

Reliability & Confidence – How reliable are your products? How confident are you about your reliability?

Reliability is a measure of how well a product will perform under a certain set of conditions for a specified amount of time. Many times, reliability will be stated along with a minimum confidence level. Reliability and confidence are two separate concepts. Reliability refers to a failure rate, while confidence refers to the minimum certainty that the claimed failure rate is accurate.

As suppliers are now required to share greater responsibilities toward the cost of replacing any defective products due to premature failure, it is critical that they have a solid understanding of the functional life spans of their products. Increasingly, companies need to warrant their products to meet specific safety or regulatory requirements for a given life. In addition, purchasing specifications for very high pressure products often require proof of performance to specific reliability levels.

Many factors can influence the reliability of a product, including the consistency of the manufacturing processes as well as the controls on the specifications that govern the product design. The remainder of this article deals with the benefits of reliability testing and how to determine the reliability of your products through testing.

The benefits of understanding your product's reliability.

All manufactured products will usually display some variation in being able to perform to specific requirements over a defined life span. The more consistently a product is manufactured, the more consistently it will perform over a defined period of time. A product that has a lot of scatter in its life span over which it will perform satisfactorily will not be as reliable as a product with lower scatter.

Once the reliability of a product can be assessed, a manufacturer can:

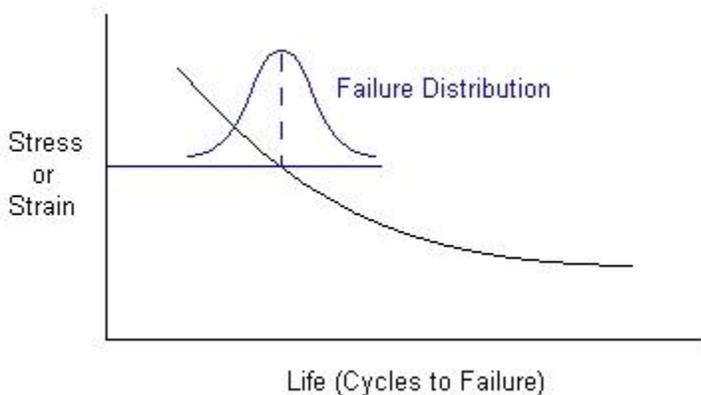
- Determine if a product meets specific reliability requirements or targets as dictated by customer specifications or regulatory requirements.
- Determine if product weaknesses exist.
- Estimate anticipated warranty claims. Also, it is of benefit to be able to reduce warranty costs by verifying effectiveness of product design enhancements or process control improvements on reliability levels.
- Benchmark competitor products. In some cases, your company may be able to gain a competitive advantage by demonstrating a higher reliability than your competitors when performing to similar operational conditions.
- Baseline a particular product and plan for future product enhancements or cost savings. In this way, a company may be able to save money by eliminating costly manufacturing steps which do not lead to greater reliability.

Determining product reliability - Understanding the consistency of your product

Reliability is a statistical measurement which defines the ability of the product to perform to specification consistently. In essence, reliability can be measured by determining the amount of scatter from failure data and performing calculations after fitting the failure data to a model. Two major methods of are modeling can be used: fitting data to a probability density function, and fitting data to a probability density function through the use of a probit model. The former makes use of variable data and measures the *variation in finite life* at a specific pressure level(s). The latter uses attribute “pass/fail” data to measure *variation in survival probability* at differing pressure levels for a given life (possibly infinite).

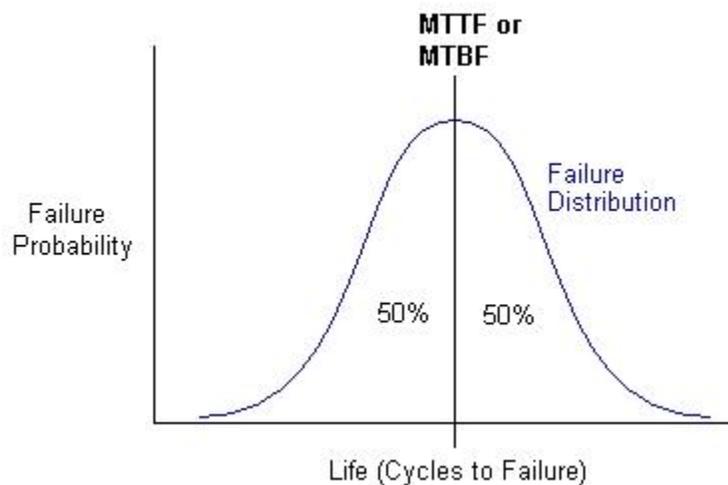
Fitting of variable failure data to a probability density function model in order to determine product reliability. Fitting failure data directly to a probability density function (probability distribution) is the simplest and most direct approach to determining product reliability. This involves testing at a minimum of one pressure range and measuring the scatter in the number of cycles to failure (variable data). This method is best used when the design life for a product is finite.

Failure probability modeling involves using a probability density distribution to express failure rate as a function of life span. Below is shown a typical failure distribution as a function of life in cycles. It should be noted that in the following illustration, the y-axis units of stress or strain can be directly replaced with units of pressure.



In the above illustration a Wöhler curve is drawn as component stress versus cycles to failure. It should be noted that the curve really represents the average fatigue life or 50% failure probability level. The line is the same as the Mean Time To Failure (MTTF) or Mean Time Between Failures (MTBF). Any point on this curve will have a failure distribution associated with it. This means that if

multiple parts are tested at a specific stress (or pressure) level, a smaller percentage of parts (less than 50%) will fail at any lower number of cycles than the curve indicates. Alternately, a greater total percentage of parts (more than 50%) will fail at any higher cycle life. A typical failure distribution illustrating MTTF is shown below.



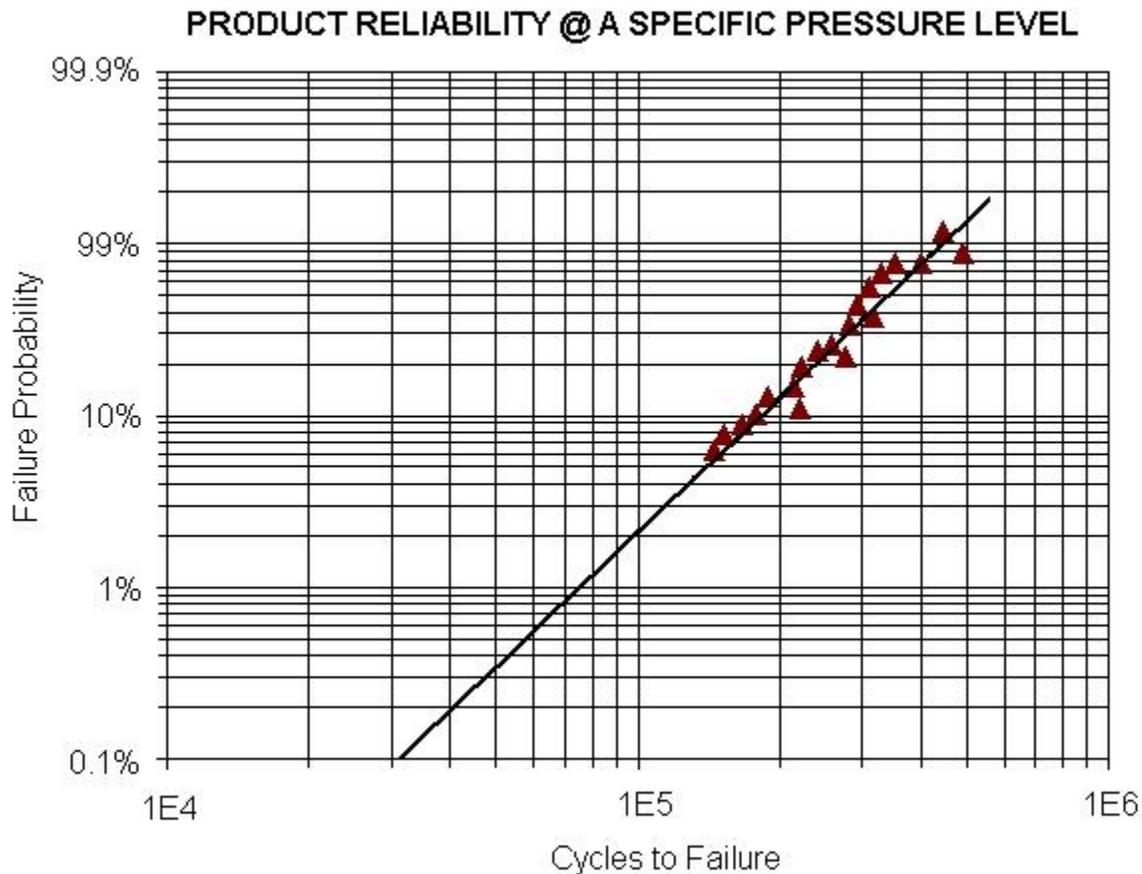
An important consideration of using this method is that it applies when focusing on a single failure mode. Mixing multiple failure modes while analyzing life data using a single distribution is not a recommended practice, unless required for system modeling.

Typical models for probability distributions include: Weibull distributions, log-normal distributions and logistic distributions. Normal distributions are fairly rare for describing fatigue data as the other distributions previously mentioned are much more accurate in fitting the failure data with greater consistency and lower error. The type of distribution actually selected may depend on company preference, customer requirements or best-fit of data.

It is the intent of this method to determine product reliability using tests that result in 100% failures. Testing that does not result in failures (when using this method) tells us very little statistically about the consistency of the product performance. In some cases, it may be necessary to increase testing pressures (stresses) in order to invoke 100% failures in all test samples.

Often, failure probability modeling can be combined with accelerated testing methods to speed up reliability testing. Of course, such techniques are only valid if they do not introduce additional plasticity effects or alter the true failure modes. Plasticity effects (stress ratcheting, stress hardening or stress softening) can occur which can impact the fatigue life of the product. If a test is accelerated, it is always best to test at or slightly above the maximum operating conditions that will be seen in service. A good rule of thumb when testing pressure vessels is to never accelerate a test by raising pressure levels more than 50% above the maximum operating pressure, and scrutinize any testing that requires more than a 10%-15% increase in pressure level beyond maximum operating pressure. Also, it is equally important to avoid testing near the fatigue limit of the product when using this method, as some of the product may statistically not incur failure, resulting in suspended testing. It is difficult to treat these “suspended test” data points unless performing a “Probit” analysis which is discussed later.

Using the fatigue life data from multiple test samples, a cumulative failure distribution can be plotted as shown in the following illustration.



By curve fitting the data to an appropriate distribution, a life span can be determined at various probabilities of failure by reading the corresponding cycle life from the x-axis. From the data in the above graph, the MTTF (50% failure rate) can be read directly as about 330,000 cycles. Also, a 2% failure rate (98% reliability) can be expected at about 100,000 cycles.

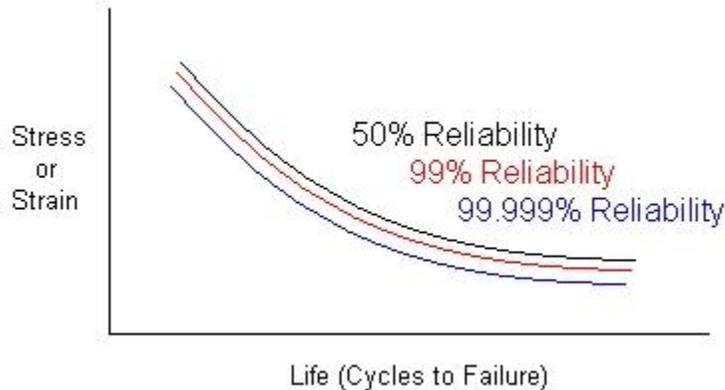
Notice that: $\text{Reliability} = 1/(\text{Failure Probability})$

In a similar way, reliability can be extrapolated from the curve all the way down to the parts-per-million (PPM) failure probabilities.

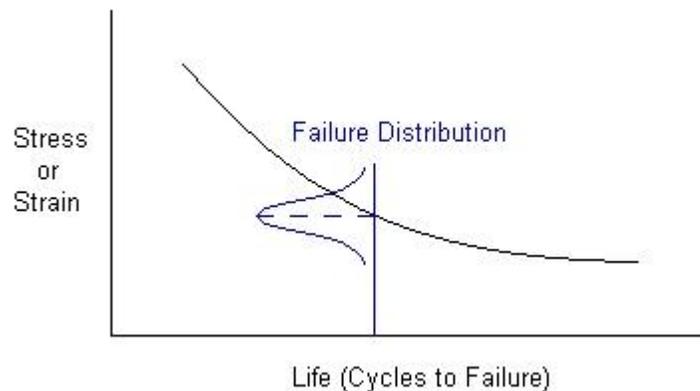
Of special note is that when a Weibull distribution is used to fit the test failure data, the curve will display both a characteristic life (average life) and slope (scatter). The higher the value of the Weibull slope, the less the data scatter. Tables of Weibull slopes for typical product types have been published for years and are available through literature searches.

In some cases, the performance requirements for a product are defined by a set of operating conditions rather than a simple single load case. In these circumstances, it may be necessary to test to a repeated block cycle. Block cycle testing, however, can be cumbersome as it may require numerous repetitive changes in test setups. An alternative to the block cycle approach is to test to a single equivalent damage cycle. In equivalent damage testing, a calculated number of cycles at a user defined pressure range replaces the entire set of operating conditions. The damage assessment for equivalence can be calculated using Palmgren-Miner's rule. Ideally, the calculated life is finite and the user defined pressure range is reasonable.

A more extensive approach to reliability testing a product may involve qualification for multiple applications, or even future applications where performance requirements are still undetermined. In such cases, it is recommended to test at several pressure levels, effectively building a fatigue curve (Wöhler curve) for the product. A minimum of three levels are recommended, with each of the levels in a finite, but realistic operating range (not highly over-accelerated) for reasons as mentioned earlier. In this way, a fatigue curve can be derived for any given reliability level.



Fitting of attribute failure data to a probability density function model in order to determine product reliability using the probit method. Fitting failure data indirectly to a probability density function (probability distribution) through the use of a probit model is a more complex method of determining product reliability. This involves testing at a minimum of two pressure ranges and measuring the scatter in the number of failures (attribute data) prior to the defined end of test. This method is best used when the design life for a product is considered infinite (1 Million – 100 Million cycles) or fixed at a specific number of cycles.



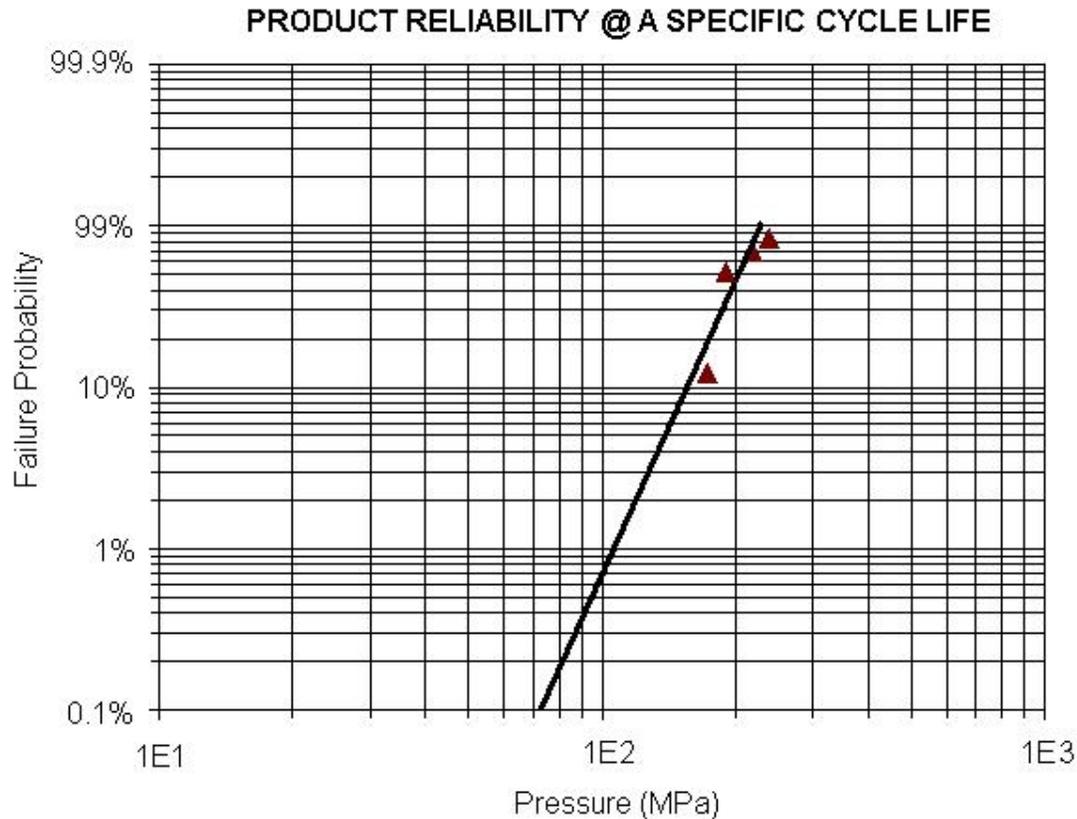
In the above illustration a Wöhler curve is drawn as component stress versus cycles to failure. As stated earlier, the curve really represents the average fatigue life or 50% failure probability level. Any point on this curve will have a failure distribution associated with it (both horizontally and vertically).

In the previous section, we were interested in determining the scatter in the number of cycles to failure (horizontal distribution). In this section, we will consider the failure distribution as it is affected by the stress level (or pressure level) which explores the vertical distribution.

This means that if a fixed number of parts are tested at a pressure level just higher than the mean failure pressure, the majority of the batch should fail. Alternately, if a fixed number of parts are tested at a pressure level just lower than the mean failure pressure, the minority of the batch should fail.

Using this method, we are interested in obtaining the percentage of parts that fail. Testing at any pressure test level is always suspended when the specified number of cycles is reached, and the percentage of failed parts is counted. The failure percentages are then plotted against their respective pressure or stress test levels. It is recommended that a minimum of 10-25 parts be tested at each pressure test level in order to assure accurate, repeatable results. Also, it is recommended that at least 3 pressure test levels be used and that one pressure test level results in a large majority of failures, while another results in a bare minimum of failures. This will result in the most accurate leverage possible to define the fatigue distribution. Either 100% failures or 100% suspensions are not desired, however that limit data can be treated as a statistical maximum or minimum failure rate respectively with some justification, however, that usually tends to add variation to the distribution.

Using the fatigue “percentage failure” data from multiple test batches, a cumulative failure distribution can be plotted as shown in the following illustration.



By curve fitting the data to an appropriate distribution, reliability can be determined at various pressure levels by reading the corresponding axis values. From the data in the above graph, the 50% failure rate can be expected at about 200MPa. Also, a 1% failure rate (99% reliability) can be expected at just over 100MPa.

Recall that: $\text{Reliability} = 1/(\text{Failure Probability})$

In a similar way, reliability can be extrapolated from the curve all the way down to the parts-per-million (PPM) failure probabilities.

Confidence Levels - How many test samples are needed?

Any time reliability of a product is assessed, the assessment itself always comes along with a certain amount of confidence. The only way to increase the confidence level of the testing performed is to increase the number of samples being tested.

If only a small number of samples are to eventually be produced in serial production, only a small number of test samples are needed to accurately represent the total population. It is more likely, however, that a very large number of samples are to be produced. This actually increases the number of test samples needed in order to accurately represent the entire population.

There is a significant amount of software out there to assess the confidence interval associated with any curve fitted data. Also, the further any curve is extrapolated outside of the range of test parameters, the greater the spread in the confidence intervals. These equations are quite complex and vary with the distribution type. Due to this complexity, the following is intended to give a guideline to the number of test samples that will generally be needed.

One way to express the confidence is to state that the mean value of the test data is within a measurable error of the true mean of the entire population with a given confidence, usually in the 90%-99% range. For example, we could set up a test with 95% confidence that the test data accurately predicts the true mean life (or any other reliability level for that matter) within a maximum error of +/- 25%. When the confidence interval is taken into account, there needs to be sufficient gap (conservatism) between the reliability curve and the customer requirement to cover this uncertainty.

A conservative estimate for the number of test samples needed given an acceptable margin of error for a 95% confidence level is:

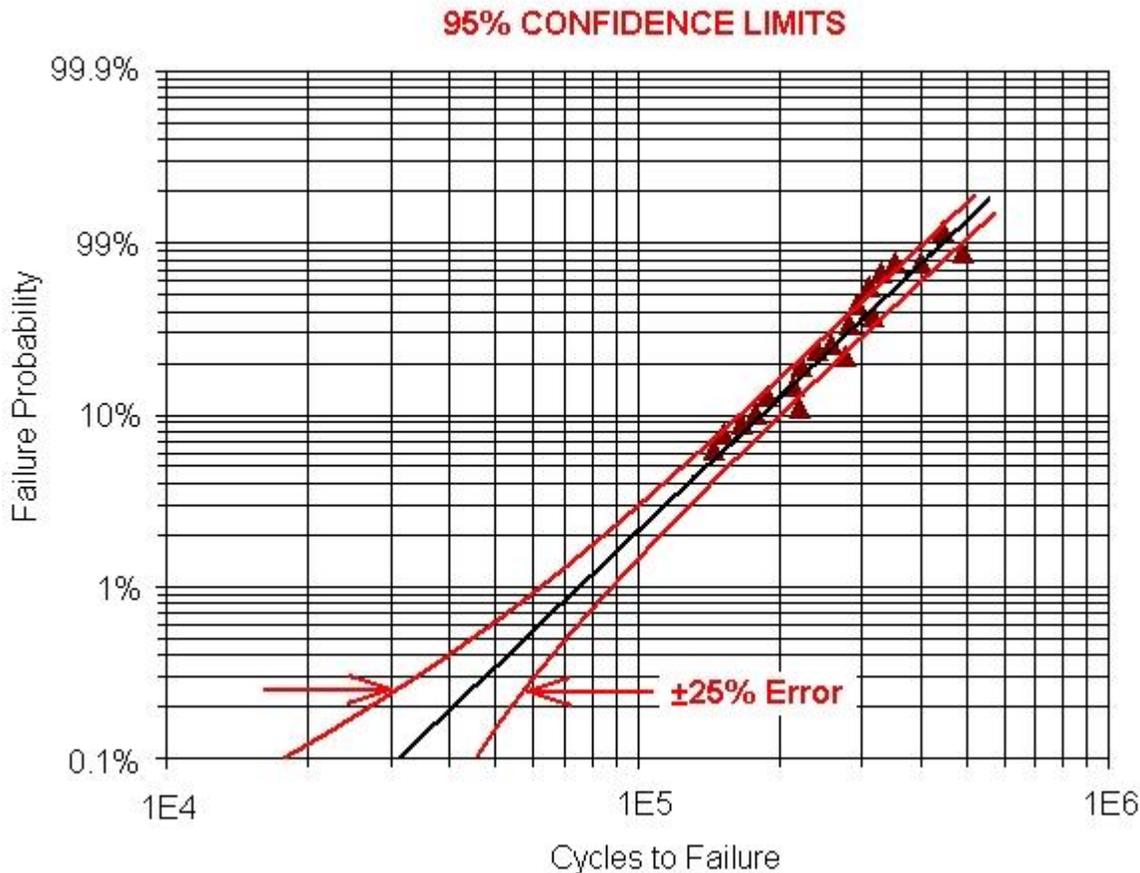
$$N = 1/e^2$$

where:

N = Total Test Samples Needed

e = Acceptable margin of error (+/- %)

Using the above example of 25% maximum acceptable error and 95% confidence, the minimum number of test samples is 16. This can be used to illustrate meeting a specific customer requirement as follows:



In the above example, the customer may have a requirement for 5,000 cycles at the tested pressure level with 99% reliability and 95% confidence. The fitted curve indicates that at 5,000 cycles, a lower-bound 95% confidence limit reliability of 99.4% can be achieved, therefore meeting the customer requirement. Another way of looking at this is that the product is capable of 6,000 cycles at the tested pressure level with 99% reliability and 95% confidence.