
NEW METRICS FOR EVALUATING MONTE CARLO TOLERANCE ANALYSIS OF ASSEMBLIES

Robert Cvetko

Systems Engineering
Rocketdyne
Canoga Park, CA

**Kenneth W. Chase
Spencer P. Magleby**

Mechanical Engineering Department
Brigham Young University
Provo, UT

ABSTRACT

Monte Carlo simulation may be used by engineering designers to predict the effects of manufacturing variations in mechanical assemblies. Validity of the predictions depends on the accuracy of the input variations, simulation sample size and the fit of a statistical distribution to the resultant assembly data. New metrics are presented for assessing the accuracy of variation analysis methods. Errors due to sample size are estimated for the mean, variance, skewness and kurtosis of the resultant distribution, and the predicted rejects.

Simple algebraic expressions are derived, which can be used to predict the error in the assembly variation parameters without having to repeat the simulation.

1. INTRODUCTION

Variation analysis is an important design tool for bringing manufacturing considerations into design decisions. Dimensional variations occurring in production cause critical assembly dimensions to vary, which can adversely affect performance. Monte Carlo simulation may be used to simulate production variations in mechanical assemblies, resulting from the accumulation of dimensional variation, commonly called tolerance stackup. In a Monte Carlo analysis, sets of sample component dimensions are generated and input to an assembly function, which calculates the resultant size of critical assembly dimensions or features for each assembly. A statistical distribution may be fit to the resultant assembly data and used to predict percent rejects due to exceeding the upper and lower design limits.

The output of a Monte Carlo analysis is a set of parameters defining a statistical distribution that describes the variation in a critical assembly dimension. The accuracy of these parameters depends upon the number of sample assemblies generated in the simulation. Typical sample sizes range from 1000 to 5000. It is not clear to a designer what sample size is required to achieve a desired accuracy. He may repeat the simulation with a different set of random numbers to see how much the parameter values

change, but he has no clear-cut quantitative measures to guide him in choosing the appropriate sample size for each parameter.

1.1 Definition of General Terms and Symbols

The variation in a set of randomly varying dimensions x_i may be described by a statistical distribution. The first four moments of dataset x_i , defined in Table 1-1, are typically used to define the shape of the distribution.

Table 1-1. First four moments of a statistical distribution.

$m_1 = \frac{1}{n} \sum_{i=1}^n x_i$	1 st moment: the mean of x_i (average)
$m_2 = s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - m_1)^2$	2 nd moment: the variance of x_i (spread)
$m_3 = \frac{1}{n} \sum_{i=1}^n (x_i - m_1)^3$	3 rd moment: the skewness of x_i (asymmetry)
$m_4 = \frac{1}{n} \sum_{i=1}^n (x_i - m_1)^4$	4 th moment: the kurtosis of x_i (peakedness)

It is useful to non-dimensionalize each parameter by the standard deviation σ of the distribution, so the distribution parameters can be compared to standard table values. For a standard Normal distribution, the four moment values are 0, 1, 0 and 3, respectively.

Table 1-2. Standardized distribution moments.

$a_1 = \mathfrak{R}(x_i - m_i) / nS = 0$	Standardized 1 st moment: the mean of x_i / σ
$a_2 = m_2 / s^2 = 1$	Standardized 2 nd moment: the variance of x_i / σ
$a_3 = m_3 / s^3$	Standardized 3 rd moment: the skewness of x_i / σ
$a_4 = m_4 / s^4$	Standardized 4 th moment: the kurtosis of x_i / σ

1.2 One-way Clutch Assembly Model

A Monte Carlo simulation of a simple assembly will be used as a case study to illustrate error estimation procedures. The one-way clutch assembly shown in figure 1-1 illustrates a critical assembly feature dimension whose variation is dependent on a chain of component dimensions.

The clutch consists of four different parts: a hub, a ring, four rollers, and four springs. The springs push the rollers into the wedge-shaped space between the ring and the hub. If the hub is turned counter-clockwise the rollers bind, causing the ring to turn with the hub. Turning the hub clockwise causes the rollers to slip and prevents the transmission of torque to the ring.

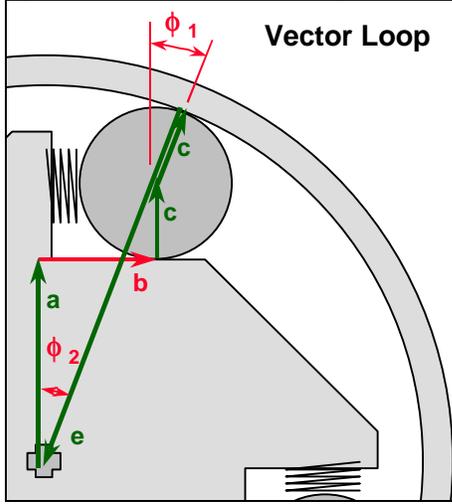


Figure 1-1: Vector Loop for the Clutch Assembly

The contact angle ϕ_1 is a dependent assembly variable which is critical to the performance of the clutch. The variable b , the location of contact between the roller and the hub, is also a dependent assembly variable. The variation in ϕ_1 and b are functions of the variations of the component dimensions a , c , and e , as well as a complex assembly function. The nominal contact angle, when all of the independent variables are at their mean values, is 7.0184 degrees. For proper performance, the upper and lower specification limits for ϕ_1 are based upon this nominal angle and ± 0.60 degree design limits. Table 1-3 below summarizes these values.

Table 1-3: Contact Angle Specifications

Contact Angle	Value (degrees)
Upper Limit	7.6184
Nominal Angle	7.0184
Lower Limit	6.4184

The objective of variation analysis for the clutch assembly is to determine the variation of the contact angle relative to the design limits. Table 1-4 below shows the three independent dimensions and the standard deviation of each. Each of the independent variables is assumed to be statistically independent (not correlated with each other) and a normally distributed random variable.

Table 1-4: Independent Dimensions for the Clutch

Variable	Mean	Standard Deviation
a - hub radius	27.645 mm	0.01666 mm
c - roller radius	11.430 mm	0.00333 mm
e - ring radius	50.800 mm	0.00416 mm

The contact angle for the clutch is described by the explicit assembly function of a , c and e , below, which is nonlinear.

$$f = \arccos\left(\frac{a+c}{e-c}\right) \quad \text{Eq. 1-1}$$

Eq. 1-1 demonstrates that even a simple assembly can result in a complex assembly function. Most assemblies are too complicated to be reduced to an explicit equation. Implicit equations can be found for an assembly using the vector loop method [Chase 1995].

1.3 Vector Loop Model and Assembly Function for the Clutch

The vector loop method uses the assembly drawing as the starting point. Vectors are drawn from part-to-part in the assembly, passing through the points of contact. The vectors represent the independent and dependent dimensions which contribute to tolerance stackup in the assembly. Figure 1-1 above shows the resulting vector loop for a quarter section of the clutch assembly.

The vectors pass through the points of contact between the three parts in the assembly. The vector b and the angles ϕ_1 and ϕ_2 are assembly variables. Since the roller is tangent to the ring, both the roller radius c and the ring radius e are colinear. Thus, $\phi_2 = \phi_1$. Once the vector loop is defined, the implicit equations for the assembly can easily be extracted. Eq. 1-2 shows the set of implicit equations for the clutch assembly derived from the vectors. h_x and h_y are the sum of vector components for the loop and h_q is the sum of relative angles between consecutive vectors.

$$\begin{aligned} h_x &= 0 = b + c\sin(f_1) - e\sin(f_1) \\ h_y &= 0 = a + c + c\cos(f_1) - e\cos(f_1) \\ h_q &= 0 = 90 - 90 + 90 - f_1 - 180 + f_2 + 90 \\ &= -f_1 + f_2 \end{aligned} \quad \text{Eq. 1-2}$$

From the implicit equations, it is seen that ϕ_1 and ϕ_2 are equal. Therefore, the equations can be simplified to just two equations for the two unknowns b and ϕ as shown by Eq. 1-3. The h_y equation may be solved for ϕ , yielding the explicit assembly Eq.

1-1. However, most assembly functions are much more difficult to convert to explicit form.

$$h_x = 0 = b + c \sin(\mathbf{f}) - e \sin(\mathbf{f}) \quad \text{Eq. 1-3}$$

$$h_y = 0 = a + c + c \cos(\mathbf{f}) - e \cos(\mathbf{f})$$

Each of the two assembly equations for the clutch equals zero because the vector loop is closed, and must maintain closure for the parts to assemble. However, vector loop methods may be applied equally well to open loop assemblies [Gao, et al 1998].

2. MONTE CARLO SIMULATION: OVERVIEW

This section presents an overview of the Monte Carlo simulation method for analysis of variation in assemblies and the current methods for estimating its accuracy.

2.1 Background on Monte Carlo

The Monte Carlo Method estimates the accumulation of variation in an assembly due to dimensional variations within the assembly. Variations are described by statistical probability distributions. Figure 2-1 graphically shows how the process works. The output distribution is a function of the distributions of the input variables and the assembly function. Thousands of sets of the input variables are combined to get a reliable measure of the output distribution.

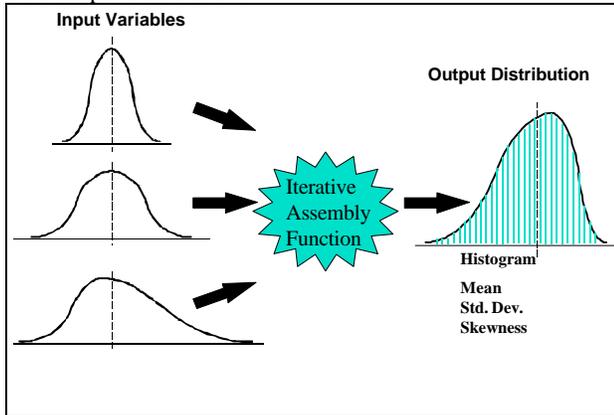


Figure 2-1: Graphical Representation of Monte Carlo Method

An assembly function is derived which describes an assembly dimension or feature in terms of the component dimensions. A Monte Carlo simulation consists of selecting random values for the independent dimensions from their respective probability distributions, and calculating the resultant assembly dimensions from the assembly function (an iterative process is required if the function is implicit). By repeating this procedure for a large number of assemblies, a histogram of the predicted variation in the resultant assembly feature can be plotted.

The results can be compared to specified design limits for the assembly feature. The rejects (assemblies that fall outside the specification limits) can be counted during the simulation, or a distribution can be fit to the moments or percentiles of the Monte Carlo output and the rejects estimated analytically or from standard tables. The most common distribution used to describe assembly variation is the normal or gaussian. Skewed, flat or peaked data may be approximated by more general distributions [Early 1989].

The number of samples required depends on the desired accuracy of the output. Very large samples may be required for accurate results. The relationship between sample size and accuracy will be examined in detail.

2.2 Monte Carlo Simulation of the Clutch

The one-way clutch assembly was easily modeled and analyzed with the Monte Carlo spreadsheet program *Crystal Ball* (since the equation for the contact angle is explicit). More complex assemblies with implicit assembly functions, requiring an iterative solution for each assembly, would be more difficult to model, and would be impractical to analyze with a spreadsheet. Figure 2-2 below shows sample output of a Monte Carlo simulation of the clutch assembly. The histogram estimates the probability distribution for the contact angle ϕ .

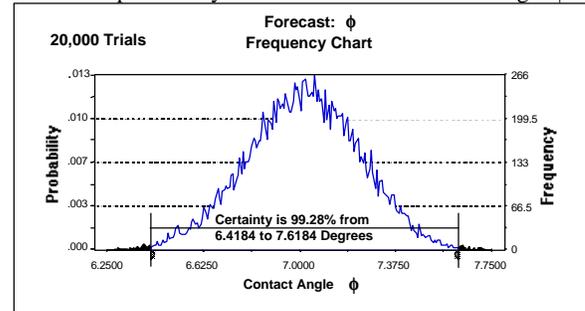


Figure 2-2: Monte Carlo Output from *Crystal Ball*

The histogram shows an estimate of the output distribution for the contact angle. The tails of the histogram represent the assemblies that fell outside the upper and lower specification limits (rejects). With the specification limits as indicated, 99.28% of the sample assemblies are good, and 0.72% of the assemblies are outside the limits. The sample size was 20,000 for this run.

A study was made of the effect of sample size on the results of the clutch simulation. The first four moments of the resultant contact angle distribution were calculated as the sample size was successively doubled. Data was accumulated for successive simulations for efficiency.

A custom Monte Carlo simulation program, using a random number generator, was used to provide greater flexibility and speed in outputting the distribution moment information. This simulation of over a billion samples took 8.6 hours to complete. Figure 2-3 shows how the value of the distribution moments changed during the simulation as the sample size increased. The final moment values are used as the benchmark for calculating the error throughout the following sections.

An earlier study of the effect of sample size compared eight different test problems, using 100,000 samples as the benchmark [Gao 1995]. We shall examine the comparative confidence levels of these widely differing sample sizes in the next section.

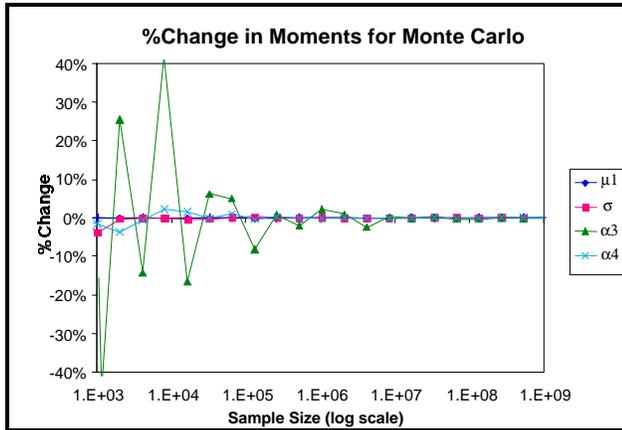


Figure 2-3: Percent Change in Moments for Monte Carlo Simulation as the Sample Size Doubles

Using the percent change in moments as an error measure distorts the results. This is because percent error is amplified when the nominal value is near zero. For example, the percent change of the mean (μ_1) is very small due to the large value of the mean. According to the graph, the error in estimating the skewness appears greater than the error in kurtosis. Since α_3 is near zero, the percent change is large, while α_4 , being much larger, has a much smaller percent change. Using a percent change in the raw moments or the percent error in the normalized moments is not a good estimate for accuracy. A better method of calculating errors will be presented in Sec. 3.

As Figure 2-3 shows, the percent change in the moments as the sample size is doubled approaches zero as the sample size reaches a billion. The benchmark moments and rejects for the clutch assembly are presented below in Table 2-1.

Table 2-1: Benchmark Results from Monte Carlo Simulation at One Billion Samples

Contact Angle for the Clutch	Value
μ_1 (Mean).....	7.014953
σ (Standard Deviation).....	0.219668
α_3 (Skewness).....	-0.094419
α_4 (Kurtosis).....	3.023816
Lower Rejects (ppm).....	4,406
Upper Rejects (ppm).....	2,166
Total Rejects (ppm).....	6,572

2.3 Estimating the Effectiveness of Monte Carlo

How can we estimate the error in a Monte Carlo simulation without having to run a billion samples? If a Monte Carlo simulation is performed for 10,000 sample assemblies, then repeated, the second assembly parameters will have different values from the first, due to the random nature of the simulation process. Each time we repeat the simulation, the results will vary, giving us a "feel" for the reliability of the estimated values. The scatter or probable error may be expressed quantitatively in terms of the standard deviation σ of the repeated values, which corresponds to the spread of the resultant histogram plot of the repeated values. An interval of $\pm 1\sigma$ defines a confidence interval of 68%, $\pm 2\sigma$ corresponds to 95%, etc, which means that

95% of our repeated Monte Carlo runs would fall within the $\pm 2\sigma$ limits.

The confidence interval in using Monte Carlo to predict an assembly parameter may be estimated by means of the binomial distribution [Shapiro 1981]. For example, the reject rate is expressed in parts-per-million (PPM), that is, the number of reject assemblies per million. The standard deviation in predicted PPM rejects may be calculated from Eq. 2-1. An equivalent non-dimensional form, percent of the rejects, is also shown. As the sample size (n) doubles, the error estimating the rejects decreases by $\sqrt{2}$ (for $n \gg 1$).

$$S_{PPM} = \sqrt{\frac{PPM(10^6 - PPM)}{n - 1}} \tag{Eq. 2-1}$$

$$S_{\%PPM} = \frac{S_{PPM}}{PPM} = \sqrt{\frac{10^6 - PPM}{PPM(n - 1)}}$$

Where:

PPM = the estimated parts per million rejects

σ_{PPM} = the one-sigma bound on the error in predicted rejects

$\sigma_{\%PPM}$ = the one-sigma bound on percent error in PPM rejects.

The one-sigma confidence limits calculated by Eq. 2-1 can be multiplied by an appropriate value to obtain any desired confidence level for the estimate of accuracy. The figures presented in this paper will use the one-sigma confidence limits for the errors for easy comparison.

To explain what the one-sigma error means, consider the lower PPM rejects for the one-way clutch assembly. If a Monte Carlo simulation of 10,000 clutch assemblies is run, the rejects correspond to assemblies with contact angles outside of the specification limits ± 0.6 degrees, presented in Table 1.1. The lower PPM rejects correspond to about a 2.6 sigma quality level, or 4,600 PPM rejects. Figure 2-4 below shows a histogram of the percent error for repeated simulations. Each data point on the graph represents one cycle (or run) of 10,000 Monte Carlo samples. A total of 1,000 such runs were performed to generate the histogram. The histogram shows the spread in values obtained by Monte Carlo simulation for a sample size of 10,000.

A sample size of 10,000 would be considered by most Monte Carlo practitioners to be large enough to guarantee accurate results. However, we see from the histogram of repeated runs that there is considerable variation in the results.

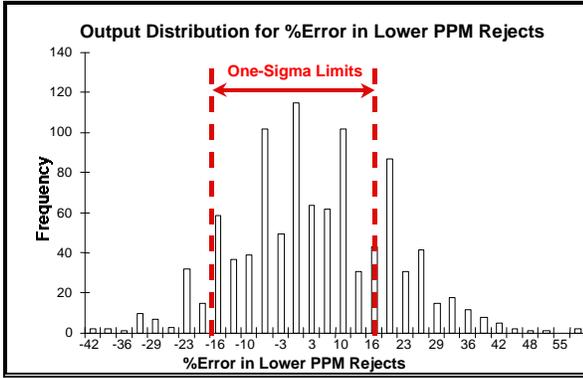


Figure 2-4: Histogram of %Error for Lower Rejects for 1,000 Simulations of 10,000 Samples Each

The histogram portrays the one-sigma limits for the percent error, as calculated with Eq. 2-1.

The mean of the percent error for all 1,000 runs (or cycles of Monte Carlo) was about two percent. The one-sigma limits are about $\pm 16\%$. If the sample size were doubled (20,000), then the one-sigma bound on the percent error ($\sigma_{\%PPM}$) would be divided by the square root of two, and would then be reduced to 11%. Thus, as the sample size increases, the histogram becomes more narrow, and the accuracy increases.

Table 2-2 shows the mean and standard deviation for the percent error of the lower, upper, and total PPM rejects calculated directly from the 1,000 Monte Carlo simulations. The estimates of standard deviation of the percent errors are also compared as calculated by Eq. 2-1.

Table 2-2: Distribution for %Error in PPM Rejects for 10,000 Samples

%PPM	Calculated from	%Error		
		Lower PPM	Upper PPM	Total PPM
Mean	Histogram	2.2%	-9.9%	-2.1%
Stdev	Histogram	15.8%	20.0%	12.6%
Stdev	Eq 2-1	15.2%	20.4%	12.2%

Even at 10,000 samples the one-sigma bound on the percent error for the lower PPM rejects is still about fifteen percent. The important result in the table is that Eq. 2-1 gives a very close estimate of the standard deviation without the need of repeating the simulation.

2.4 Error in Estimating First and Second Moments with Monte Carlo

The method described above for estimating the error in assembly rejects obtained from Monte Carlo simulation may be extended to other assembly parameters of interest to designers. The error in the estimate of the mean of the output distribution is presented in Eq. 2-2.

$$\frac{\hat{m}_1 - m_1}{s/\sqrt{n}} = \text{Standard Normal Variable} \quad \text{Eq. 2-2}$$

$$s_{m1} = \frac{s}{\sqrt{n}}$$

Where: $\hat{\mu}_1$ = the estimate of the first moment of a distribution

μ_1 = The benchmark value of the mean

σ_{μ_1} = the one-sigma bound on the error of the mean estimate.

The standard Normal variable is a non-dimensionalized error estimate. It has been used as an estimate for the error in the first moment [Shapiro 1981, Creveling 1997].

The quantity σ_{μ_1} is also called the "Standard Error of the Means" [Jamieson 1982]. It is a fundamental statistical parameter which requires some explanation. Suppose a large production run is sampled n parts at a time, and the mean of each sample is calculated and a histogram plotted, as was done in the previous section for PPM rejects. As we accumulate samples, the histogram will approach a normal distribution, whose mean (that is, the mean of sample means) approaches the mean of the whole production run. The spread, or standard deviation, of this histogram of mean values is called the standard error to distinguish it from the standard deviation of the whole population. As seen in Eq. 2-2, the standard error for the mean decreases with the square root of n (sample size).

By extending this concept to other assembly parameters besides the mean value of the assembly resultant, new metrics for evaluating Monte Carlo accuracy have been developed, which are presented in the next section.

Estimates of the variability in the second moment of the assembly distribution, the variance, as calculated by Monte Carlo simulation, have been made by applying a Chi-Square distribution. The Chi-Square distribution for variance can be approximated by the normal distribution (with the same mean and variance) for large sample sizes [Vardeman 1994]. This is appropriate, since Monte Carlo generally uses large sample sizes. Thus, the accuracy of counting rejects, and estimating the mean and the variance of an assembly distribution can be evaluated quantitatively. The accuracy of each of these parameters improves as the number of runs increases. Section 3 will expand upon this to estimate the accuracy of the skewness and kurtosis.

3. NEW METRICS FOR ESTIMATING THE ACCURACY OF MONTE CARLO SIMULATION

Section 2 presented the state-of-the-art in terms of estimating the accuracy of Monte Carlo simulation. Eq. 2-1 was used to estimate the one-sigma bound on the error of estimating the PPM rejects. Eq. 2-2 was used to estimate the error for the mean. This section will investigate the standard errors for the first four moments of a distribution as a function of sample size.

3.1 Standard Errors for Estimating Moments with Monte Carlo

The estimate for all of the moments of the output distribution improves as the sample size increases. The one sigma bound curve is a good indicator. The accuracy of the estimate of the mean and PPM rejects of the output distribution was presented in Section 2.

Four error measures for each of the four moments of a distribution are defined in Table 3-1. Each is non-dimensionalized in terms of the distribution variance μ_2 raised to an appropriate power.

Table 3-1. Standard Error Metrics for the Moments of a Distribution

Standard Error Variable	Standard Deviation of the Error	
$SER1 = \frac{\hat{m}_1 - m_1}{\sqrt{m_1}}$	$S_{SER1} = \frac{1}{\sqrt{n}}$	Eq. 3-1
$SER2 = \frac{\hat{m}_2 - m_2}{m_2}$	$S_{SER2} = \sqrt{\frac{2}{(n-1)}}$	Eq. 3-2
$SER3 = \frac{\hat{m}_3 - m_3}{m_2^{3/2}}$	$S_{SER3} = \sqrt{\frac{4}{n-2}}$	Eq. 3-3
$SER4 = \frac{\hat{m}_4 - m_4}{m_2^2}$	$S_{SER4} = \sqrt{\frac{100}{n-6}}$	Eq. 3-4

where

$\hat{\mu}_i$ = the *i*th moment of the distribution estimated from measured samples, simulation, or other variational analysis methods.

μ_i = The benchmark value of the moment

SER_{*i*} = the non-dimensional standardized error

σ_{SERi} = the Standard Error of the Moment.

The standard moment error for each of the four moments decreases with \sqrt{n} (sample size). Using standard error, instead of just the error in the moment itself, eliminates having to scale the estimate of error by the standard deviation or percent.

$$\frac{\hat{m}_i - m_i}{S/\sqrt{n}} = \text{Standard Normal Variable}$$

$$S_{SER1}^2 = \text{Variance} \left(\frac{\hat{m}_1 - m_1}{S} \right) = \frac{1}{n} \quad \text{Eq. 3-5}$$

$$S_{SER1} = \frac{1}{\sqrt{n}}$$

Where σ_{SER1} = the one sigma bound of the estimated mean.

The one-sigma error bound for the estimate of the variance is a little more complicated. The distribution for estimating the accuracy of the variance is the Chi-Square, whose mean and variance is shown in Eq. 3-6. After a little manipulation, the standard error of the variance σ_{SER2} is shown as the last equation.

$$\frac{(n-1)\hat{m}_2}{m_2} = \text{Chi - Square Variable}$$

$$\text{Mean} = (n-1)$$

$$\text{Variance} \left(\frac{(n-1)\hat{m}_2}{m_2} \right) = 2(n-1)$$

$$\text{Variance} \left(\frac{\hat{m}_2}{m_2} - \frac{m_2}{m_2} \right) = \frac{2}{(n-1)} \quad \text{Eq. 3-6}$$

$$S_{SER2} = \sqrt{\frac{2}{(n-1)}}$$

The errors for the skewness and kurtosis are not as easy to calculate as with the mean and variance. Using data from several Monte Carlo simulations, the equations for the one-sigma error bounds for SER3 and SER4 were found empirically to be described by equations 3-3 and 3-4.

To visualize the probability distribution for SER4 of the clutch assembly at $n = 10,000$, figure 3-1 shows a histogram of SER4 for 1,000 runs of Monte Carlo (each at 10,000 samples).

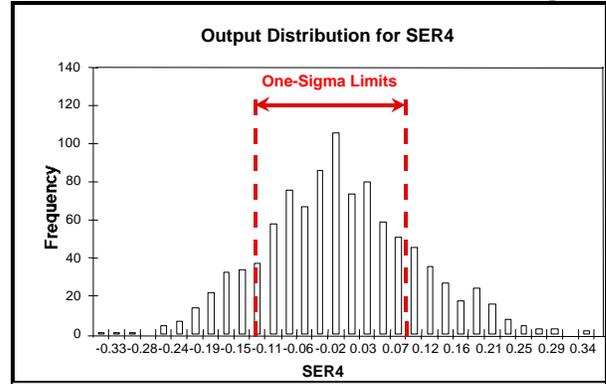


Figure 3-1: Histogram of SER4 for 1,000 Cycles of Monte Carlo, each with 10,000 Samples

The mean SER4 for the 1,000 cycles is very close to zero. The standard deviation of SER4 is about 0.1. The distribution looks slightly skewed to the right, but overall it is not far from normal.

Table 3-2: Probability Distributions for Estimating SER1—SER4 at 10,000 Samples

	Calc. from	SER1	SER2	SER3	SER4
Mean	Histogram	0.0003	-0.0002	-0.0005	-0.0001
Stdev	Histogram	0.010	0.015	0.025	0.108
Stdev	Eq. 3-1-4	0.010	0.014	0.020	0.100

The one-sigma error bounds for the standard errors from the histogram are very close to those predicted by Eq. 3-1 through Eq. 3-4. SER4 is significantly larger than the other standard errors (ten times greater than SER1). To convert the one-sigma error bounds in the standardized moments back to one-sigma bounds on the original moments, Eq. 3-7 below is used.

$$S_{m_i} = S_{SERi} S^i \quad \text{Eq. 3-7}$$

Where: $\sigma^i = i^{\text{th}}$ power of the standard deviation of the distribution (if the actual standard deviation is not known, the estimate can be used, if the sample size is not too small)

3.2 Graphically Representing Error versus Sample Size

A useful graphical demonstration of the error will now be presented. The percent error in PPM rejects for the clutch problem, along with the one-sigma bounds on the error, are plotted in Figure 3-2 versus the log of sample size.

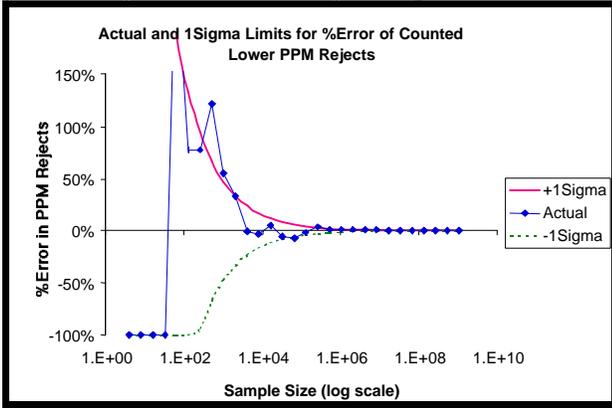


Figure 3-2: Percent Error of Counted Lower Rejects with Monte Carlo Showing One-Sigma Bounds for the Error

The +1 and -1 sigma limits in the figure above show the sixty-eight percent confidence range for the percent error of PPM rejects. The absolute lower bound is -100 percent, as any lower would be predicting negative rejects. The one-sigma bound estimates are calculated directly from Eq. 2-1.

Figure 3-2 illustrates the percent error versus the confidence limits up to a sample size of 100,000; afterward the error is too small to be seen on the graph clearly. To improve the resolution, the same information is displayed on a log-log scale graph below in Figure 3-3. In order to plot the percent error on a log scale, the absolute value of the error was taken. Additionally, the confidence limit bounds are symmetrical, and thus form a single bound. All subsequent graphs of error versus sample size use this same format.

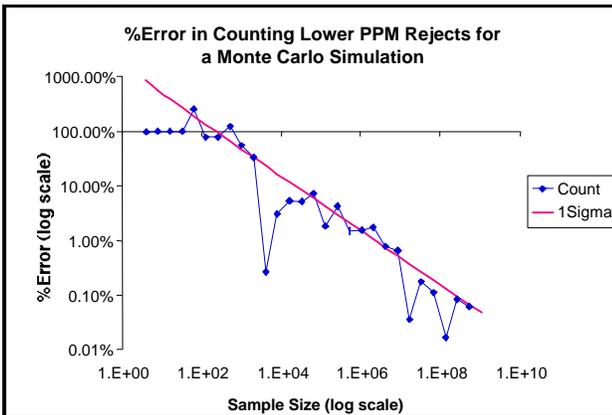
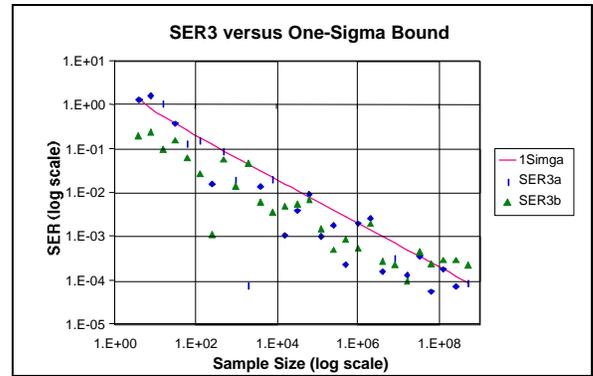
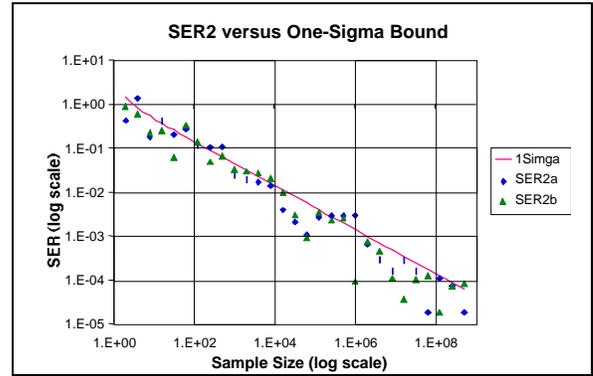
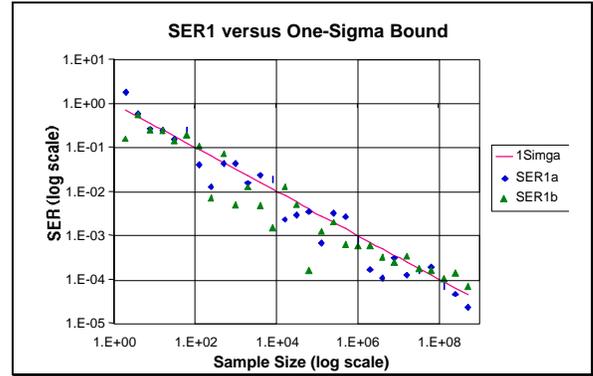


Figure 3-3: Percent Error for Counting Rejects with Monte Carlo versus Sample Size

This graph allows the errors for all of the samples to be viewed together relative to the one-sigma bound. It appears that Eq. 2-1 for the one sigma bound generally describes the error of

using Monte Carlo to count rejects. The term “count” refers to the method of counting rejects during the Monte Carlo simulation and then calculating the resulting PPM rejects.

The theoretical one-sigma limits versus sample size for the standard errors SER1 through SER4 are shown below in Figure 3-4. The actual errors shown are from two different Monte Carlo simulations of a billion samples. Additionally, the one-sigma bounds on error are shown for comparison.



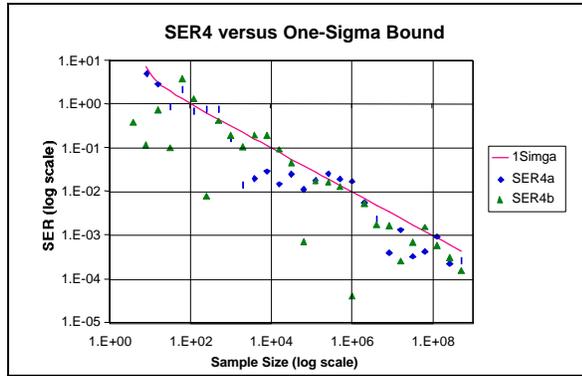


Figure 3-4: Error in Moments versus Sample Size and One-Sigma Bound

The one-sigma bound curves on the graphs are calculated by Eq. 3-1 through Eq. 3-4. Each of the one-sigma bound curves has the same general slope, but is shifted up. Figures 3-4 show that the equations for the one-sigma bounds for SER1 through SER4 estimate the Monte Carlo error well for the whole range of sample sizes. Having an estimate for the error in the output distribution moments as a function of only the sample size is a valuable tool to determine the number of samples to run. Figure 3-5 below shows the four one-sigma bounds on the same graph to allow for easy comparison.

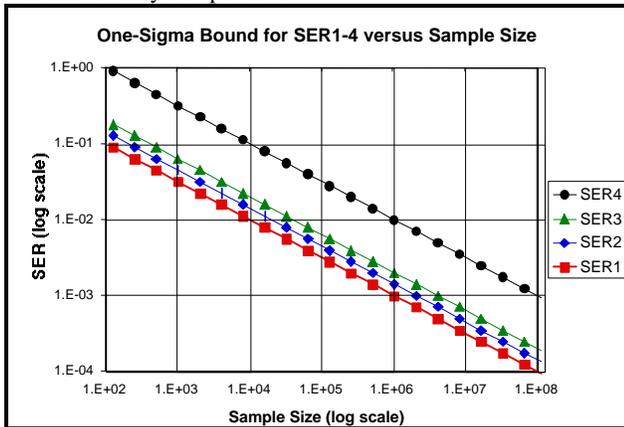


Figure 3-5: The One-Sigma Bound on SER1-4 versus Sample Size

The one-sigma confidence intervals for all four standard moment errors have a very similar slope (decreasing as the square root of n , which is $-1/2$ on a log-log plot). The error in SER4 is 10 times the error in SER1. But, to achieve an error level of .01 for SER4, one million samples is required, versus 10,000 for SER1. The plot can be used to estimate the error in the moments for any sample size.

Other sources of error have been considered: 1) the error inherent in fitting a distribution to the assembly data, 2) the error inherent in the input data. The results are presented in [Cvetko 1997].

3.3 Application of the Standard Moment Errors

The Standard Moment Errors, σ_{SERi} , may be used to estimate the error due to sample size in a Monte Carlo simulation without performing a large sample benchmark test and without repeating the simulation to see how much the results change. Equations 3-

1 through 3-4 and 3-7 can predict the Standard Moment Error for each moment and convert it to dimensional quantities. The actual value will be within one standard deviation of the estimate with a 68% confidence level.

$$\begin{aligned}\mu_1 &= \mu_{1est} \pm \sigma_{\mu 1} \\ &= \mu_{1est} \pm \sigma_{SER1} \sigma_{est} \\ &= \mu_{1est} \pm \frac{1}{\sqrt{n}} \sigma_{est}\end{aligned}$$

Similarly,

$$\mu_2 = \mu_{2est} \pm \sqrt{\frac{2}{n-1}} \sigma_{est}^2 \quad \text{Eqs 3-8}$$

$$\mu_3 = \mu_{3est} \pm \sqrt{\frac{4}{n-2}} \sigma_{est}^3$$

$$\mu_4 = \mu_{4est} \pm \sqrt{\frac{100}{n-6}} \sigma_{est}^4$$

where

μ_i = the i^{th} moment of the distribution for sample size n

μ_{iest} = value of the i^{th} moment estimated by Monte Carlo

$\sigma_{\mu i}$ = the 68% confidence interval of the moment for repeated simulations

$\sigma_{est} = \sqrt{\mu_{2est}}$, estimated by Monte Carlo

Table 3-3 shows the estimated one sigma error bounds using Eqs. 3-8 for the clutch assembly, compared to 1000 repeated simulations of 10,000 samples each. The standard deviation is calculated from 1000 values of the moments obtained from the 1000 Monte Carlo runs. The 68% confidence interval is estimated from the Standard Moment Errors. Monte Carlo Run #1 is shown as a typical example of the results of a single run.

α_3 and α_4 are standardized moments and have more recognizable values. α_3 and α_4 tell us how close the distribution is to a normal distribution ($\alpha_3 = 0$ and $\alpha_4 = 3.0$).

Table 3-3. Comparison of Errors for the Clutch Assembly (Sample Size=10,000)

	Monte Carlo	Monte Carlo - 1000 Runs		68% Conf Int
	Run #1	Max/Min	Std Dev	
μ_1	7.01111	7.02288/ 7.00846	.002203	$\pm .002212$
μ_2	0.04893	0.05036/0.04598	.000717	$\pm .000692$
μ_3	-.00086	-.00011/-0.00184	.000263	$\pm .000212$
μ_4	0.00732	0.00788/ 0.00628	.000251	$\pm .000233$
α_3	-.079445	-.010124/-0.16833	.024465	
α_4	3.05901	3.20252/ 2.85782	.053057	

Most analysts would assume a sample size of 10,000 to be large enough to obtain accurate results. But, the Max/Min values shown in Table 3-3 demonstrate the wide range of values which result from repeated simulations. A single run could be anywhere in this range.

There is close agreement between the standard deviation calculated from repeated Monte Carlo runs and the $\pm 1\sigma$ confidence intervals predicted from the Standard Moment Errors. Comparing the interval variation estimates to the Run #1 moments, μ_1 and μ_2 appear to be sufficiently accurate, while α_3

and α_4 may not be. To reduce the error intervals by a factor of 2 would require 4 times the sample size, or 40,000.

4. CONCLUSION

The principle contribution of this paper is the demonstration of new metrics for estimating the accuracy vs. sample size for five common assembly parameters obtained by Monte Carlo simulation. Further application to a variety of assemblies is needed to validate the new metrics.

The confidence interval predicted from the Standard Moment Error offers an alternative method of estimating the error, without having to repeat the simulation. It is both simple and accurate. It also helps in selecting the sample size, if error reduction is desired.

REFERENCES

- [Chase 1991] Chase, Kenneth W. and Alan R. Parkinson. "A Survey of Research in the Application of Tolerance analysis to the Design of Mechanical Assemblies." *Research in Engineering Design*, 3 (1991): 23--37.
- [Chase 1995] Chase, K. W., J. Gao, and S.P. Magleby. "General 2-D Tolerance Analysis of Mechanical Assemblies with Small Kinematic Adjustments." *J. of Design and Manufacturing*, 5 (1995): 263--274.
- [Chase 1997] Chase, K. W., J. Gao, and S. P. Magleby. "Tolerance Analysis of 2- and 3-D Mechanical Assemblies with Small Kinematic Adjustments." *Advanced Tolerancing Techniques*, John Wiley (1997), Ed. by H-C.Zhang, Chap. 5, 103--137.
- [Creveling 1997] Creveling, C. M. *Tolerance Design: a Handbook for Developing Optimal Specifications*. Addison-Wesley, 1997.
- [Cvetko 1997] Cvetko, Robert, "Characterization of Assembly Variation Analysis Methods", MS Thesis, Brigham Young University, Dec. 1997.
- [Early 1989] Early, R. and Thompson, J., "Variation Simulation Modeling-Variation Analysis Using Monte Carlo Simulation," Failure Prevention and Reliability -1989, ASME Publ. DE-Vol. 16, pp.139-144.
- [Gao 1995] Gao, J., Chase, K. W., and S. P. Magleby, "Comparison of Assembly Tolerance Analysis by the Direct Linearization and Modified Monte Carlo Simulation Methods," Proc. of the ASME Design Engineering Tech. Conf., 1995, 353-360.
- [Jamieson 1982] Jamieson, A., *Introduction to Quality Control*, Reston Publishing, 1982.
- [Shapiro 1981] Shapiro, S. S. and A. Gross. *Statistical Modeling Techniques*. Marcel Dekker, 1981.
- [Vardeman 1994] Vardeman, S. B. *Statistics for Engineering Problem Solving*. PWS Publishing Co., 1994.