SERT

Systems Engineering Review Topics

Untangling Reliability Distributions

Introduction

In the field of reliability there are several probability distributions that regularly come up, but it is not always obvious why they are used or how they are related. For example, electronic equipment is often modelled using an exponential distribution because of its memory-less property. However, sampling plans for reliability typically invoke the Poisson distribution or the Chi-Square distribution. This SERT Sheet will explain the main distributions that are invoked for reliability tasks and explain how they relate to one another.

Binomial Distribution

The binomial distribution is a discrete distribution that predicts the number of failures when *k* items are sampled from a batch of *n* items. When developing an acceptance sampling plan we typically use the binomial distribution because it applies to events that can be given a pass or fail criterion, which is typically some form of attribute test. For example, in a batch of light bulbs does the sample bulb work or not? In a packaging process does the product weigh at least a kilogram? In a communication system are calls connected within less than five seconds?

Let's assume that we are going to test *n* items and the true probability of failure is *p*. Then the probability that *k* failures will occur in the batch of *n* items is given by:

$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

This is simply the probability of *k* fails and *(n-k)* items not failing, when all possible groupings of *k* items are included. For example, consider a batch of *n*=3 items labelled A, B and C. If the underlying probability of any single item being defective is p=0.1 then the probability of exactly one item being defective in a batch of three is given by

$$\Pr(X = 1) = Pr(\text{A is defective AND B is OK AND C is OK}, \\ OR\ A\ is\ OK\ AND\ B\ is\ defective\ AND\ C\ is\ OK, \\ OR\ A\ is\ OK\ AND\ B\ is\ OK\ AND\ C\ is\ defective)$$

$$Pr(X = 1) = 3(0.1^1 0.9^2)$$

Cumulative Binomial Distribution

A successful test occurs when *k or less* failures occur during the test. The probability of a successful test is therefore the probability of zero fails or a single fail or two fails - up to *k* fails. This is expressed as:

$\Pr(X \le k) = \sum_{i=0}^{i=k} \binom{n}{i} p^i (1-p)^{n-i}$, and is known as the Cumulative Binomial Distribution. Using this formula we can calculate values for *n* and *k* to ensure that we have a specific probability of passing a test if the underlying true probability of failure is *p*. How to apply this technique to an attribute sampling task is discussed in SERT 1-2 *Acceptance Sampling Plans*.

Exponential Distribution

Modelling reliability is similar in principle to modelling acceptance sampling, but the distributions are continuous rather than discrete. This is because we typically model the Mean Time Between Failure (MTBF), which is a continuous random variable. Hence, the data for the model is simply a list of measured times between successive failure events, such as 1,506 hours, 1,457 hours, 1,83 hours, etc.

Matching this data to a probability distribution is usually based on historical evidence and our experience of the failure performance of similar equipment. Each potential distribution is defined by one or more parameters which change the overall *shape* of the distribution. For example, the Normal distribution is defined by two parameters, the mean μ and the standard deviation σ.

The exponential distribution is defined by a single parameter, λ, called the failure rate and is defined as

$f(t;\ \lambda) = \lambda e^{-\lambda t}$ for $x > 0;\ 0$ otherwise

An interesting property of the exponential distribution is that it is *memoryless*. This means that the probability of a failure occurring is not affected by any previous occurrences of failure. Or in other words, the fact that three failures have occurred does not increase the likelihood of the next failure. The distribution has a constant *hazard rate*, or a constant conditional probability.

Historically, the constant hazard rate made this distribution attractive because calculations were easier than with some other distributions. It has continued to be popular, despite the availability of computing power, although the reality of the constant hazard rate should be assessed against the operation and failure of the system being considered. Generally, to apply the exponential distribution the system needs to be complex and should not experience significant wear out failures. Electronic systems are often modelled with exponential failure behaviour.

For systems modelled with the exponential distribution the MTBF is defined as

$$MTBF = \frac{1}{\lambda}$$

Gamma Distribution

The Gamma Distribution is a two parameter continuous distribution, characterised by a shape parameter, *r*, and a scale parameter, θ. If a random variable *X* has a Gamma distribution then we say

$X \sim Gamma(r, \theta)$.

If we run a timed system trial in order to measure failure data, and the time between each failure event is modelled as t_i , then it would be useful to predict the actual time of the next failure event, T_i. This idea is illustrated in Figure 1.

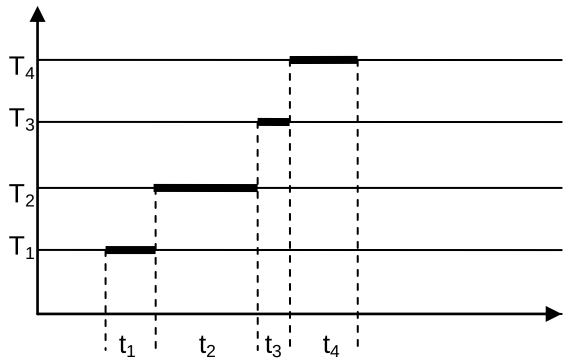

Figure 1 - Time Between Failure vs Time to Failure

Along the *x-axis* we see the time between failures, t_i, which we assume are distributed in accordance with an exponential distribution with MTBF= 1/λ. Along the *y-axis* we see the actual time at which each failure event occurs, T_i. Clearly, $T_i = t_0 + t_1 + \ldots + t_i$.

It can be shown that a random variable which is the sum of a series of exponential variables will have a Gamma distribution. Hence, the time T_i of each failure will be distributed as a Gamma variable. For reasons of convenience it is preferable to measure λT rather than just T, which then allows us to set the shape parameter θ=1 and the scale parameter becomes *r+1*, where *r* is the number of failures that occur during the total test period T. Hence we can say

$\lambda T \sim Gamma(r+1, 1)$

Chi-Square Distribution

The Chi-Square distribution is a single parameter continuous distribution that is a special case of the Gamma distribution. Given that λT ~ Gamma(*r*+1, 1), by the properties of the Gamma distribution 2λT ~ Gamma(2(*r*+1), 1). But Gamma(2(*r*+1), 1) is equivalent to Chi-Square(2(*r*+1)). Hence we can say that for a reliability trial of duration *T* that includes *r* failures the test statistic 2λT will be distributed as a Chi-Square distribution with 2(*r*+1) degrees of freedom.

Although we arrived here along a somewhat circuitous route from our initial assumption of exponential failure data, this is a useful result. For a specific trial of duration *T* which experiences a particular number of failures, *r*, we can now use the Chi-Square distribution to calculate the probability that our system has a particular MTBF.

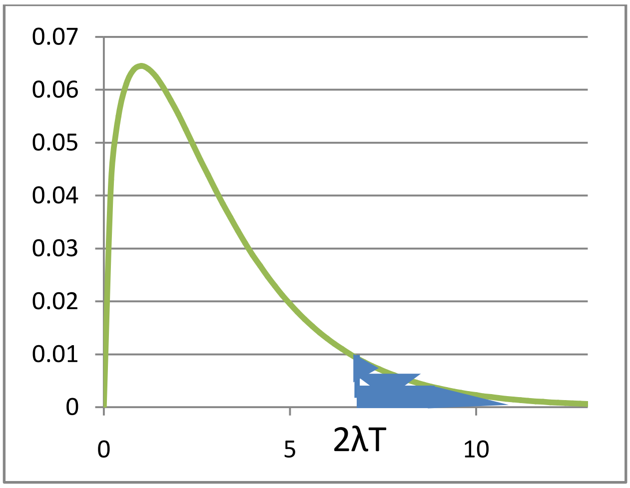

Figure 2 - Chi-Square Distribution with 3 Degrees of Freedom

Figure 2 shows the failure distribution for such a test - a Chi-Square distribution with 2(*r*+1)=4 degrees of freedom. To determine a 90% lower confidence limit on the MTBF we need only calculate the value of λ such that the upper tail area is 10%. [Note that since λ=1/MTBF, the larger the value of λ the worse (smaller) the MTBF.]

Intuitively, the reason this works is that Chi-Square is a special case of the Gamma distribution, and the Gamma distribution describes the probable time, T_{i+1}, of the next and subsequent failure events. Hence, the upper tail integral is the probability (10%) that, for the given λ, the $r+1^{th}$ failure event will occur after time *T*. This is the same as calculating the 90% probability that, for the given λ, only *r* failures would occur prior to time *T*.

Hence, the 90% lower confidence limit for a time truncated reliability test can be expressed as

$$MTBF = \frac{2T}{X^2_{(0.9,2(r+1))}}$$

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.