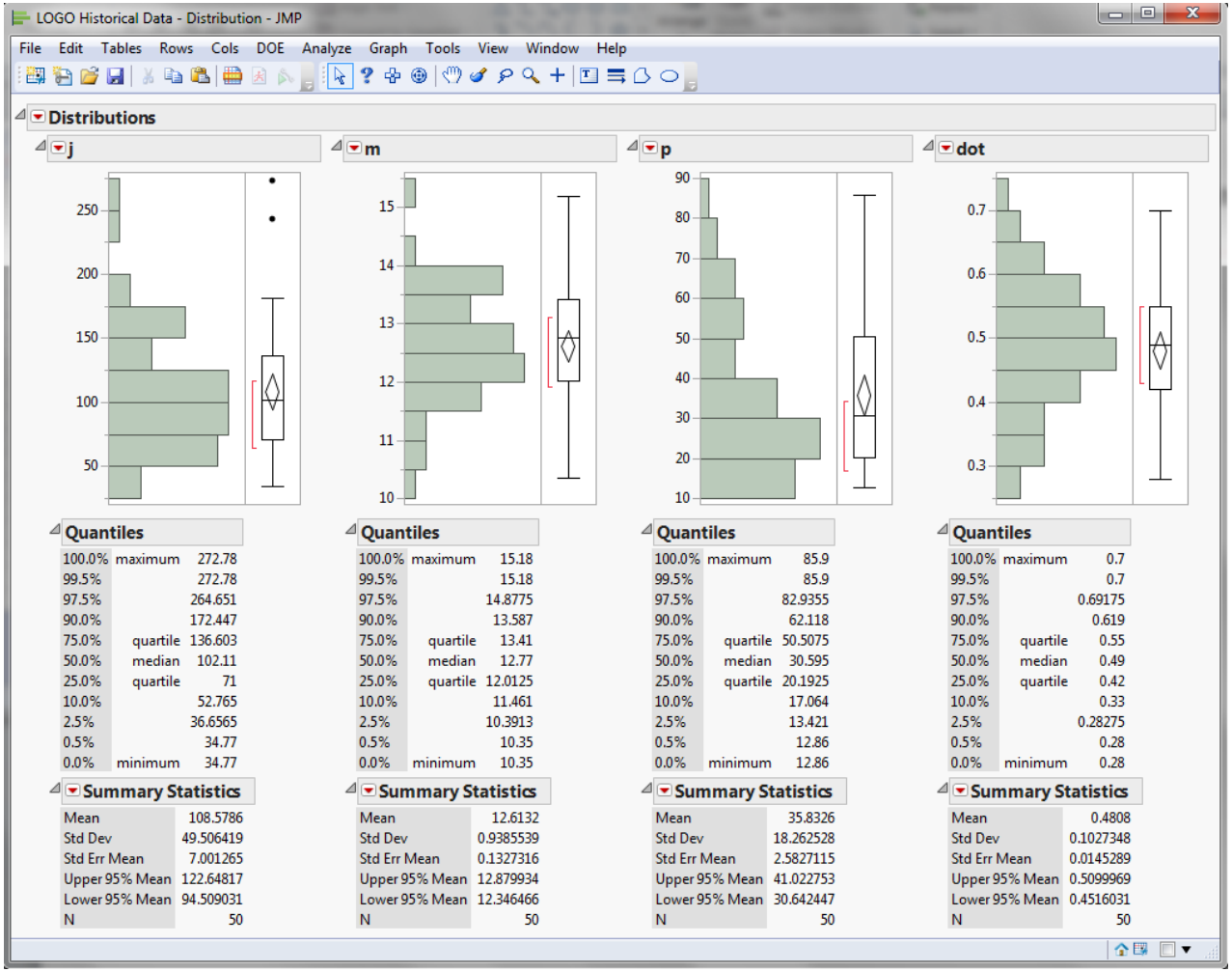
1. Collect historical data on inputs. 50 of each of the four subassemblies was made and the hours to complete each subassembly was collected.



	j	m	p	dot
1	34.77	12.31	32.54	0.55
2	67.66	12.95	23.08	0.45
3	272.78	15.18	19.8	0.43
4	81.4	13.51	28.29	0.3
5	114.9	13.4	41.79	0.46
6	80.07	12.32	53.86	0.43
7	172.94	13.59	58.59	0.66
8	129.92	13.01	34.33	0.43
9	177.86	10.35	25.3	0.33
10	159.23	11.92	18.03	0.67
11	135.2	13.56	18.04	0.37
12	90.17	13.53	53.3	0.7
13	56.07	12.05	18.32	0.3
14	101.79	12.9	35.25	0.57
15	41.63	12.77	61.83	0.33
16	54.52	12.34	62.97	0.62
17	131.23	13.1	14.9	0.53
18	181.15	11.89	16.07	0.38
19	68.93	11.73	28.79	0.49
20	106.37	13.54	17.1	0.5

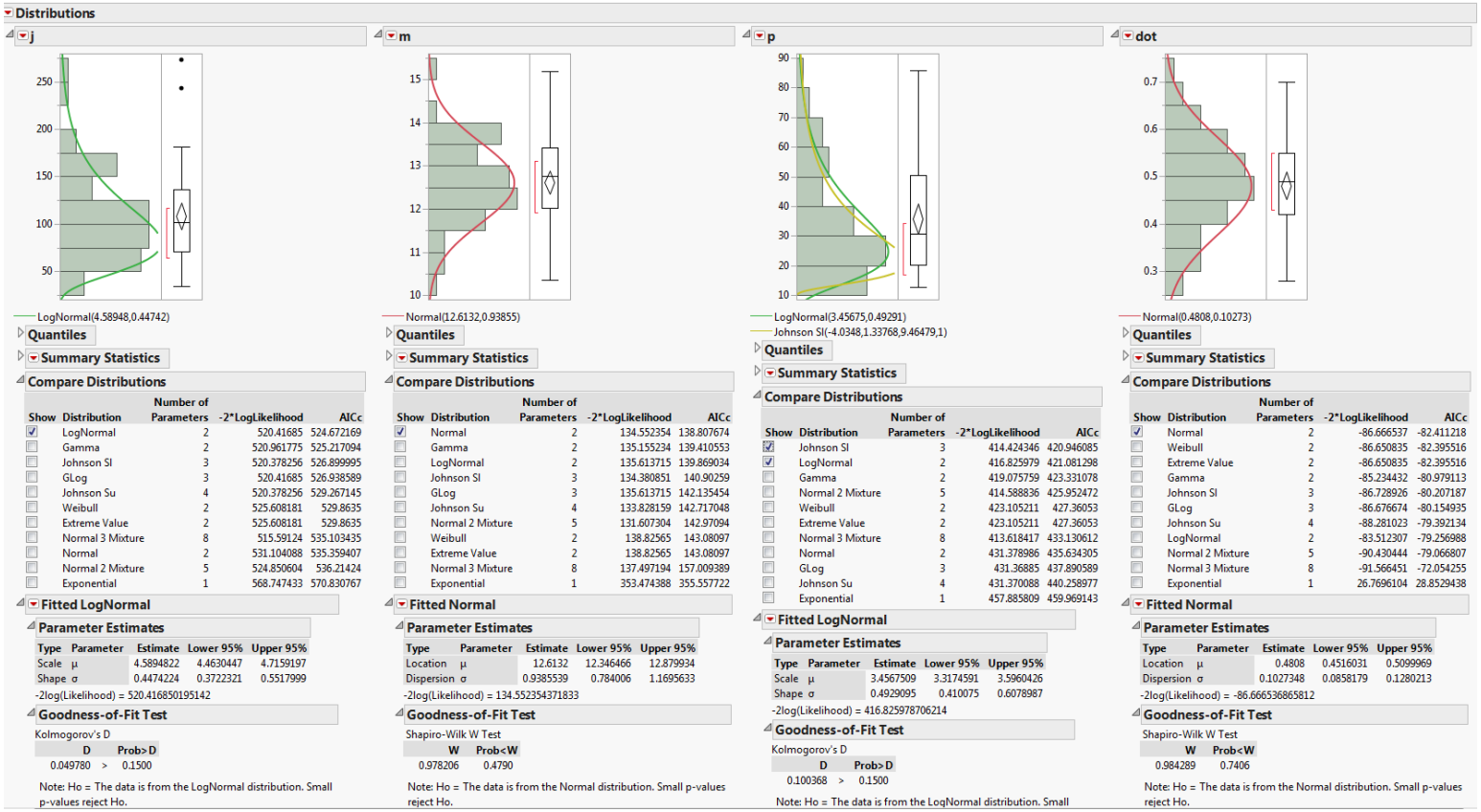
Monte Carlo Steps (cont.)

2. Fit distributions to historical subassembly data from four subassemblies (Analyze, Distribution)



Monte Carlo Steps (cont.)

- Add distribution fit to histograms (red down arrow, Continuous Fit, All)



Distributions

j

LogNormal(4.58948,0.44742)

Quantiles

Summary Statistics

Compare Distributions

Show	Distribution	Number of Parameters	-2*LogLikelihood	AICc
☑	LogNormal	2	520.41685	524.672169
☐	Gamma	2	520.961775	525.217094
☐	Johnson SI	3	520.378256	526.899995
☐	GLog	3	520.41685	526.938589
☐	Johnson Su	4	520.378256	529.267145
☐	Weibull	2	525.608181	529.8635
☐	Extreme Value	2	525.608181	529.8635
☐	Normal 3 Mixture	8	515.59124	535.103435
☐	Normal	2	531.104088	535.359407
☐	Normal 2 Mixture	5	524.850604	536.21424
☐	Exponential	1	568.747433	570.830767

Fitted LogNormal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Scale	μ	4.5894822	4.4630447	4.7159197
Shape	σ	0.4474224	0.3722321	0.5517999

-2log(Likelihood) = 520.416850195142

Goodness-of-Fit Test

Kolmogorov's D

D		Prob>D
0.049780	>	0.1500

Note: Ho = The data is from the LogNormal distribution. Small p-values reject Ho.

m

Normal(12.6132,0.93855)

Quantiles

Summary Statistics

Compare Distributions

Show	Distribution	Number of Parameters	-2*LogLikelihood	AICc
☑	Normal	2	134.552354	138.807674
☐	Gamma	2	135.155234	139.410553
☐	LogNormal	2	135.613715	139.869034
☐	Johnson SI	3	134.380851	140.90259
☐	GLog	3	135.613715	142.135454
☐	Johnson Su	4	133.828159	142.717048
☐	Normal 2 Mixture	5	131.607304	142.97094
☐	Weibull	2	138.82565	143.08097
☐	Extreme Value	2	138.82565	143.08097
☐	Normal 3 Mixture	8	137.497194	157.009389
☐	Exponential	1	353.474388	355.557722

Fitted Normal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	μ	12.6132	12.346466	12.879934
Dispersion	σ	0.9385539	0.784006	1.1695633

-2log(Likelihood) = 134.552354371833

Goodness-of-Fit Test

Shapiro-Wilk W Test

W	Prob<W
0.978206	0.4790

Note: Ho = The data is from the Normal distribution. Small p-values reject Ho.

p

LogNormal(3.45675,0.49291)

Johnson SI(-4.0348,1.33768,9.46479,1)

Quantiles

Summary Statistics

Compare Distributions

Show	Distribution	Number of Parameters	-2*LogLikelihood	AICc
☑	Johnson SI	3	414.424346	420.946085
☑	LogNormal	2	416.825979	421.081298
☐	Gamma	2	419.075759	423.331078
☐	Normal 2 Mixture	5	414.588836	425.952472
☐	Weibull	2	423.105211	427.36053
☐	Extreme Value	2	423.105211	427.36053
☐	Normal 3 Mixture	8	413.618417	433.130612
☐	Normal	2	431.378986	435.634305
☐	GLog	3	431.36885	437.890589
☐	Johnson Su	4	431.370088	440.258977
☐	Exponential	1	457.885809	459.969143

Fitted LogNormal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Scale	μ	3.4567509	3.3174591	3.5960426
Shape	σ	0.4929095	0.410075	0.6078987

-2log(Likelihood) = 416.825978706214

Goodness-of-Fit Test

Kolmogorov's D

D		Prob>D
0.100368	>	0.1500

Note: Ho = The data is from the LogNormal distribution. Small

dot

Normal(0.4808,0.10273)

Quantiles

Summary Statistics

Compare Distributions

Show	Distribution	Number of Parameters	-2*LogLikelihood	AICc
☑	Normal	2	-86.666537	-82.411218
☐	Weibull	2	-86.650835	-82.395516
☐	Extreme Value	2	-86.650835	-82.395516
☐	Gamma	2	-85.234432	-80.979113
☐	Johnson SI	3	-86.728926	-80.207187
☐	GLog	3	-86.676674	-80.154935
☐	Johnson Su	4	-88.281023	-79.392134
☐	LogNormal	2	-83.512307	-79.256988
☐	Normal 2 Mixture	5	-90.430444	-79.066807
☐	Normal 3 Mixture	8	-91.566451	-72.054255
☐	Exponential	1	26.7696104	28.8529438

Fitted Normal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	μ	0.4808	0.4516031	0.5099969
Dispersion	σ	0.1027348	0.0858179	0.1280213

-2log(Likelihood) = -86.666536865812

Goodness-of-Fit Test

Shapiro-Wilk W Test

W	Prob<W
0.984289	0.7406

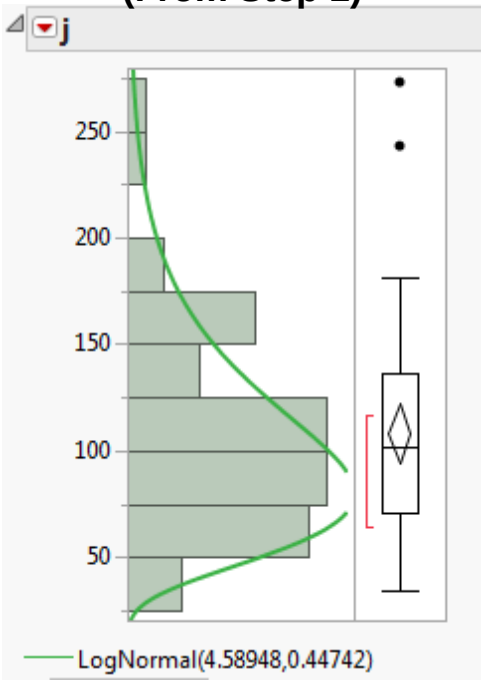
Note: Ho = The data is from the Normal distribution. Small p-values reject Ho.

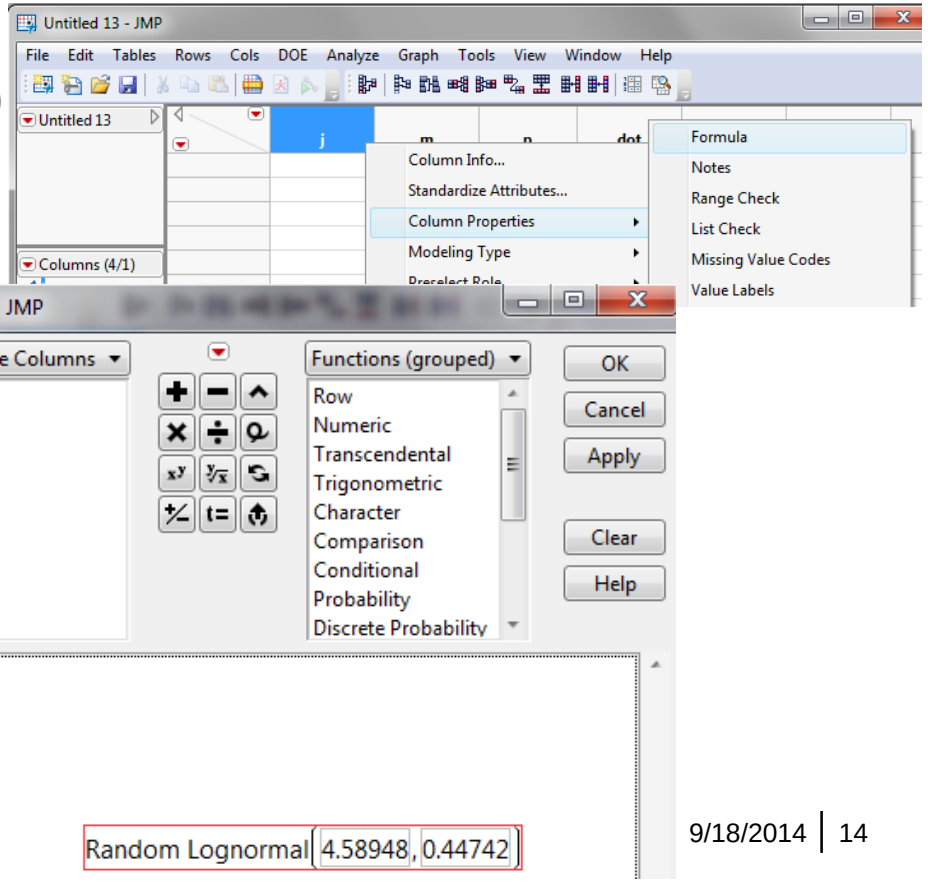
- Smallest AICc is gives the “Best Fit” distribution.
- Use Goodness-of-Fit Test to test the quality of the “Best Fit” distribution. (p-values < 0.05 => distribution does not fit the data)
- Use simplest distribution from those that fit. For example for subassembly ‘p’ the Johnson SI has the lowest AICc but it is very close to the LogNormal. Since the LogNormal is a simpler distribution use it.

Monte Carlo Steps (cont.)

3. Use distributions found in Step 2 to create new data set of with random draws from the distributions for the four subassemblies
   - Create columns (j,m,p,dot) with formulas based on distributions from Step 2
     - Random Normal(mu, sigma)
     - Random Lognormal(mu, sigma)

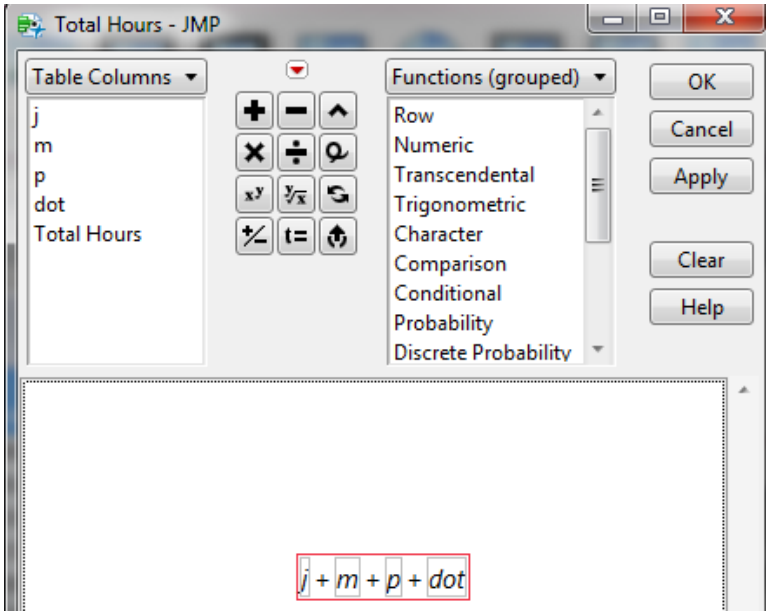
(From Step 1)



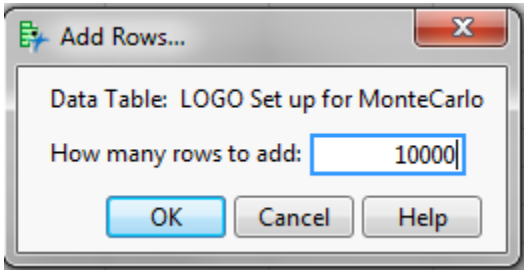


Monte Carlo Steps (cont.)

4. Create column “Total Hours”= j+m+p+dot

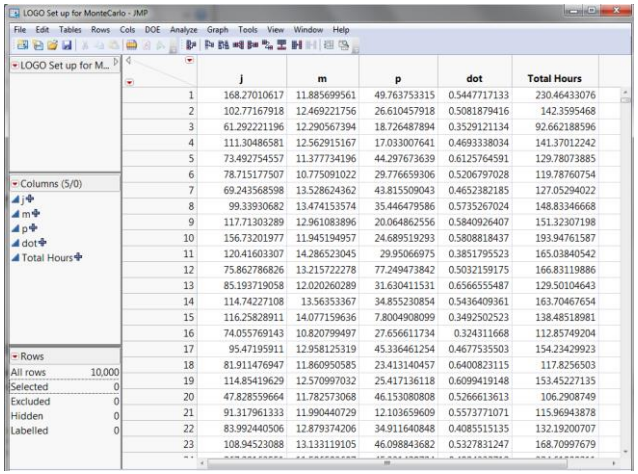


5. Add 10,000 rows (Rows, Add Rows)



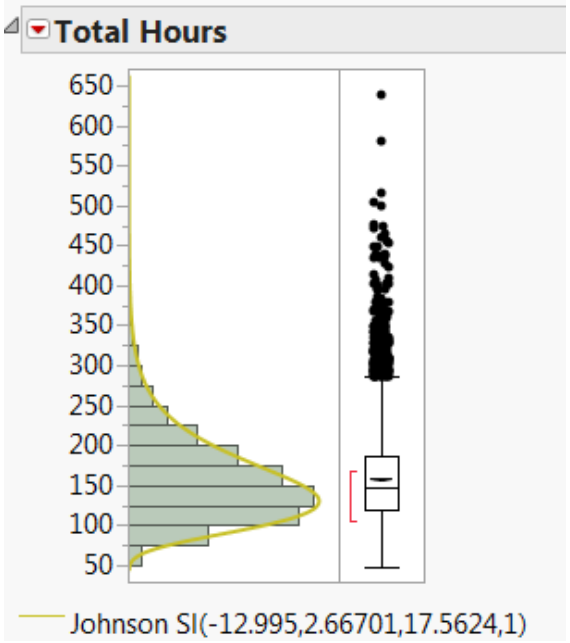
Monte Carlo Steps (cont.)

6. Analyze “Total Hours” column



	j	m	p	dot	Total Hours
1	168.27010617	11.885699561	49.763753315	0.5447717133	230.46433076
2	102.77167918	12.469221756	26.610457918	0.5081879416	142.3595468
3	61.292221196	12.290567394	18.726487894	0.3529121134	92.662188596
4	111.30486581	12.562915167	17.033007641	0.4693338034	141.37012242
5	73.492754557	11.377734196	44.297673639	0.6125764591	129.78073885
6	78.715177507	10.775091022	29.776659306	0.5206797028	119.78760754
7	69.243568598	13.528624362	43.815509043	0.4652382185	127.05294022
8	99.33930682	13.474153574	35.446479586	0.5735267024	148.83346668
9	117.71303289	12.961083896	20.064862556	0.5840926407	151.32307198
10	156.73201977	11.945194957	24.689519293	0.5808818437	193.94761587
11	120.41603307	14.286523045	29.95066975	0.3851795523	165.03840542
12	75.862786826	13.215722278	77.249473842	0.5032159175	166.83119886
13	85.193719058	12.020260289	31.630411531	0.6566555487	129.50104643
14	114.74227108	13.56353367	34.855230854	0.5436409361	163.70467654
15	116.25828911	14.077159636	7.8004908099	0.3492502523	138.48518981
16	74.055769143	10.820799497	27.656611734	0.324311668	112.85749204
17	95.47195911	12.958125319	45.336461254	0.4677535503	154.23429923
18	81.911476947	11.860950585	23.413140457	0.6400823115	117.8256503
19	114.85419629	12.570997032	25.417136118	0.6099419148	153.45227135
20	47.828559664	11.782573068	46.153080808	0.5266613613	106.2908749
21	91.317961333	11.990440729	12.103659609	0.5573771071	115.96943878
22	83.992440506	12.879374206	34.911640848	0.4085515135	132.19200707
23	108.94523088	13.133119105	46.098843682	0.5327831247	168.70997679

- Create histogram and fit distribution of “Total Hours”



Quantiles

100%	maximum	639.016
90%		227.668
80%		196.334
70%		176.533
60%		161.459
50%	median	148.208
40%		136.117
30%		124.788
20%		113.277
10%		98.5849
0%	minimum	48.2964

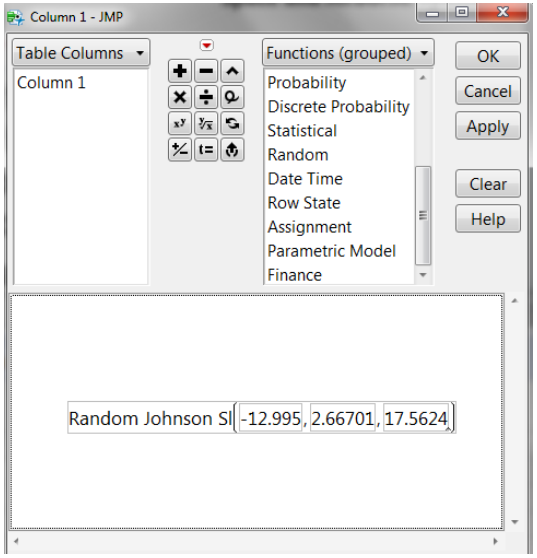
Summary Statistics

Mean	157.76466
Std Dev	54.826028
Std Err Mean	0.5482603
Upper 95% Mean	158.83937
Lower 95% Mean	156.68996
N	10000

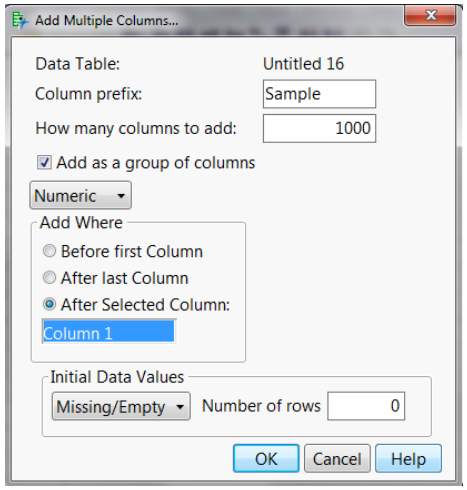
Monte Carlo Steps (cont.)

7. Remaining 30 units should follow a similar distribution in the Total Hours to build each unit. To estimate the ETC for the 30 units run another Monte Carlo simulation that takes random draws of size 30 from the Total Hours distribution.

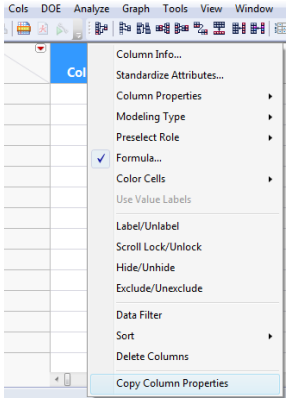
a. Create 1 column with fitted "Total Hrs" distribution from Step 6



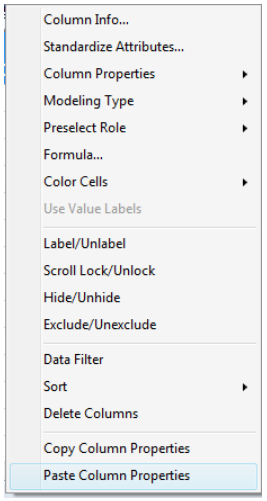
b. Add 1000 columns as a group. (Columns, Add Multiple Columns)



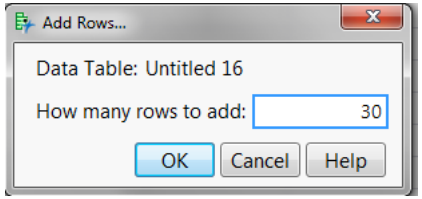
c. Copy formula from Column 1 (right mouse, Copy Column Properties)



d. Paste formula to 1000 group columns (select grouped columns, right mouse, Paste Column Properties)



e. Add 30 rows (Rows, Add Rows)

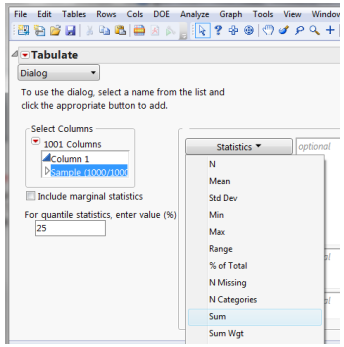


Monte Carlo Steps (cont.)

Raytheon
Space and Airborne Systems

8. To get an estimate of ETC need to sum each column to get total hours for all 30 units. Then save as a data table and transpose the data table.

a. Sum each column (Analyze, Tabulate)

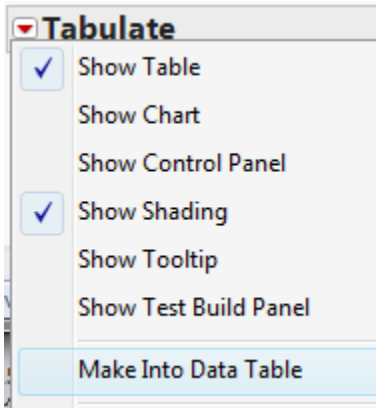


b. Results from Tabulate. Need to make into data table and transform to one column.

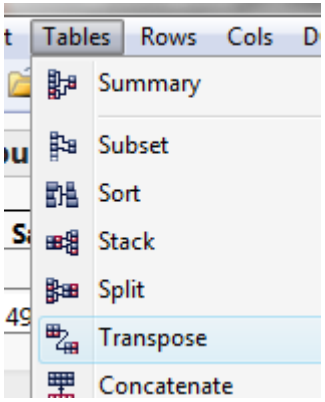
Tabulate

	Sample 0001	Sample 0002	Sample 0003	Sample 0004	Sample 0005	Sample 0006	Sample 0007	San
	Sum	Sum	Sum	Sum	Sum	Sum	Sum	
	4919.1282081	4533.9083558	5035.2592257	4910.6922202	4872.7355664	4743.6046971	4383.3695216	507

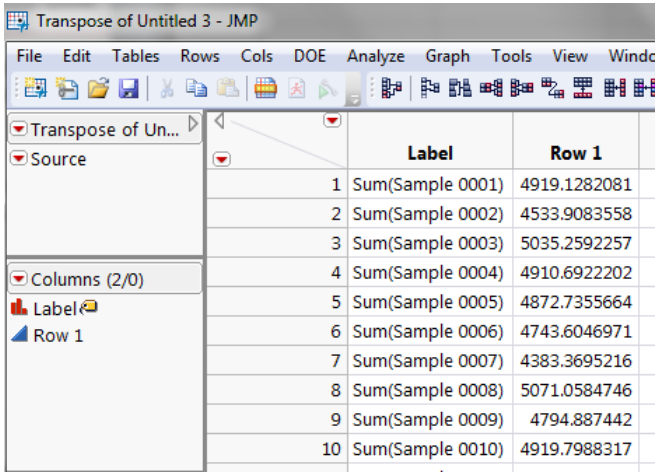
c. Make into data table



d. Transpose data table



e. Resulting data table

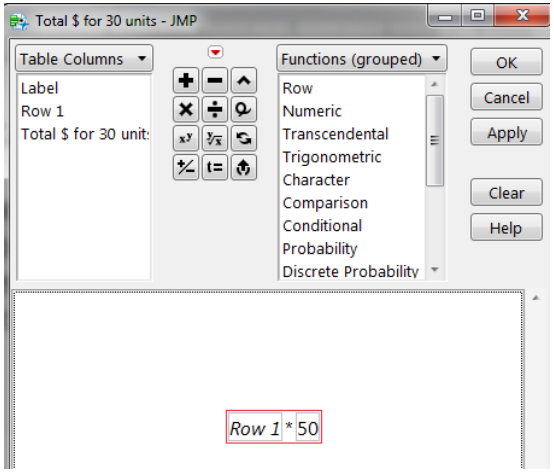


Monte Carlo Steps (cont.)

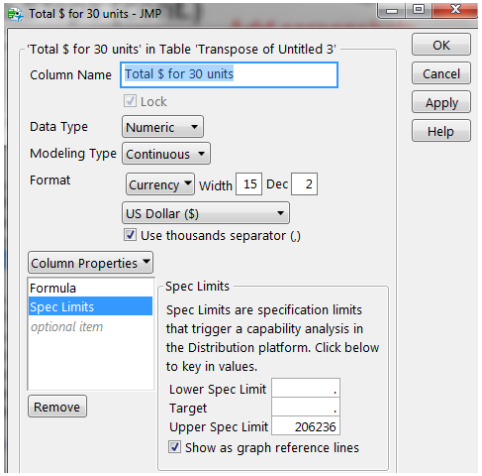
Raytheon
Space and Airborne Systems

9. Create a new column for Total Cost for 30 units by multiplying by $50/hour and add “Spec Limit” based on ETC needed to meet BAC. (assume EAC=BAC).
   - ETC = EAC-Actuals(cum to date) = $600,000 - $393,764 = $206,236

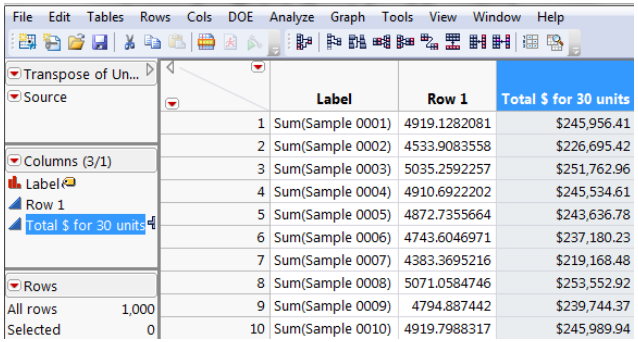
a. Create new column “Total $ for 30 units”



b. Add “Spec” limit



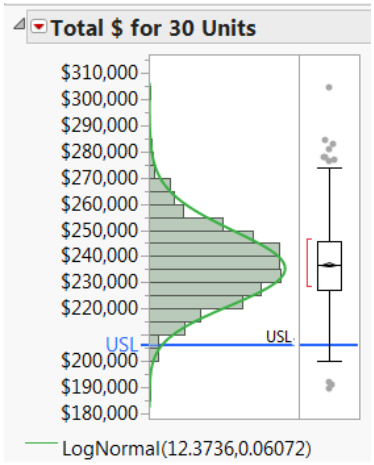
c. Resulting data set



	Label	Row 1	Total $ for 30 units
1	Sum(Sample 0001)	4919.1282081	$245,956.41
2	Sum(Sample 0002)	4533.9083558	$226,695.42
3	Sum(Sample 0003)	5035.2592257	$251,762.96
4	Sum(Sample 0004)	4910.6922202	$245,534.61
5	Sum(Sample 0005)	4872.7355664	$243,636.78
6	Sum(Sample 0006)	4743.6046971	$237,180.23
7	Sum(Sample 0007)	4383.3695216	$219,168.48
8	Sum(Sample 0008)	5071.0584746	$253,552.92
9	Sum(Sample 0009)	4794.887442	$239,744.37
10	Sum(Sample 0010)	4919.7988317	$245,989.94

Monte Carlo Steps (cont.)

10. Analyze resulting column “Total $ for 30 units”



Quantiles

100%	maximum	$304,480
90%		$254,342
80%		$248,372
70%		$243,721
60%		$240,475
50%	median	$236,735
40%		$232,926
30%		$229,371
20%		$224,807
10%		$219,482
0%	minimum	$189,570

Capability Analysis

Specification	Value	Portion	% Actual
Lower Spec Limit	.	Below LSL	.
Spec Target	.	Above USL	98.8000
Upper Spec Limit	206236	Total Outside	98.8000

- LOGO program has >98% chance of overrunning BAC
- Based on Avg hours, we had potential for $30,022 overrun. Now know more exact overrun. 80% of the time the overrun will be less than $42,136 (80% percentile – BAC left=$248,372-$206,236)

Monte Carlo Simulation gives an estimate of Risk of overrun or Opportunity of underrun

Summary

- Accurate financial predictions of Estimate at Completion (EAC) is vital to Raytheon
- Performing Monte Carlo simulations provides a way to add calculated risk/opportunity to the EAC predictions.
  - Determine probability of overrunning/underrunning
  - Determine range of overrun or underrun
- Monte Carlo simulations have many uses. Can also been used for:
  - Cost estimates
  - Proposal cycle time
  - Engineering tolerance stack up
- JMP provides an easy way to perform Monte Carlo simulations