

2001 Joint Statistical Meetings
Atlanta, GA

***An Excel Add-In for
Capturing Simulation
Statistics***

David C. Trindade
david.trindade@sun.com
Sun Microsystems, Inc.
Cupertino, CA

David Meade
david.meade@amd.com
Advanced Micro Devices, Inc.
Austin, TX

Uses of Computer Simulation

Powerful tool to recreate possible outcomes of random processes

Allows checking models and analysis procedures without extensive experimentation

Where theory is limited, may be the only avenue for solving complex problems

Source of data for training

Outline

- Illustrate the capability of computer simulation using spreadsheets for solving many different types of simple and complex problems in design, manufacturing, analysis, etc.
- Describe a general procedure for simulating data for any distribution
- Apply to example for estimating parameters of Weibull distribution
- Provide a macro to capture distribution statistics from specified simulation runs
- Present example results to show power of spreadsheet simulation for analyzing and modeling random processes

Simulating a Distribution

We want to **produce failure times** from a **specific distribution**.

To simulate **any distribution**, we will use a special property of the cumulative distribution function (**CDF**).

*Let **U** be a uniform random variable in the interval (0, 1).*

*If **F(t)** is any **CDF** with an inverse $t = F^{-1}$, then substituting **U** for **F** in the inverse expression generates a **random variable T** distributed according to **F**.*

Weibull Distribution Example

The CDF for a Weibull distribution with characteristic life ***c*** and shape parameter ***m*** is

$$F(t) = 1 - e^{-(t/c)^m}$$

F(t) is the probability of failure by time ***t***.

The inverse expression is :

$$t = F^{-1} = c[-\ln(1 - F)]^{1/m}$$

Simulating Weibull Times to Failure

Substitute the unit uniform random variable U in the inverse expression to get :

$$T = F^{-1}(U) = c[-\ln(1-U)]^{1/m}$$

Since both $1-U$ and U are a uniformly distributed variate in the interval $(0,1)$, simplify the inverse expression as :

$$T = c[-\ln U]^{1/m}$$

Each U value produces a **time to failure** T from a Weibull distribution with parameters c and m .

Weibull Simulation Example in Spreadsheet

- Estimating the parameters c and m of a Weibull distribution is not simple, even for complete data. For MLE's, iterative numerical methods are required.
- Let's investigate the suitability of the following simple procedure for estimating the parameters.
- Write the 25th and 75th percentiles for a Weibull distribution as:

$$t_{25} = c[-\ln(1 - 0.25)]^{1/m} = c[\ln(4 / 3)]^{1/m}$$

$$t_{75} = c[-\ln(1 - 0.75)]^{1/m} = c[\ln(4)]^{1/m}$$

Parameter Estimation for Weibull Distribution

Solve the second equation for c and substitute into the first to get

$$t_{25} = t_{75} \frac{[\ln(4/3)]^{1/m}}{[\ln(4)]^{1/m}} = [0.207519]^{1/m}$$

or

$$m \ln\left(\frac{t_{25}}{t_{75}}\right) = \ln(0.207519) = -1.572534$$

Solve for m directly in terms of the ratio of the two percentiles.

$$\hat{m} = \frac{-1.572534}{\ln\left(\frac{t_{25}}{t_{75}}\right)}$$

Parameter Estimation for Weibull Distribution

With m , we can easily solve for c using any convenient percentile such as the median, that is,

$$\hat{c} = \frac{t_{50}}{[\ln(2)]^{1/\hat{m}}}$$

How well does this procedure work for fitting a Weibull model to data?

To get an answer, let's try simulation in a spreadsheet.

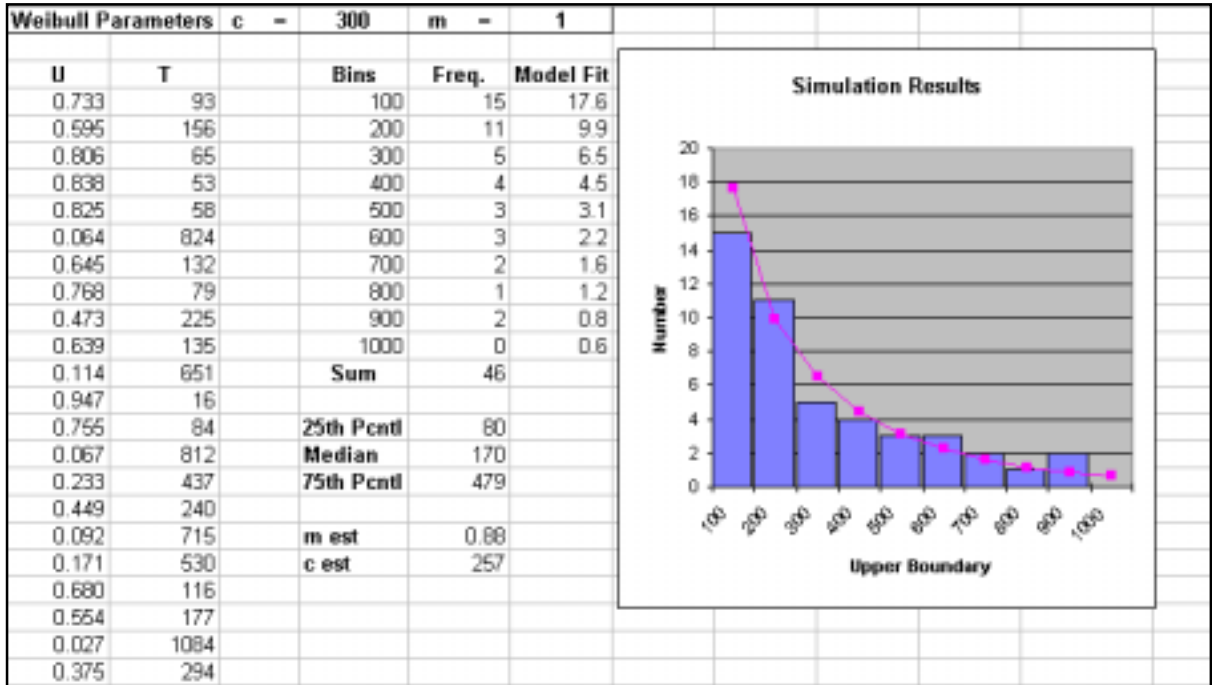
Weibull Distribution Simulation in Spreadsheet

We generate 50 pseudo-random numbers U in a column using the spreadsheet function **=rand()**.

The adjacent column, based on the inverse CDF calculation, contains the random times distributed according to a Weibull with specified parameters c and m .

After estimating the parameters according to the described procedure, we produce a histogram of the data along with a model Weibull fit.

Simulation Results After One Run



Adding Simulation Runs

By making additional simulation runs, and collecting parameter estimates from each simulation, we can determine many facets of the problem, such as:

- the sampling distribution of the estimates
- confidence intervals on the estimates
- the suitability of the Weibull model based on these estimates

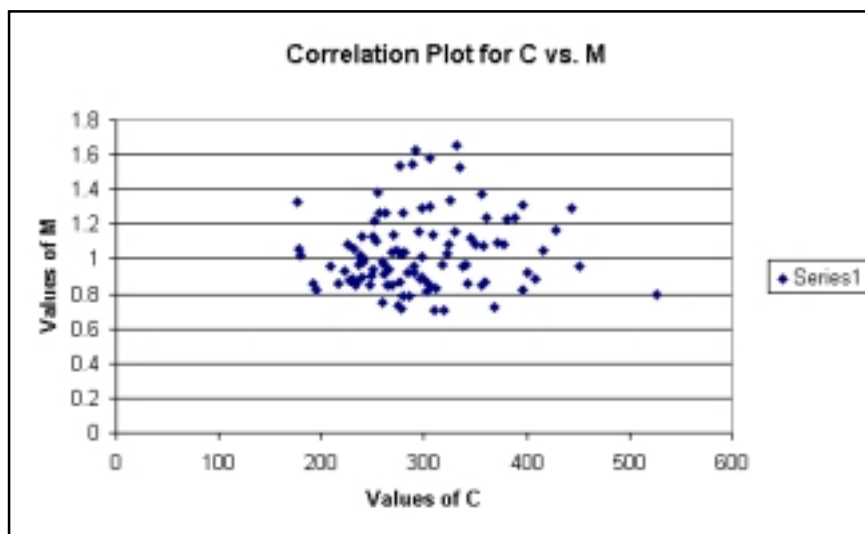
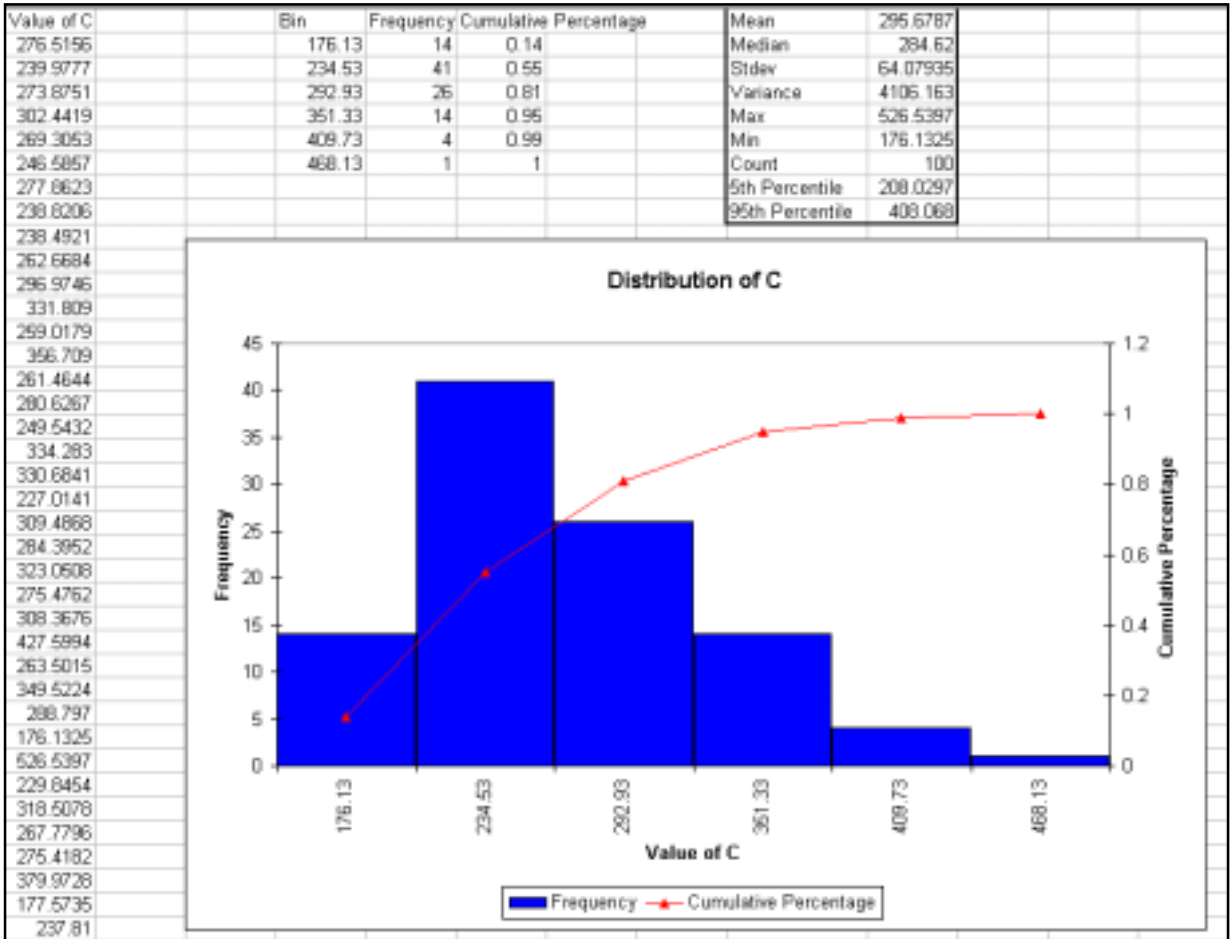
EXCEL Macro Add-In

Simulation Entry Table [?] [X]

Please Enter		Cell References	
Number of Repetitions	<input type="text" value="100"/>	Cell 1	Cell Label
Round to how many digits?	<input type="text" value="2"/>	<input type="text" value="\$E\$20"/>	<input type="text" value="M"/>
<input checked="" type="checkbox"/> Include Correlation Plots		Cell 2	Cell Label
		<input type="text" value="\$E\$21"/>	<input type="text" value="C"/>
		Cell 3	Cell Label
		<input type="text"/>	<input type="text" value="Cell3"/>

OK
Cancel

Macro Output for One Parameter



Benefits of Simulation Studies

Check adequacy and variation in parameter estimation formulas

Determine confidence intervals along with model estimates

Try different representations of displaying results

Experiment with new estimation formulas

Show range of variability in possible system behavior

Summary

Simulation is a powerful tool for understanding and analyzing repairable system processes.