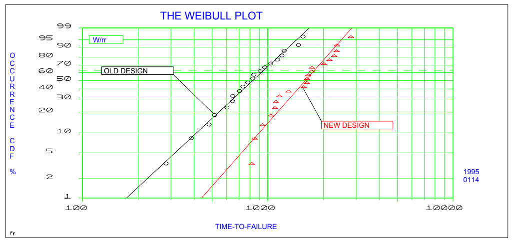
Dr. Bob Replies To E-mail Questions



Dr. Robert B. Abernethy
536 Oyster Road
North Palm Beach, FL 33408-4328, USA
Phone/FAX: 1-561-842-4082
Email: Weibull@worldnet.att.net

Over the years I've answered many Weibull analysis questions from students, friends, and colleagues. This is a short list of many issues and many definitions that you may find helpful in rounding out your Weibull background. Some comments are answers to E-mail. Other answers are comments to questions that flow over the transom. Fast forward to the index.

Why aren't these details listed in **The New Weibull Handbook**, 4th edition, ISBN 0-9653062-1-6? Answers below are intended to be disjointed, spur of the moment questions and answers that many people find helpful---they are not intended to flow as required in a text book, nor are they as terse as you would find in a dictionary. They are intended to educate, inform, and round-out information. Think of this document as my friends say, **"Words From The Master To The Student"**—they will change with time and your background.

The index is upfront so your search is simple and easy. Each word in the index is hyperlinked to the answer. Another hyperlink returns you to the index. If you want to find every instance of the word in this document, use your search function—remember some of the definitions you need are answered under different issues.

This document was written in Word. It is saved as a searchable PDF file to preserve the integrity of the document. You can download the PDF (200 KB).

Send your Email comments and other questions for inclusion into this document to the address shown above.

Last revision: July 3, 2003

INDEX-	

Adjusted Rank Algorithm- (return to index)

April 2003. I want to know if there is more information about how the *adjusted rank algorithm* (that adjusts the plotting positions for suspensions) was calculated. You reference

a Drew Auth as the person that came up with the simplified version of something originally done by Leonard G. Johnson. His book, "The Statistical Treatment of Fatigue Experimentation," has been out of print for some time. Any help you could give in tracking down how this adjusted rank algorithm was derived would be most appreciated. *Dr. Bob's reply: The derivation in Johnson's book is not too clear; however Charles Mischke's ASME paper listed in the References is excellent, very clear.*

Aggregate Cumulative Hazard- (return to index)

ACH, developed by Rolls Royce, detects batch problems using the hazard function. The hazard function is the probability density function divided by the reliability function which gives the instantaneous failure rate at a specific age. The cumulative hazard function is summation of the hazard function over time. The comparison depends upon knowing the existing survivors to the cumulative survivors expected from the hazard function

Average & Standard Deviation- (return to index)

January 2003. Hi Dr. Bob, I was your student last year in the Weibull course at Lohmar, Germany. I see that the *average and the standard deviation* estimates in the SuperSMITH software are different from the simple average and standard deviation (as calculated by Excel, for example). Please explain why there are differences. Thanks a lot for your help. *Dr. Bob's reply: You cannot use the usual equations for the mean and standard deviation because they do not account for suspensions. Instead we solve for eta and beta by whatever method you have selected and then use the Weibull equations in Appendix G to estimate the mean, μ, and standard deviation σ.*

Weibull Mean

$$\mu = \eta\Gamma\left(1 + \frac{1}{\beta}\right) = \text{MTTF}$$

Weibull Standard Deviation

$$\sigma^2 = \eta^2\left(\Gamma\left(1 + \frac{2}{\beta}\right) - \left[\Gamma\left(1 + \frac{1}{\beta}\right)\right]^2\right)$$

Batch Problem- (return to index)

July 2003. In my research, I have come across an example wherein I get some conflicting results when comparing the standard Weibull 2-parameter regression model to the MLE-RBA model (four failures in the example with lots of suspensions). The MLE and regression solutions agree quite well, but the MLE-RBA does not. When doing risk analysis, the MLE and regression models predict quite closely what the actual number of failures are, but the RBA model would lead one to believe that a batch problem may exist. Without delving deeper into the handbook, perhaps you could provide some guidance on this example, as it could impact the research I am doing when I have to use small failure samples. *Dr. Bob's reply: I will have to give you the prize for coming up with interesting problems. Although I admit the "now risk" and the ACH batch detection methods show only a small probability of a batch problem, the number one test - lots of right suspensions, shows strong evidence of a batch problem or phony data. The probability of having 47 right suspensions at the B10 life without a failure is significantly less than 1%. The fact that MLE beta is less than the MRR beta is more evidence, (discovered by Geoff Cole of Rolls Royce). For this small sample we expect the MLE beta to be greater than the MRR beta. You either have a batch problem or bad data. Note that if you delete the 47 right suspensions you have almost a perfect fit and a steeper beta. I suspect there is something wrong about the right suspensions.*

Confidence Intervals- (return to index)

June 2003. I **need to compute the confidence intervals (upper and lower) for the failure rate of a system with three identical components in series. I have the individual component reliability and of course the individual failure rate with its confidence bounds based upon a particular set of measurements. The distribution is log normal and I have the estimated parameters for the probability density function. This data can give me a median CDF with confidence intervals. What I want is to take this same data and get CDF (probability of failure) for a system with three of these components in place and the associated confidence intervals for this system. The CDF would be for one to three failures. I know I can get probability of failure from 1-(Rel)^3 but I do not know how to get the associated confidence intervals.** *Dr. Bob's reply: You won't like my answer. Confidence interval estimates are a function of data, and you have no data on the system. Further, confidence interval estimates are not invariant under transformation which means you must always calculate the confidence estimates last in the calculation. You cannot calculate confidence intervals from confidence intervals as the result is garbage. Therefore, the answer is there is no way you can make the confidence interval estimate for the system without data on the system. ... Can you play games with this? Of course you can, but not rigorously. You might divide the time on the components tested by three and say "if we had tested them in series" we would have had F failures in T time and go from there...As long as your customer understands the meaning of "if" you could do this. Strictly speaking the answer would not be a confidence interval...but your customer might buy it.*

Cramer-Mise and Kolmogorov-Smirnoff- (return to index)

May 2003. My statisticians recommend Cramer-Mise and Kolmogorov-Smirnoff as measures of *goodness of fit* but you do not. Why? *Dr. Bob's reply: Comparing all the known goodness of fit tests the two best goodness of fit tests are the likelihood ratio test and our "p" value for the correlation coefficient squared (Coefficient of Determination), [as shown in Chi Chao Liu's thesis]. The likelihood ratio test described in Chapter 5, is used with MLE or better with MLE Reduced Bias Adjustment. The p value for r squared presented I Chapter 3, goes with median rank regression, X on Y. Of course, either technique may be used with either method. For testing goodness of fit both are excellent and we recommend using both. For distribution analysis, both work well, but you need at least 21 failures to have enough information to make a credible choice with any method. Less than 21 failures always use the Weibull 2p even if you know the data are log normal. If you do not have Chi Chao Liu's thesis you may download in two parts: 1) Abstract (1.2 Meg) 2) Dissertation (15.8 Meg).*

Competing failure modes- (return to index)

April 2003. The failures were produced by two competitive, "dueling ," failure modes. We cannot separate the data into two failure modes. How can we do a failure forecast? *Dr. Bob's reply: The cumulative probability of failure considering both modes is [1-(F(t1))x(F(t2))]. The first step is do a mixture analysis with WSW. If the "p" value supports two Weibulls rather than one Weibull, it is now possible to do an Abernethy risk analysis with the WSW, version 4.0V and later. Alternatively, you could use Monte Carlo Simulation. "RAPTOR" would be a good choice for the simulation software.*

Exponential distribution- (return to index)

May 2003. In Chapter 8 discussing the exponential distribution, you mention the mean time between in-flight shutdowns for commercial engines of 25,000 engine operating hours. As typical commercial flights are 2-4 hours this number seems extraordinary. *Dr. Bob's reply: At the time I wrote that section, 25,000 was the standard; today it is much higher and it is extraordinary. To put that in context, years ago I had lunch with Dr. Von Ohain with the*

President of Pratt & Whitney Aircraft. Dr. Von Ohain invented the gas turbine and developed the Jumo engine for Germany. He said the mean time between in-flight shutdowns on the Jumo engines in the ME262 was 25 hours. When I expressed shock, he said that was good enough as the mean life of the ME262 was 7 hours and 10 minutes. In the six decades that have followed World War II we have made progress[with gas turbines]!

Failures (few data points) - (return to index)

July 2003 When I recently attended your excellent Weibull Workshop you said best practice for 20 failures or less was to always use the two parameter Weibull even if you know that the log normal is the parent distribution. This is very hard for me to believe. Could you give some reasoning for this recommendation? *Dr. Bob's reply: This recommendation comes from Chi Chao Liu's thesis. His massive study of thousands of Weibull and log normal data sets with and without all types of suspensions showed that the Weibull 2P is always more conservative in the lower tail than the log normal. It also showed that for betas of one or less the mean square error of the log normal was much greater, even an order of magnitude greater than the Weibull 2P even when the data was log normal. For betas greater than one the log normal MSE is slightly less than the Weibull 2p for log normal data but the log normal B lives from the lower tail are always more optimistic than the Weibull 2P B lives. Therefore, for engineering problems, we recommend the Weibull 2P for all samples with 20 failures or less. For more detail see Liu's thesis available as a free download from Paul Barringer's Website [abstract or dissertation].*

Interval and Inspection Data- (return to index)

February 2003. With *interval and inspection data* I use the data shortcut to input points at the same value. For example, 88x9 means there nine failures occurred at time eighty-eight. This data appears as the number 9 on the plot. If I change from median rank regression (MRR) to the Inspection Option the plot position of the 9 changes. Why? *Dr. Bob's reply: If you input the 9 failures in separately they would appear as a vertical column of points on the Weibull plot. With MRR the 9 is located in the middle of this column which is approximately where we expect the standard Weibull line to appear. With the Inspection Option the 9 is located at the topmost point of the column, where we expect the Inspection Option line to appear.*

Mean Time Between Failure (MTBF) and Mean Time To Failure (MTTF)- (return to index)

May 2003 What is the difference between MTTF and MTBF? *Dr. Bob's reply: Mean Time To Failure (MTTF) is the average life of the part. Mean Time Between Failures (MTBF) is the average time between failures for all the parts in the fleet. To estimate MTBF divide the total operating time on all parts by the number of failures. With a "complete" sample, (no suspensions) the two parameters are identical. MTTF does not change with or without suspensions other than statistical scatter. MTBF is heavily influenced by suspensions. For example for wearout failure modes a young fleet will have enormous MTBF but MTTF does not change. For a very old fleet that has been through many overhauls and parts replacements, MTBF will converge toward MTTF.*

These two parameters are different and have different applications. MTTF is the parameter of the exponential distribution and is related to eta, the characteristic life of the Weibull distribution. MTBF is used for fleets of repairable systems and is useful for maintainability analysis.

Median bias- (return to index)

February 2003. In the New Weibull Handbook you use *median bias* for most of the comparison of methods and yet you write about *unbiased estimates*. I thought an unbiased

estimate is one whose expected value equals the true value. I am confused as the expected value is the mean value. Correct? *Dr. Bob's reply: You are correct that an unbiased estimate is one whose expected value equals the true value and that the mean value is the expected value. However, the mean value is not a good measure, or a typical measure, for skewed distributions and most life distributions are skewed. We recommend the median value instead of the mean value for skewed distributions. Statistical estimates will split 50/50 around the median. For example, we recommend the median rank plotting instead of the mean rank positions for life data analysis. In the case you selected, five failures, B1 and beta are highly skewed. See the simulation results below. For comparing the accuracy of alternative methods like MRR versus MLE, I use the median bias rather than the mean bias. With MLE-RBA Weibull we employ the median bias correction as the standard or default, but we also offer a mean bias correction in our software. These are quite different corrections; the median bias correction is C4^3.5 versus C4^6 for mean bias. The mean bias correction is much larger.*

MonteCarlo Simulation N=5, *True Values eta=**1000**, beta=**1**, B1=10..... Median /Mean*

	Eta	***Beta***	***B1***
MRR X on Y	***972****/1059*	*1.05/1.179*	*12.5/37.3*
MLE	*942/1028*	*1.25/1.45*	*23.9/57.2*
MLE-RBA	*945/1022*	***1.02****/.991*	***10.4****/22.4*
MRR Y on X	*1068/1124*	*.912/1.05*	*6.3/27.9*

Note that statisticians prefer mean square error to bias as a measure of accuracy. Engineers prefer bias as a measure because they want to know if the estimate is optimistic or pessimistic. If a failure produces a health, or death, or crash risk, engineers want either an unbiased estimate or a conservative estimate. Optimistic bias estimators like small sample MLE failure forecasts and B life are unacceptable. Our conclusion is to recommend methods with the smallest median bias as our best practice. MRR and MLE-RBA for small and moderate sample sizes are examples.

Mixed failure modes- (return to index)

June 2003. I have read much of your Handbook and have been learning how to apply Weibull analysis to benefit our company. I am writing to see if you could elaborate for me on a statement made on p. 46 of the Handbook: "Weibulls for a system or component with many modes mixed together will tend toward a beta of one. These Weibulls should not be employed if there is any way to categorize the data into separate, more accurate failure modes. Using a Weibull plot with mixtures of many failure modes is the equivalent of assuming the exponential distribution applies. The exponential results are often misleading and yet this is common practice." … I have performed analyses on individual failure modes pertaining to our product, but I anticipate a management will request a Weibull analysis for a product with all failure modes lumped together, for the purposes of making decisions regarding warranty terms. According to your statement above, this would appear to be a risky proposition. However, I do not fully understand why. What would be the harm in lumping all failure modes together and performing Weibull analysis in order to determine the overall failure risk for the product? *Dr. Bob's reply: Your method will forecast the same number of claims each month for a constant fleet size...This does this make sense? A better answer is use Monte Carlo simulation. There is free software available to do that called "Raptor." Download it from the net at Raptor.Com. Or better yet, make a failure-warranty forecast for each failure mode from the individual Weibulls and plot them all in WinSMITH Visual, then sum the Y values to obtain the system forecast of cumulative failures by months. For repairable systems, this method produces accurate forecasts. Lumping them all together assumes no wearout modes, no infant mortality, everything is exponential. Does that make sense? You* *might consider attending one of our Weibull Workshops. We will teach you everything about warranty forecasting with Kaplan Meier, the inspection option and Crow-AMSAA.*

Random or exponential distribution- (return to index)

May 2003. Hi Dr.Bob, I'm lost. I have a Beta of 0.80 and the BetaU is 1.03 and BetaL is 0.70. How does that tell me if the failure is close enough to be *"random"/exponential?*
Dr. Bob's reply: Your data is not significantly different from an exponential at whatever level of confidence you selected to get the bounds on beta because the value one lies within the interval. If your double sided confidence bounds are at 90% confidence, your data is not significantly different from an exponential at 90% confidence.

Statistical text references- (return to index)

May 2003. I have recently been reassigned and need to use statistics in my new postion. Could you recommend some *introductory statistical texts*?
Dr. Bob's reply: Some recommendations:
- *"Introduction to Statistical Analysis," Dixon and Massey, McGraw Hill*
- *"Introduction to the Theory of Statistics," Mood & Graybill, Mcgraw Hill*
- *"Introduction to Statistical Inference," Keeping, Van Nostrand...this one is a little heavier..*
- *"The Cartoon Guide to Statistics," Gonick and Smith, HarperPerennial, This is very easy to read and really quite good. paperback.*

The first three are old standards. Suggest AMAZON.COM second hand books.

Unbiased estimate- (return to index)

An unbiased estimate is one whose expected value equals the true value. Note though that using the proportion % as you have is meaningless as a measure of unbiasness. Actually the mean value is not a good measure or a typical measure for skewed distributions and most life distributions are skewed. We recommend the median instead of the mean for life data analysis. For example in the case you selected, B1 and beta are highly skewed. So for my analysis of goodness I use the median value rather than the mean value. Median estimates are 50/50 on the proportion which makes sense with the median. Even with MLE-RBA we have a median bias correction, the standard, but we also offer a mean bias correction in our software. These are quite different corrections; C4^3.5 versus C4^6.

I did a quicky simulation of a small sample of data:

True Values eta=1000, beta=1, B1=10….. Median /Mean				
	Eta	**Beta**	**B1**	Comment
MRR X on Y	**972**/1059	1.05/1.179	12.5/37.3	**←right way to regress**
MLE	942/1028	1.25/1.45	23.9/57.2	**-**
MLE-RBA	945/**1022**	**1.02/.991**	10.4/22.4	**-**
MRR Y on X	1068/1124	0.912/1.05	6.3/27.9	**←wrong way to regress**

Note that for Weibull failure forecasts and B1 estimates, if the failure produces a health, or death, or crash risk, engineers always want either an unbiased estimate or a conservative estimate. Optimistic bias estimators like MLE for moderate and small samples are unacceptable. Our conclusion is to ***recommend MLE-RBA or MRR X on Y as our best practice for small and moderate sample sizes***.

Warranty claims forecast- (return to index)

April 2003. For *warranty claims forecasting by age* you recommend both the Inspection Option and the Kaplan-Meier model. Which should I use? *Dr. Bob's reply: We have research underway comparing all the interval methods, Inspection Option, Probit, K-M, and Interval MLE but it is not completed. Industry seems to prefer the Inspection Option for warranty claims by age but many use K-M. There is a problem with KM. When we use the KM with the actuarial correction, some of your suspensions are eliminated from the data set. This prohibits using the Abernethy Risk forecast as the suspension histogram is wrong. However, this does not effect the estimate of per cent claims at the end of the warranty period which is the usual objective. The Inspection Option does not have this problem. If the probability of repeat warranty claims for the same problem is significant you may want to consider Wayne Nelson's Graphical Repair method described in Appendix M.*

Weibull analysis- (return to index)

Use of the Weibull distribution provides accurate failure analysis and risk predictions with extremely small samples using a simple and useful graphical plots. Solutions are possible at the earliest stage of a problem without the requirement to "crash a few more". Small samples also allow cost-effective component testing. Weibull analysis is a key discipline for reliability, maintainability, safety, and supportability (RMS) engineering largely because of new, credible, and accurate quantitative methods.

Weibull 3-parameter- (return to index)

March 2003. When I use a *three parameter Weibull* the slope, beta, is less than the two parameter beta. Which beta is correct? *Dr. Bob's reply: The 3-P Weibull is a much more complex distribution than the two parameter and we have fixed requirements to meet before we adopt the 3P solution. Remember the four hard fixed rules for using 3-parameter:*

- *you must have 21 or more failures, some experts say 100.*
- *you must be able to explain why the physics of failure support a guaranteed failure free zone,*
- *the 2-p plot should show curvature,*
- *and the distribution analysis must favor the 3-p.*

If you meet all these criteria above the 3-p distribution is the best distribution and the 3-P beta is the correct beta. The 2-p beta is irrelevant. For example, when we have some data sets with missing data and use the 3-p and compare it to data sets without missing data from the same source the 2-p with all the data fall on top of the 3-p with missing data, same beta.