APPENDIX B

THE MATHEMATICS OF RELIABILITY

B-1. Introduction to the mathematics of reliability

This appendix provides the reader with an overview of the mathematics of reliability theory. It is not presented as a complete (or mathematically rigorous) discussion of probability theory but should give the reader a reasonable understanding of how reliability is calculated. Before beginning the discussion, a key point must be made. Reliability is a design characteristic indicating a product's ability to perform its mission over time without failure or to operate without logistics support. In the first case, a failure can be defined as any incident that prevents the mission from being accomplished; in the second case, a failure is any incident requiring unscheduled maintenance. Reliability is achieved through sound design, the proper application of parts, and an understanding of failure mechanisms. **It is not achieved by estimating it or calculating it**. Estimation and calculation are, however, necessary to help determine feasibility, assess progress, and provide failure probabilities and frequencies to spares calculations and other analyses. With that in mind, let's first look at the theory of probability.

B-2. Uncertainty - at the heart of probability

The mathematics of reliability is based on probability theory. Probability theory, in turn, deals with uncertainty. Probability had its origins in gambling. Some gamblers, hoping to improve their luck, turned to mathematicians with questions like what are the odds against rolling a six on a die, of drawing a deuce from a deck of 52 cards, or of having a tossed coin come up heads. In each case, probability can be thought of as the relative frequency with which an event will occur *in the long run*. When we assert that tossing an honest coin will result in heads (or tails) 50% of the time, we do not mean that we will necessarily toss five heads in 10 trials. We only mean that in the long run, we would expect to see 50% heads and 50% tails. Another way to look at this example is to imagine a very large number of coins being tossed simultaneously; again, we would expect 50% heads and 50% tails.

a. Events. Why is there a 50% chance of tossing a head on a given toss of a coin? It is because there are two results, or events, that can occur (assume that it is very unlikely for the coin to land on its edge) and for a balanced, honest coin, there is no reason for either event to be favored. Thus, we say the outcome is random and each event is equally likely to occur. Hence, the probability of tossing a head (or tail) is the probability one of two equally probable events occurring = 1/2 = 0.5. Now consider a die. One of six equally probable events can result from rolling a die: we can roll a one, two, three, four, five, or six. The result of any roll of a die (or of a toss of a coin) is called a discrete random variable. The probability that on any roll this random variable will assume a certain value, call it x, can be written as a function, f(x). We refer to the probabilities f(x), specified for all values of x, as values of the probability function of x. For the die and coin, the function is constant. For the coin, the function is f(x) = 0.5, where x is either a head or tail. For the die, f(x) = 1/6, where x can be any of the six values on a die.

b. Probability functions. All random events have either an underlying probability function (for discrete random variables) or an underlying probability density function (for a continuous random variable). The results of a toss of a coin or roll of a die are discrete random variables because only a finite number of outcomes are possible; hence these events have an underlying probability function. The possible height of a male American is infinite (between 5' - 8" and 6', for example, there are an infinite number of heights) and is an example of a continuous random variable. The familiar bell-shaped curve describes most natural events, such as the height of a man, intelligence quotient, errors of measurement, etc. The underlying probability density function represented by the bell-shaped curve is called normal or Gaussian. Figure B-1 shows a typical normal distribution.

c. Mean value. Note that the event corresponding to the midpoint of the curve is called the mean value. The mean value, also called the expected value, is an important property of a distribution. It is similar to an average and can be compared with the center of mass of an object. For the normal distribution, half the events lie below the mean value and half above. Thus, if the mean height of a sample of 100 male Americans is 5' -9", half the sample will be less than 69" inches tall and half will be taller. We would also expect that most men will be close to the average with only a few at the extremes (very short or very tall). In other words, the probability of a certain height decreases at each extreme and is "weighted" toward the center; hence, the shape of the curve for the normal distribution is bell-shaped.

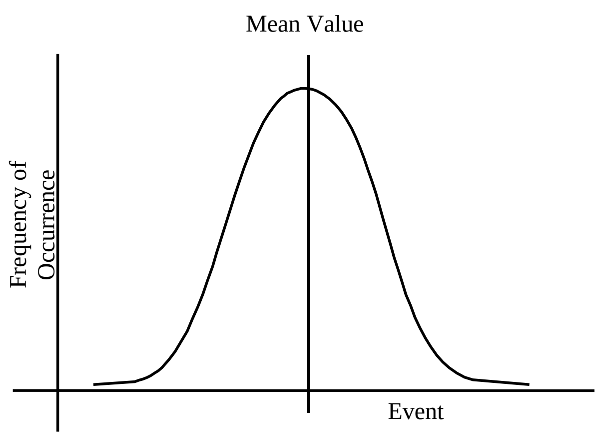

Figure B-1. Typical normal distribution curve.

d. Range of values of probability. The probability of an event can be absolutely certain (the probability of tossing either a head or a tail with an honest coin), absolutely impossible (the probability of throwing a seven with one die), or somewhere in between. Thus, a probability always can be described with equation B-1.

$$0 \leq \text{Probability} \leq 1 \qquad \text{(Equation B-1)}$$

B-3. Probability and reliability

Just as probability is associated with gambling events and naturally occurring phenomena (e.g., height of humans), it is associated with the times to failure. Since there can be an infinite number of points on a time line, time is a continuous random variable and probability density functions are used.

a. Use of the exponential distribution in reliability. Often, the underlying statistical distribution of the time to failure is assumed to be exponential. This distribution has a constant mean, λ. A reason for the popularity of the exponential distribution is that a constant failure rate is mathematically more tractable than a varying failure rate.

(1) Equation B-2 is the typical equation for reliability, assuming that the underlying failure distribution is exponential.

$$R(t) = e^{-\lambda t} \qquad \text{(Equation B-2)}$$

where:

λ is the failure rate (inverse of MTBF)
t is the length of time the product must function
e is the base of natural logarithms
R(t) is reliability over time *t*

(2) Figure B-2 shows the curve of equation B-2. The mean is not the "50-50" point, as was true for the normal distribution. Instead, it is approximately the 37-63 point. In other words, if the mean time to failure of an item is 100 hours, we expect only 37%* of the population of equipment to still be operating after 100 hours of operation. Put another way, when the time of operation equals the mean, the reliability is 37%.

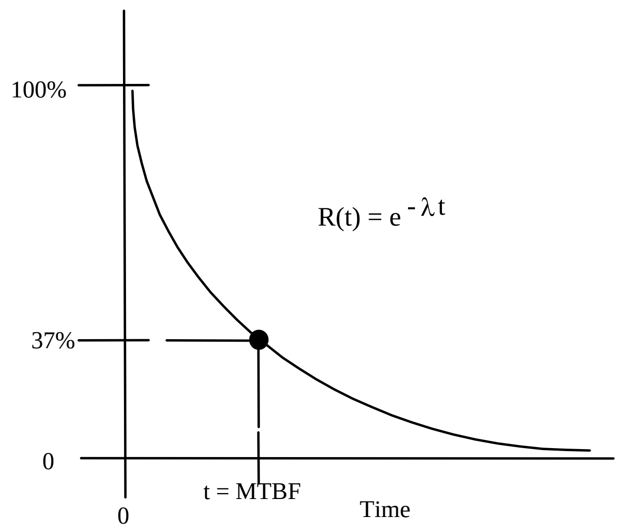

Figure B-2. Exponential curve relating reliability and time.

b. Other probability density functions used in reliability. As already stated, the popularity of the exponential distribution is its simplicity and ease of use. Unfortunately, the exponential does not accurately describe the time to failure for many parts (items that fail once and are not usually repaired, such as resistors, bearings, belts and hoses, seals, etc.). Since the failure rate, λ, is a constant, it implies that the probability of failure for a new part is the same as for an old part. But we know that many parts wear out with use or as they age. Obviously, the probability of failure increases with age for these parts. Accordingly, many other probability distributions are used to describe the time to failure of parts. The most versatile of these is the Weibull.

(1) The Weibull distribution is defined by equations B-3 and B-4 (equation B-3 is for the two-parameter Weibull and equation B-4 is for the three-parameter Weibull).

$$F(t) = 1 - e^{-(t/\theta)^{\beta}} \qquad \text{(Equation B-3)}$$

$$F(t) = 1 - e^{-[(t-t_0)/\theta]^{\beta}} \qquad \text{(Equation B-4)}$$

* If t = MTBF = $1/\lambda$, then $e^{-\lambda t} = e^{-1} = 0.367879$.

(2) The Weibull distribution can provide accurate estimates of reliability with few samples, renders simple and useful graphical results, provide clues to physics of failure, and can represent many distributions. When β is equal to 1, the Weibull is exactly equal to the exponential distribution. When, β is 3.44, the Weibull is approximately the normal distribution.

c. Applicability of the exponential to systems. Although the exponential is often inappropriate for parts (i.e., items that fail once and are discarded), it is often applicable to systems. The reason is that systems are made of many parts, each with different failure characteristics. As parts fail, they are replaced. After some time, the system has parts of varying "ages". The net result is that the times between failures of the system are exponentially distributed. This behavior of system is described by Drenick's Theorem.

B-4. Failure rate data

How do we determine the failure rate of a specific product or component? Two methods are used.

a. Method 1 - *Comparable product*. In the first method, we use failure rate data for a comparable product(s) already in use. This method has two underlying assumptions. First, the product in use is comparable to the new product. Second, the principle of transferability applies. The principle of transferability states that (failure rate) data from one product can be used to predict the reliability of a comparable product.

b. Method 2 – Testing. The other method of determining failure rate data is through testing of the product or its components. Although, theoretically, this method should be the "best" one, it has two disadvantages. First, predictions are needed long before prototypes or pre-production versions of the product are available for testing. Second, the reliability of some components is so high that the cost of testing to measure the reliability in a statistically valid manner would be prohibitive. Usually, failure rate data from comparable products are used in the early development phases of a new product and supplemented with test data when available.

B-5. Calculating reliability

If the time, t, over which a product must operate and its failure rate, λ, are known, then the reliability of the product can be calculated using equation B-2. If the information is known for individual subsystems or components, then the reliability of each can be calculated and the results used to calculate the reliability of the product. For example, consider the product represented by the reliability block diagram (RBD) in figure B-3.

a. Series calculation. Components A, B, and C are said to be in series, which means all three must operate for the product to operate. Since the components are in series, we could find the reliability of each component using equation B-2 and multiply them as follows: 0.9900 x 0.9851 x 0.9925 = 0.9680. Alternatively, the product reliability can be found by simply adding together the failure rates of the components and substituting the result in equation B-4. The product failure rate is 0.001000 + 0.001500 + 0.000750 = 0.003250. The reliability is:

$$R(t) = e^{-0.003250 \text{ x } 10} = 0.9680$$

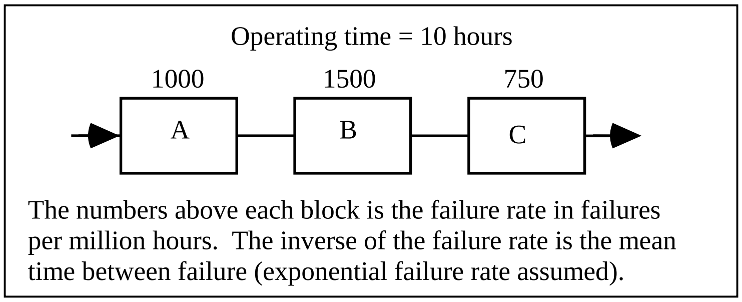

Figure B-3. Example reliability block diagram.

b. Redundancy. Now consider the RBD shown in figure B-4. The system represented by the RBD in figure B-4 has two components marked B, in a configuration referred to as redundant or parallel. Two paths of operation are possible. The paths are: A, top B, and C; and A, bottom B, and C. If either of two paths is intact, the product can operate.

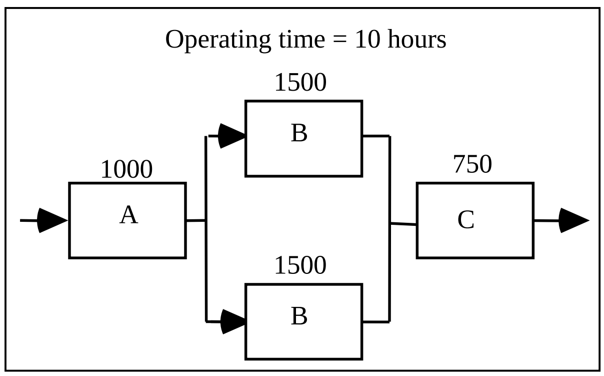

Figure B-4. RBD of a system with redundant components.

(1) The reliability of the product is most easily calculated by finding the probability of failure (1-R(t)) for each path, multiplying the probabilities of failure (which gives the probability of both paths failing), and then subtracting the result from 1. The reliability of each path was found in the previous example. Next, the probability of a path failing is found by subtracting its reliability from 1. Thus, the probability of either path failing is 1 - 0.9680 = 0.0320. The probability that both paths will fail is 0.032 x 0.032 = 0.001. Finally, the reliability of the product is 1 - 0.001 = 0.9989, about a 3.2% improvement over the series-configured product.

(2) Two components in parallel (redundant) may always be on and in operation (active redundancy) or one may be off or not in the "circuit" (standby redundancy). In the latter case, failure of the primary component must be sensed and the standby component turned on or switched into the circuit. Standby redundancy may be necessary to avoid interference between the redundant components and, if the redundant component is normally off, reduces the time over which the redundant component will be used (it's only used from the time when the primary component fails to the end of the mission). Of course, more than two components can be in parallel.

B-6. Calculating basic versus functional reliability

Reliability can be viewed from two perspectives: the effect on system performance and the effect on logistics costs.

a. Functional reliability. In the previous examples, we have seen how adding a component in parallel, i.e., redundancy, improves the system's ability to perform its function. This aspect of reliability is called functional reliability.

b. Basic reliability. Note that in figure B-4, we have added another component that has its own failure rate. If we want to calculate the total failure rate for all components, we add them. The result is 4750 failures per million operating hours (0.004750). The failure rate for the series-configured product in figure B-3 was 3,250 failures per million operating hours. Although the functional reliability of the system improved, the total failure rate for all components **increased**. This perspective of reliability is called basic or logistics reliability. Whereas functional reliability only considers failures of the function(s), logistics reliability considers all failures *because some maintenance action will be required*. Logistics reliability can be considered as either the lack of demand placed on the logistics system by failures or the ability to operate without logistics. If standby redundancy is used with the redundant component not on, the apparent failure rate of that component will be less than that of its counterpart (because the probability it will be used is less than 1 and the time it will operate less than 10 hours), but the failure rate of the switching circuits must now be considered. The logistics reliability for the system with active redundancy in figure B-4 is 0.9536 over a ten-hour period.