S | D | G
System Design Group

SERT

Systems Engineering Review Topics

Writing Reliability Requirements

Reliability requirements are notorious for being poorly written in many specifications. A good reliability requirement must set out what the desired reliability performance is, but also with what confidence that reliability must be demonstrated. This is because reliability measures are inherently statistical in nature.

In order to write a good requirement, it is important to understand how reliability is tested, as this knowledge gives the context within which the requirement will be framed. These notes are intended to describe those aspects of reliability testing that are important to deriving quality reliability requirements.

Population vs Sample Statistics

In writing reliability requirements it is often the Mean Time Between Failure (MTBF) that is specified. It is common, but inadequate, to define a system's MTBF with a requirement such as:

The system shall have an MTBF greater than or equal to 1,500 hours.

It is inadequate because the MTBF of, say, a truck is the average time between failures across the truck fleet. It is a *population statistic* because it describes a level of performance that is to apply across the whole fleet, and so the only way to test it and be 100% certain of the test is to run every truck in the fleet until it fails and average the failure results across the whole fleet. Of course, if the fleet will be manufactured for several years then this approach is impractical.

The alternate approach is to test a *sample* of trucks and average their failure rates to create the *sample MTBF statistic.* It is then possible to *infer* the population MTBF from this limited sample, using statistical methods. This makes the calculation tractable, but introduces several new problems.

The Failure Distribution

Failure data can be characterised as following a range of different probability distributions, depending on the physical system being considered. For systems that exhibit wear out effects the Weibull distribution is popular because the wear out parameter can be adjusted to match historical data. For electronic systems an exponential distribution is often popular because it reflects a memory-less failure history. Thus, just because one component has failed does not mean that another component is also likely to fail.

For reliability purposes many types of systems are routinely assumed to have an exponential failure distribution. Historically this was because exponential failure data produces a constant hazard rate, which made the calculations tractable. Despite the availability of powerful computers, the exponential distribution is still commonly used.

The choice of underlying distribution is important for two reasons. Firstly, it affects how the reliability can be calculated because the mathematical relationship between parameters is different for the different distributions. Secondly, if the distribution selected does not match the actual system failure performance then the model will produce elegant results which will be just wrong! For example, if a system experiences significant wear out, but the exponential distribution is assumed, then the predicted failure rate is likely to vary considerably from the actual failure rate.

A Good Reliability Requirement

An example of a complete reliability requirement is:

The system shall have an MTBF greater than or equal to 1,500 hours at a significance of 10%, demonstrated at a 90% lower confidence limit of 1,000 hours.

It is made up of three key components:

1. *MTBF greater than or equal to 1,500 hours.* This is the benchmark reliability performance requirement.

2. *at a significance of 10%.* This indicates the confidence we need that a system with this MTBF would pass the reliability test.

3. *demonstrated at a 90% lower confidence limit of 1,000 hours.* This indicates how confident we want to be that if the system MTBF was less than 1,000 hours it would fail the test.

The complete requirement puts an upper and lower confidence bound on the test that we want conducted in order to prove this requirement.

Example – The Exponential Distribution

Since the exponential distribution is still widely used in practise, this example will illustrate how to calculate each component of the requirement when an underlying exponential distribution is assumed and the

desired MTBF is 1,500 hours. Calculating these parameters *before* the requirements are released is important because if the test requirements are too stringent then your supplier will have to charge you significantly for the test program. Therefore, calculations should be performed to at least determine how many test hours are required to meet the proposed specification.

The Significance Level (Producer's Risk, Alpha)

Since we are only testing a sample of the items, we must decide how many to test, how long to conduct the test and how many failures would be acceptable during the test. Essentially, this is a decision about the test quality, or the ability of the test to discriminate between a genuinely poor outcome and a poor outcome that is merely a statistical anomaly.

There are two undesirable outcomes that the test should avoid. The first is the risk that we might incorrectly reject a working system, commonly called a Type I error, or an alpha error or the Producer's Risk. This risk is expressed as a probability, typically 1%, 5% or 10%. By declaring alpha=10% we are saying that we want only a 10% likelihood that we might reject the system when in fact the MTBF really is 1,500 hours or greater. In other words, we want a 90% probability of acceptance. We can incorporate this Producer's Risk into our requirement as:

The system shall have an MTBF greater than or equal to 1,500 hours at a significance of 10%.

It is worth considering how we might come to reject such a system, given that its true MTBF really is 1,500 hours. It comes about because the *mean* is just an average, which means that for every test the actual time to failure will be slightly different, sometimes longer than the mean and sometimes shorter. Since these things are random, it is possible to conduct the test and just happen to record a string of results shorter than the mean, or in other words the number of failures could be greater than r, which would fail the test.

By setting an alpha risk of 0.1 we are saying that we want the likelihood of a string of short failure times occurring by random chance to be less than 10%. Of course, it still might happen, but we have designed our test so that the likelihood is now quite small. Thus, if we run the test and it fails we can be 90% sure that it has failed because the MTBF really is less than 1,500 hours, not because of a random chance event.

Exponential and Poisson Distributions

Given that the system is assumed to fail in accordance with an exponential distribution, it can be shown that the probability of a particular number of failures occurring will have a Poisson distribution with parameter T/MTBF. The Poisson distribution is a discrete distribution that answers questions such as "what is the probability of four failure events in time T, given the average number of failures in time T is T/MTBF?".

If we nominate the test duration time, T, and agree to accept r failures then we can use the Poisson distribution to determine the probability of passing this test, given a fixed MTBF. Alternately, for the same MTBF, if we set an alpha=0.1 then we need a 90% probability of experiencing r or less failures during a test of duration T. There are a range of combinations of *r* and T that can meet this 90% requirement, by adjusting either the test duration, T, or the number of permitted failures, r. The simplest combination to pass the test is to run a single item under test for T=160 hours and experience r=zero failures. Under such a test the probability of passing is 90% if the true MTBF really is 1,500 hours.

Operating Characteristic Curves

Operating Characteristic (OC) Curves are a useful way to display the achievable alpha percentage for varying underlying MTBF figures when a fixed T and r have been chosen. Figure 1 is an operating characteristic curve for this test, (T=160 hours, r=0) which shows that if the true MTBF is 1,500 hours then this approach provides a 90% likelihood of passing the test (ie, experiencing zero failures over 160 hours).

Unfortunately, the OC curve also shows that if the true MTBF is actually only 1,000 hours this test is still 85% likely to pass. From the customer's perspective this is unlikely to be acceptable and it is referred to as the Consumer's Risk, or the risk of accepting a system with lower reliability performance than that desired. The Consumer's Risk is managed by demanding that the test has a certain *power*, characterised by the beta number.

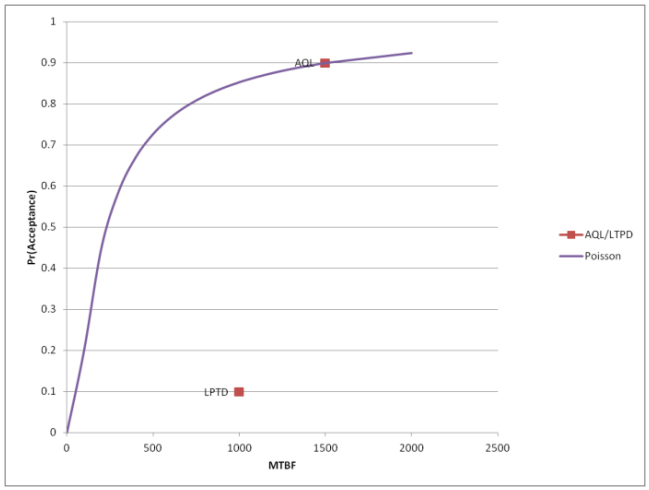

Figure 1 - OC Curve T=160 hrs; r=0

The Test Power (beta)

If the true MTBF of the system is only 1,000 hours then we need to design a test that the system would be very unlikely to pass. The proposed test (T=160 hours, r=0) does not provide this assurance, so it needs to change. A typical requirement might be that the reliability must be demonstrated at a 90% lower confidence limit of 1,000 hours. This means that passing the test should prove that there is not more than a 10% chance of the true MTBF being 1,000 km or less. In this case the beta=0.1 and we say that the power of the test is 1-beta=0.9

To make the test this powerful we will need to find a different combination of T and r that can achieve this beta requirement. Typically, we need to extend the duration, T, in order to increase the power of the test. Of course, if we increase T, but not r, then we effectively increase the reliability required, which increases the risk that our system will now fail the test. In other words, we have reduced the probability of acceptance (alpha) that we set up earlier. We can compensate by permitting additional failures (increasing r), but this tends to reduce the test power. Increasing the test duration generally reduces the Consumer Risk, but increasing the number of failures generally reduces the Producer's Risk, so the two measures are related. A longer test is also more expensive, so we would prefer to keep them short.

Finding a pair of values for T and r that can satisfy both the Producer's Risk and the Consumer's Risk is a non-linear optimisation problem. There are tabulated values in early military handbooks, however, with modern spreadsheets it is easier to let the computer do the calculations.

Figure 2 shows the OC curve for a test of 47,080 hours duration that permits 38 failures. This test has a 90% probability of acceptance if the true MTBF is 1,500 hours and also a healthy 90% probability of rejection if the true MTBF is less than 1,000 hours. Notice that by reducing the Consumer's Risk the test duration is several orders of magnitude larger than for the first example.

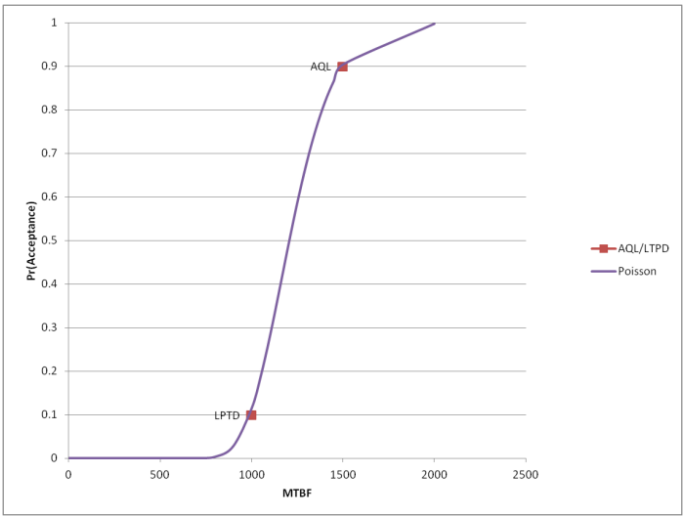

Figure 2 - OC Curve T=47,080 hrs; r=38

A Complete Requirement

Hence we achieve the complete requirement:

The system shall have an MTBF greater than or equal to 1,500 hours at a significance of 10%, demonstrated at a 90% lower confidence limit of 1,000 hours.

Note that in developing this requirement we now know the impact of the requirement on the test program. If this were a vehicle program then we must budget for 47,000 kilometres of testing. Of course that could be a single vehicle, or it could be four vehicles travelling around 12,000 kilometres each. In either case, the test must not produce more than 38 failures. A basic calculation shows that 38 failures would produce a sample MTBF of 1,239 kilometres, which is less than the required population statistic of 1,500 kilometres. Nevertheless, we can be 90% confident that the true MTBF is greater than 1,000 kilometres and under our chosen test conditions we would accept the hypothesis that the true population mean is 1,500 kilometres.

About the Author:

Dr Ian Brace is a systems engineer with a background in communication systems and signal processing. He has consulted widely to industry and government in the areas of systems engineering and capability analysis and he has over twenty years experience in the defence domain.

Ian.Brace@systemdesigngroup.com.au