Interconnect Failure Rate Estimation Based on the Extreme Value Distribution

Mark Plucinski
Mark Hoffmeyer
IBM Corporation
3605 Hwy North
Rochester MN 55901
(507) 253-6010 plucinsk@us.ibm.com
(507) 253-6606 hoffmeyr@us.ibm.com

Abstract

Evaluating and predicting the failure rate of interconnect is a time consuming and expensive process. Non parametric techniques to analyze the qualification data, such as employing the Chi-square distribution require large sample sizes to achieve an accurate estimate.

Recently there has been a resurgence in the use of extreme value theory (EVT). Increases in temperature records, the numbers of strong storms, and flooding events have fueled this interest. A novel method that is based on EVT and an accelerated degradation model for estimating the failure rate from a set of stress data is proposed and described.

There are many advantages of this technique, and recommendations on sample size are discussed. Advice is given as to how the total sample should be sectioned before the maximum is taken of each subset. Interconnect examples are given, generated from Monte Carlo simulations of known distributions, and used for a comparison of the extreme value technique versus Chi-square and Johnson distribution methods.

Key Words: Extreme Value Interconnect Failure Rate

Introduction

In order to evaluate the reliability performance of new interconnect a series of accelerated life tests are performed. A sufficient number of test vehicles are made using the new interconnect. Low level contact resistance measurements are made at the start and then throughout the regimen of tests that include mixed flowing gas and temperature cycling. At the start of the test, and after subtracting out the bulk contact resistances, the distribution is tightly packed around the nominal value with a normal distribution. For certain mechanisms, failures appear at random on only a small number of contacts. In fact, no contacts may fail during the test, but the interconnect system may have a failure rate that is considered too large for the application. The distribution of resistance changes from the start of the test to the maximum values reached during the test is considered to be the estimate of the field performance probability density for this contact system from a qualification standpoint. If the final histogram reveals a tightly grouped set of resistance changes that fall well below the failure criterion, then the remaining failure analysis and failure rate calculations are straightforward. If the main body of the resistance changes is closely grouped together but there are a small number of outliers, then a failure analysis is necessary to determine the cause of the oddities. However, if the histogram of resistances is positively skewed, then this is distinct evidence of an accelerated stress failure mechanism in effect.

In practice this empirically measured distribution is not well defined. A statistical investigation will in general not identify a particular distribution. Moreover, what is happening in the upper positive tail of the distribution will determine how the contact system will perform. This paper will briefly describe some of the standard techniques used to determine interconnect failure rate, and then describe one technique that focuses its analysis on the tail of this unknown distribution. This procedure can be applied to any other type of data that is similar in nature. That is, where one wants to determine the percentage of the population that exceeds a certain level from a set of data from an unknown distribution.

Standard Failure Rate Methods

The standard failure rate techniques for this type of data are broadly divided into two types, discrete and continuous. The discrete types assume that each contact either passes or fails. These can be modeled with the binomial distribution, Chi-square, or the Poisson. The disadvantage of these techniques is that they generally require more samples than the continuous methods. The continuous methods use measured information from the sample and information about a particular distribution to provide a more efficient and accurate estimate for the same sample size.

The Johnson family of distributions uses four parameters to fit a wide range of data [1]. They can be used to fit the distribution of resistance changes. Once this is done, then the percentage of contacts that would fall above a certain criterion can be estimated, even if no values fell in the failure region during the test. Figure 1 shows an example of a test data set with a Johnson curve fit and a depiction of the area of the curve above the 10 mOhm failure criterion.

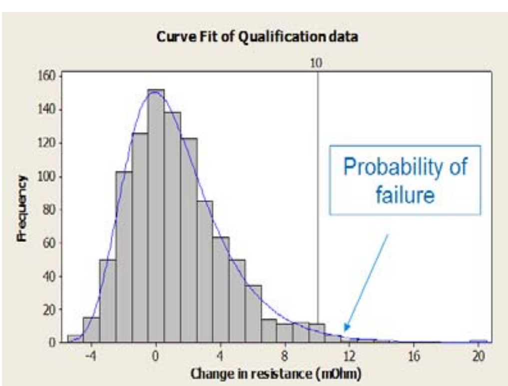

Figure 1: The empirical Johnson distribution fit to resistance change data.

The area under the curve to the right of the criterion is our best estimate of the probability that the population will fall above that value. In order to calculate the failure rate this probability is divided by the accelerated time of the test [2]. Since the actual distribution is not known, but estimated to be a transformation of the normal distribution, there are no confidence limits that can be placed on the estimate.

This technique uses the whole population to estimate the distribution. The middle and each end are given equal weight in estimating the parameters of the empirical distribution [1]. Perhaps it would be better to concentrate on just the positive tail that is in the region of the failure criterion. That is, one should look at the high values to estimate the failure rate.

Largest Extreme Value Background

The average of a set of numbers ($\Sigma x_i/n$) gives an estimate of the center of a population. The more samples one takes, the better the estimate is of the population mean. For contact systems we are not interested in the average, since even for contact systems with low reliability, the average contact usually has acceptable performance.

So consider taking the maximum number (max x_i) from a sample. For interconnect qualifications this would entail measuring the resistances after a prescribed set of stresses, ordering the changes in resistance from the start of the test, and then selecting the highest change for each contact. Notice that the maximum value one selects is a function of sample size. The more samples one tests, the higher the likelihood of getting something in the tails. It turns out that in order to adjust for this, a normalization can be made such that one can determine where you are in the distribution. In fact Gnedenko and others have proved that the distribution of maximum values for any distribution will converge to one of only three possible distributions [3]. These three distributions can be further generalized into one, but the majority of connector qualification data is positively skewed and the maximums converge to the Largest Extreme Value (LEV) distribution. This convergence of maximums can be thought of as the extreme value equivalent of the central limit theorem for taking averages.

As an example consider the daily rainfall for a city somewhere in the southwestern United States. Every day, for 30 years, the rainfall is measured. The distribution is plotted in Figure 2 for one year. When data for all 30 years are collected, then the maximum rainfall for each year is collected as a set. Figure 3 shows the histogram of the maximums for 30 years. Also shown in the figure is the curve fit of the Largest Extreme Value (LEV) distribution, which is what this histogram will converge to as data for more years are collected.

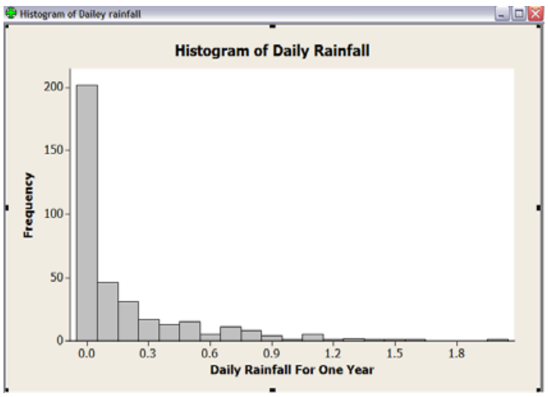

Figure 2: An example of daily rainfall for one year that is shown as a histogram.

Now that one knows the (approximate) distribution of extreme values, estimates of the likelihood of having extremely large rainfall amounts in one day can be quantified. The Largest Extreme Values enables one to extrapolate from the data, but do so in a mathematically precise way. This type of analysis is used in many areas, such as predicting maximum flood levels, maximum wind speeds, and loads on bridges.

Extreme values can also be used to detect unusual outliers. If one has a set of data that is expected to be normally distributed, its maximum value, after being converted to its standard normal deviate, can be compared to what is expected for various sample sizes. For example, if one has a set of 18 values, and the maximum value is calculated to be a mean plus 4 sigma number, Figure 4 shows that this has a less than 1 percent chance of occurring. Figure 4 was generated using Monte Carlo runs on the standard normal distribution, and is an approximate. Gumbel [4] calculates the curves exactly.

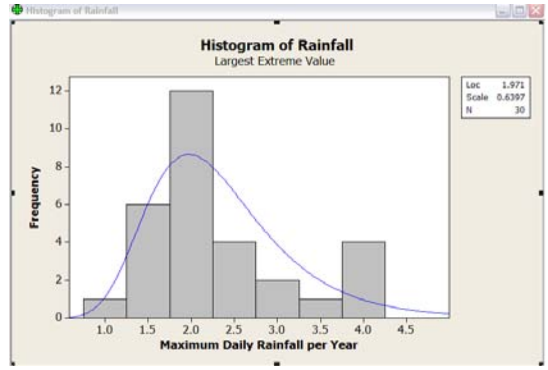

Figure 3: Maximum rainfall for each of 30 years summarized into a histogram.

Largest Extreme Value Used to Calculate the Failure Rate

The process of calculating the failure rate using the LEV distribution is as follows. The set of resistance changes are input and then the sample is segmented into n equal groups of N contacts. The LEV distribution is formed by selecting the maximum from each of the n groups. The percentage of maximums that fall above a 10 milliohm maximum is then calculated by using the maximum likelihood technique to estimate the parameters of the LEV distribution. This is similar to what was done in Figure 1, only just for the set of maximums. This percentage of maximums is then adjusted to the whole population by diving by N. Figure 5 summarizes this process visually.

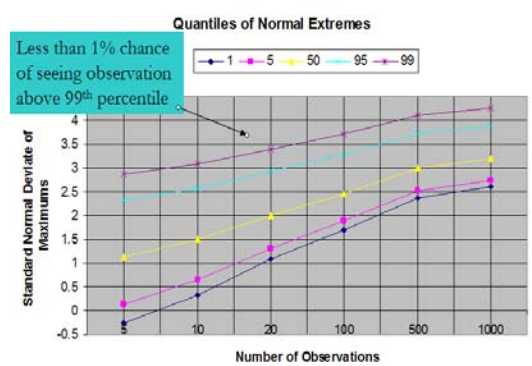

Figure 4: A plot showing the probability of seeing rare events from a normal distribution.

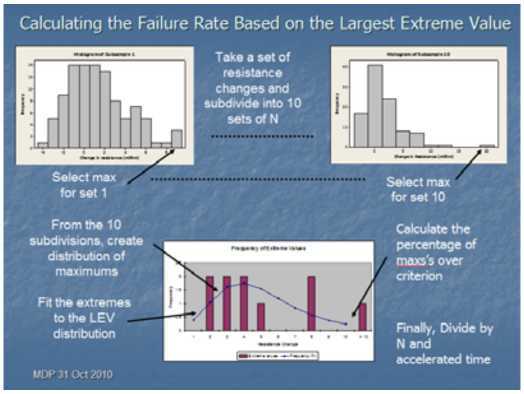

Figure 5: The process showing how to create the LEV distribution from a single data set.

For evaluation purposes this process was programmed into spreadsheet form. Figure 6 shows a screen view of the spreadsheet. The data are plotted to give the user a picture of the distribution, and the failure rate as calculated by using the LEV distribution is given.

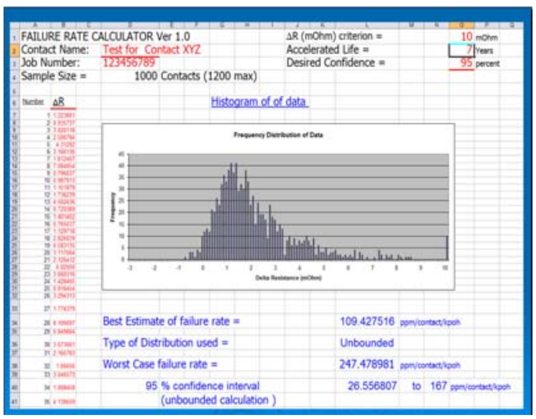

Figure 6: LEV calculation of interconnect failure rate programmed into spreadsheet form.

Evaluation of Largest Extreme Value Technique

In order to evaluate this technique a series of Monte Carlo simulations were conducted. Samples were taken from a known unbounded Johnson distribution that had a failure rate of 1 ppm/contact/kpoh. Total sample sizes varied from 50 to 10,000, and the number of segments the sample was split up into, n, varied from 5 to 1,000. For each sample, the spreadsheet was used to calculate the failure rate. 10 samples were taken for each combination of n and N. The averages of the results are shown in Table 1 and the standard deviations are in Table 2.

Table 1: Average LEV failure rates shown as a function of sample size.

		N										
		5	10	20	25	40	50	100	200	250	500	1000
	5		0.302	0.631		1.14		2.27	3.3		1.83	1.65
	10	0.819	0.194	0.895			1.49	3.08		2.12	2.51	1.71
n	20	0.159	0.889		1.07		1.78			2.03	1.51	
	25			0.821		2.27		1.71	1.8			
	40	0.482			1.1					1.38		
	50		0.803	1.19			1.27	1.11	1.34			

Table 2: Standard deviation of LEV failure rates shown as a function of sample size.

		N										
		5	10	20	25	40	50	100	200	250	500	1000
	5		0.694	1.51		2.48		3.18	6.71		2.35	2.79
	10	1.97	0.49	1.14			1.35	4.19		2.4	2.98	1.73
n	20	0.391	1.45		1.36		1.79			2.23	1.41	
	25			0.878		3.87		1.18	1.89			
	40	0.571			1.02					0.978		
	50		1.13	1.57			0.485	0.83	0.868			

Note that in general, for a given N, as n increases the standard deviation of the samples decreases, so increasing sample sizes reduce the error of the estimate. From Table 1 for small n and small N, the average failure rates fall below the actual value. This must occur because there just isn't the number of samples needed to get any high readings, so the LEV method will tend to under predict for small sample sizes. Also, for small values of n, the variances appear to be large. By inspection, it appears that n must be at least 10. Similarly, at least 25 is needed for N to have better results.

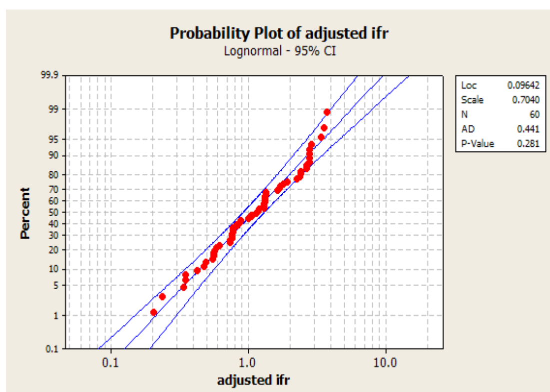

Figure 7: Probability plot showing failure rate results from 60 samples of 10,000 each.

The distribution of results from a sample of 60 simulations of 10,000 contacts each are shown to be approximately lognormal in Figure 7 (p=.28), so the distribution of the estimated failure rates can be transformed to a normal distribution, and tolerance limits can be estimated. The 90% failure rate tolerance limits are shown in Figure 8 as a function of sample size.

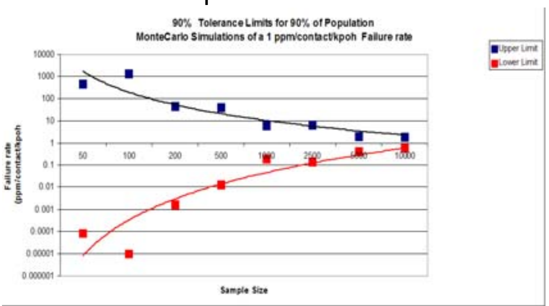

Figure 8: 90% tolerance limits shown for LEV failure rate predictions as a function of sample size.

Figure 8 shows that for a sample size of 50, a wide range of failure rate predictions can occur that cover almost 7 orders of magnitude. Even with 200 samples the errors can be quite large. With 1000 samples the errors become more reasonable, and fall within about plus and minus 1 order of magnitude of the actual failure rate.

For 1 ppm/contact/kpoh the probability of getting a resistance above 10 milliohms, and therefore the area of the curve past 10 milliohms was 0.00006. So for large N, such as 1,000, the expected number of contacts above 10 milliohms is low or 0.06. This number appears to have yielded acceptable results based on the results of Figure 8. For any segmentation this number should be kept low, since the technique as described assumes there is only one contact that falls above the failure criterion level for each set of N contacts. For a more general technique one could extend this process to add in the succeeding order statistics that fall below the maximum. One could also calculate the percentage over the criterion by using the Generalized Pareto Model [5]. However, for relatively low values of expected failures in the sample of size N, the next section shows that this technique is as least as accurate as the one that fits the empirical Johnson distributions.

Comparison of various methods

The LEV method of failure rate estimation is much more accurate than assuming a discrete distribution. Assuming a discrete distribution, and for a test where there have been no fails (no resistance changes over 10 mOhm) out of 1,000 contacts, the best estimate of the failure rate is 0 ppm/contact/kpoh. This can be misleading, as this happens frequently for contact systems that have high failure rates. In this case the best estimate provides little information. Using the Chi-Square formula [6] or assuming a Poisson [7] distribution for the same situation ones calculates the high end of the 95% confidence interval at 48.9 ppm/contact/kpoh. This turns out to be exceedingly conservative. The methods that assume a distribution are considerably more efficient with the data. Figure 9 shows the comparison of the failure rate estimates using the LEV and Johnson methods. Ten data sets of 1,000 contacts each were created by sampling from a known population that had a 1 ppm/contact/kpoh failure rate. Note that both methods are roughly equivalent, and both are susceptible to large variations with this sample size. This experiment was repeated with ten sets of 10,000 contacts each and the results are shown in Figure 10. As expected, the variability is lower. With respect to accuracy, both techniques are approximately equivalent, but note that for two datasets, the Johnson technique did not converge on a solution.

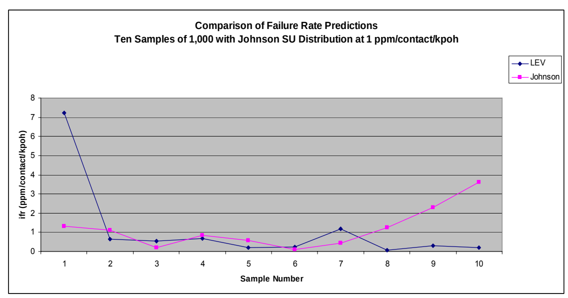

Figure 9: Failure rate comparison using LEV and Johnson techniques for 10 data sets of 1,000 each coming from a 1 ppm/contact/kpoh.

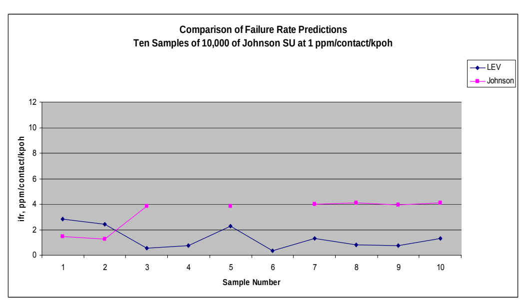

Figure 10: Failure rate comparison using LEV and Johnson techniques for 10 data sets of 10,000 each coming from a 1 ppm/contact/kpoh.

The above figures suggest that the LEV technique is as accurate as the Johnson method. More samples can be taken to quantify this difference, but from a practical point of view the difference appears minimal. However, the LEV method is much simpler to use in practice. No user intervention or judgment is required to set the points where the moments are calculated [1]. Also, the LEV method is much more likely to converge on a solution than the Johnson method, especially with sample sizes less than 1,000. Since the method described here uses the maximum likelihood technique to estimate the parameters of the approximate but known LEV distribution, confidence interval information can be calculated. And as shown, the method is amenable to programming in a simple spreadsheet where the user only has to enter the data and the result is immediate.

Conclusions

A method for analyzing continuous sampled data and extrapolating the fraction of the population that will fall above a critical value was described. The method was compared to other techniques, and found to be at least as accurate and efficient as the current techniques. Moreover, the method is easy to use since it can be programmed and no user judgment or interaction, other than entering the data is needed. The technique also provides confidence bounds along with nominal estimates.

References

[1] J. Slifker and S. Shapiro, "The Johnson System: Selection and Parameter Estimation" Technometrics vol. 22, no. 2, May 1980, pp. 239-246.

[2] A. Choudhury *et al*, "Methodology for Connector Reliability Analysis" Proceedings of the Electronic Components and Technology Conference (ECTC) San Jose, California, May 18-21, 1997, pp. 928 to 935.

[3] Stuart Coles, "An Introduction to Statistical Modeling of Extreme Values", Springer, London, Chapter 3, pp.46-47.

[4] E. Gumbel, "Statistics of Extremes", Columbia University Press. New York, Chapter 4, pp. 129-136.

[5]. Stuart Coles, "An Introduction to Statistical Modeling of Extreme Values", Springer, London, Chapter 4.

[6] Wayne Nelson, "Applied Life Data Analysis", John Wiley and Sons, New York, Chapter 7, pg. 253.

[7] Gerald van Belle, "Statistical Rules of Thumb", John Wiley and Sons, pp. 49-50.