USAAVSCOM TECHNICAL REPORT 77-16

POCKET HANDBOOK ON RELIABILITY

James K. Byers

Henry A. Weibe

COMPCO COMPANY

P.O. Box 632

Rolla, Missouri 65401

Final Report

Prepared for

U.S. ARMY AVIATION SYSTEMS COMMAND

R&M Division, Product Assurance Directorate

P.O. Box 209

St. Louis, MO 63166

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1 REPORT NUMBER: USAAVSCOM Technical Report 77-16	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)*: POCKET HANDBOOK ON RELIABILITY.	5. TYPE OF REPORT & PERIOD COVERED: Final Report.
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): James K. Byers, Henry A. Weibe	8. CONTRACT OR GRANT NUMBER(s): DAAJ01-74-C-1091 [redacted]
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Compco Company, P. O. Box 652, Rolla, Missouri 65401	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Directorate for Product Assurance, US Army Aviation Systems Command, St. Louis, Missouri 63166	12. REPORT DATE: Sep[redacted]75 / 13. NUMBER OF PAGES: 73
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*: [illegible] TR-77-16	15. SECURITY CLASS *(of this report)*: Unclassified / 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for Public Release; Distribution Unlimited

17 DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18 SUPPLEMENTARY NOTES

19 KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Reliability definitions, Series and parallel systems, exponential distributions, Weibull distribution, extimating reliability, confidence intervals, reliability growth, O. C. curves, Bayesian analysis.

20 ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

This handbook is intended as an educational tool aimed at engineers and managers who do work directly in reliability and who must communicate with other individuals who do likewise.

(Continued) —>

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

It attempts to explain the basic techniques that are used in reliability analysis and how these techniques are used to solve real problems.

The handbook is built around six important subject areas within reliability. These subject areas are; (1) Reliability of a single component versus multiple components, (2) The exponential and Weibull models, (3) Estimating reliability using test data, (4) Techniques for computing confidence intervals, (5) O. C. curves in reliability analysis, (6) Bayesian methods in reliability analysis.

The pamphlet is structured to stress the techniques used within the subject areas. Examples are frequently used to show how these techniques are applied to solve real problems. The emphasis is on the practical application rather than on the theoretical discourse.

Pocket Handbook on Reliability

R & M DIVISION
DIRECTORATE FOR PRODUCT ASSURANCE
U.S. Army Aviation Systems Command

Forward

The Army Aviation Systems Command has the responsibility for assuring the reliability and combat readiness of the total fleet of Army aircraft. This responsibility can only be carried out through the cooperation and understanding of a large number of people--both inside and outside AVSCOM.

This handbook has been developed to promote this understanding by presenting the techniques that must be conscientiously applied to assure a reliable product. Although some statistical and mathematical background is required to understand the techniques, the handbook has been designed for engineers and scientists who have not frequently been exposed to these techniques. Examples and illustrations are frequently used for ease of reading.

It is hoped the handbook will be a good introduction for those not familiar with reliability and a good refresher for those who are currently working in the area.

LEWIS NERI, CHIEF

RELIABILITY AND MAINTAINABILITY DIVISION
DIRECTORATE FOR PRODUCT ASSURANCE

Acknowledgement

This pamphlet was prepared by COMPCO Company under the technical direction of the R & M Division, Product Assurance Directorate, AMSAV-LR, AVSCOM (Contract Number DAAJOI-74-C-1091(P1G)). At COMPCO, the project was conducted under the joint direction of Drs. Henry Wiebe and Jim Byers. The illustrations were done by Bill Juedemann.

On the part of the government, the project was under the technical cognizance of Mr. Lewis Neri, Chief of Reliability and Maintainability Division, Directorate for Product Assurance, U.S. Army Aviation Systems Command. COMPCO is grateful to Mr. Edward Hollman, Director of Product Assurance for overall guidance and direction concerning Army needs and requirements. Mr. Elmer Lueckerath provided valuable technical review of the handbook.

Respectfully submitted,
COMPCO Company

Henry A. Wiebe
General Partner

James K. Byers
General Partner

Preface

This handbook is intended as an educational device aimed at engineers and scientists who do not work directly in reliability but must deal with individuals who do. It attempts to communicate the techniques that are used in reliability analysis and how these techniques are used to solve real problems.

The handbook is built around five important subject areas within reliability. These subject areas are:

1. Reliability of a Single Component versus Multiple Components
2. The Exponential and Weibull Models
3. Estimating Reliability using Test Data
4. O.C. Curves in Reliability Analysis
5. Bayesian Methods in Reliability Analysis

The pamphlet is structured to stress the techniques used within the subject areas. Examples are frequently used to show how these techniques are applied to real problems.

Table of Contents

Section 1.0
Introduction to Reliability

1.1 Why Bother with Reliability

The measure of an equipment's reliability is the frequency with which failures occur over a specified period of time. In the past few years, the concepts of reliability have become increasingly important and have become a primary concern in the development of most large weapon systems.

The reason for this concern and the increased emphasis on reliability is found in the technological revolution which mankind has been experiencing during the last several decades. This revolution has been accelerated by the various wars, dating back to the 2nd World War, and the stress on military preparedness since that time. In addition to accelerating technological developments, armed conflict dramatically emphasizes the consequences of unreliability. These consequences range from minor inconveniences to matters that can affect national security.

A somewhat disturbing fact is that the problem of avoiding these consequences can only become more severe as time progresses. Highly refined and sophisticated equipment is a necessity in order to accomplish the missions facing today's military forces. The ability to respond to military situations can easily be compromised by potential equipment failures if reliability is not held at a high level.

Reduced operating budgets serve to further compound the problem of equipment readiness by limiting the number of backup systems and units that are available to respond when needed. The fact that back up systems are often not available means that the primary units **must** function properly. Sound reliability and maintainability practices can insure that existing systems are capable of functioning properly. More importantly, perhaps, is the fact that sound

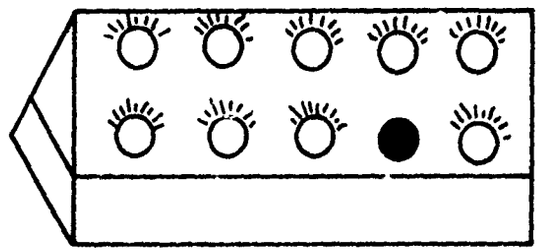
R&M practices will also insure that a minimum number of dollars will be expended to achieve a required level of operational readiness. For example, reliability analyses can be used to determine whether it is better from a cost effectiveness viewpoint to use redundant systems or to upgrade the reliability of the primary unit in order to achieve a given level of operational capability. Sound R&M practices will also insure that limited operations and maintenance dollars are spent in the correct areas to assure that maximum benefit is obtained from the dollars available. For example, proper reliability analyses can show which problem areas are the ones in real need of attention from an operational capability viewpoint and which ones are of less critical nature. The net effect of conscientiously applying adequate R&M procedures is to bring down the overall acquisition and operational costs and increase the operational readiness of most systems.

The inescapable conclusion that one must reach is that the probability of failure must be carefully controlled for the highly complex equipment required by today's military forces to function properly when called upon. The only way this can be accomplished is for a great deal of emphasis to be placed in the area of reliability and maintainability during the equipment development and operation phases.

1.2 How to Calculate the Reliability of a System

Reliability can be defined in its simplest form as "the probability of successful operation". A number representing this probability can be obtained from test data and, again in its simplest form, is the ratio of the number of components surviving a test to the number of components present at the beginning of the test.

As a hypothetical example: Ten helicopter warning lights could be placed on a test stand, turned on, and observed for



$$\text{Rel.} = \frac{9}{10} = 0.9$$

Figure 1-1 Test Stand

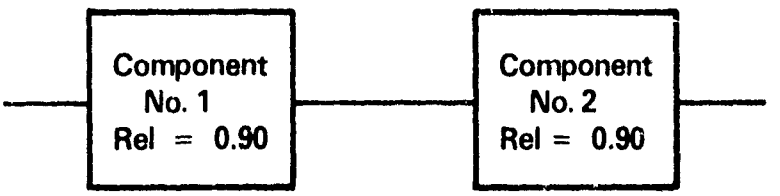
a one week period. If at the end of this time there were nine warning lights still burning, one could conclude that the reliability of this device was 0.90. (See Figure 1-1). That is--the probability of successful operation for the warning light tested is 0.90.

A complete definition of reliability is somewhat more complex than the one given above and is stated in MIL-STD-721B as follows: Reliability is "The probability that an item will perform its intended function for a specified interval under stated conditions". This definition indicates that matters may not be quite as simple as they were stated in the first paragraph of this section. For example: The warning lights of Figure 1-1 will eventually be used in a hostile environment where they will be subject to stress from vibrations that are always present on a helicopter. To yield proper results, the test would have to be conducted to account for this hostile environment. Otherwise, the reliability calculated would not have any meaning. The time over which the test was conducted is also important since the light must perform "---for a specified interval---". This means that our previously calculated reliability of 0.90 must be modified to reflect the actual period of time over which the light must function in actual use.

1.2.1 Probability Laws and their use in Reliability Calculations

Further complications in the determination of reliability are introduced when we start talking about system reliability instead of component reliability. A system is made up of several components of which one or more must work before the system can function. Figure 1-2 indicates a situation where both components must be working in order for the system to function. The components are said to be connected in series and when one component fails, the entire system fails. In this case we are interested in the reliability of the entire system and not the reliability of the individual components. We can calculate the system reliability by multiplying the reliabilities of the individual components together as follows:

$$\begin{aligned} \text{Reliability of System} &= (\text{Rel. Component \#1}) \times (\text{Rel. Component \#2}) \\ &= 0.90 \times 0.90 \\ &= 0.81. \end{aligned}$$



System Reliability = 0.90 x 0.90 = 0.81

Figure 1-2 Series Components

The above operation made use of a probability law called "The Law of Multiplication". This law can be stated by saying "that if two or more events are independent, the probability that all events will occur is given by the product of their respective probabilities." In the example above, our event consisted of a single component working. The product of the two events yielded the probability that the system would function (i.e., both components would work at the same time).

A more realistic example is shown in Figure 1-3 where a portion of the aft section, anti-torque controls installation for the AH-1G attack helicopter is illustrated. A description of this series subsystem is given in Table 1-1 below along with a list of hypothetical component reliabilities.

Table 1-1

COMPONENTS OF ANTI-TORQUE CONTROL SUBSYSTEM FOR AH-1G

COMPONENT NUMBER	COMPONENT DESCRIPTION	HYPOTHETICAL COMPONENT RELIABILITY
1	Control Rod #1	0.98
2	Pivot Arm	0.99
3	Pivot Bolt #1	0.92
4	Pivot Bolt #2	0.92
5	Pivot Bolt #3	0.92
6	Control Rod #2	0.98
7	Quadrant Assembly	0.97
8	Quadrant Connecting Bolt	0.92
9	Quadrant Pivot Bolt	0.91
10	Quadrant Cable Connecting Pin	0.90

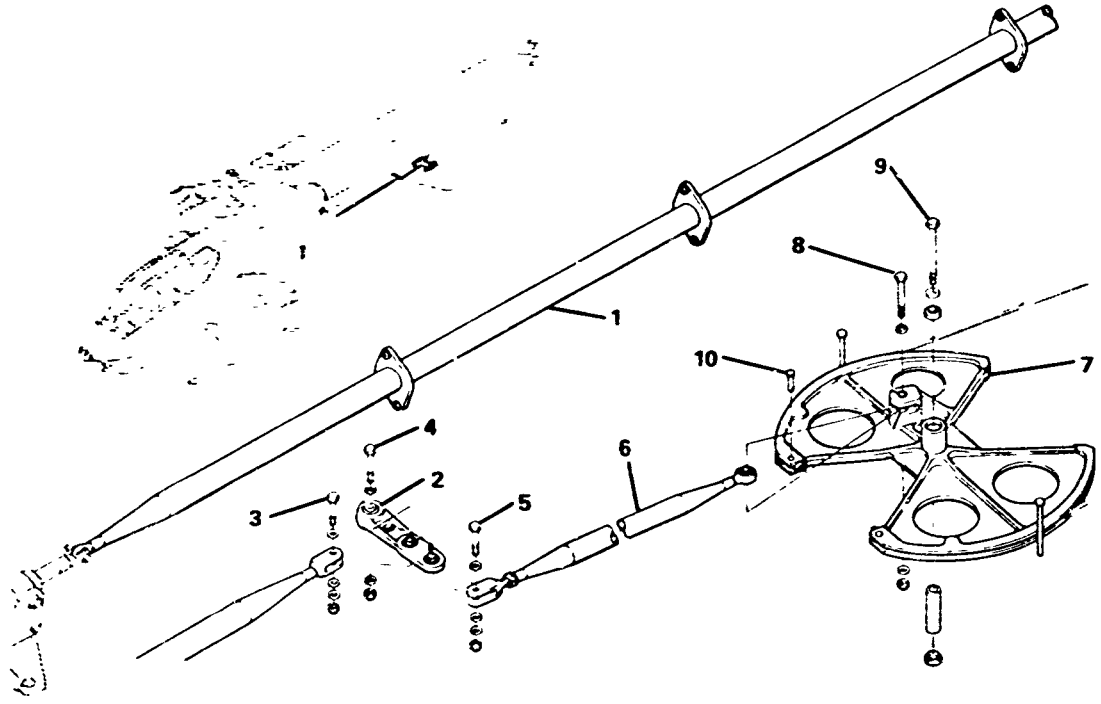


Figure 1-3 Aft Section, Anti-torque Controls Installation

The reliability of the anti-torque control subsystem is calculated as follows:

Rel = .98 x .99 x .92 x .92 x .92 x .98 x .97 x .92 x .91 x .90

Rel = 0.54.

It is interesting to note the low overall reliability caused by using multiple components in the above example. The impact of using multiple components will be discussed in detail later.

A different type of system arrangement utilizing two components is shown in Figure 1-4 below. This system requires that only one component be functional. These components represent a parallel or redundant system where one can serve as a backup unit for the other in case of a single failure.

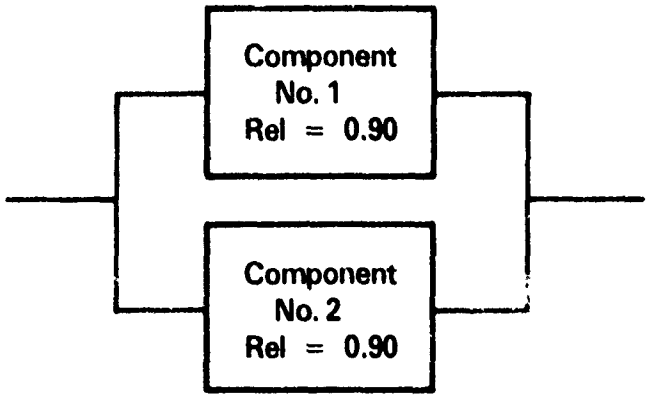


Figure 1-4 Parallel Components

To determine the probability that either component No. 1 or component No. 2 or both will operate we simply add the two

probabilities and then subtract their product as follows:

$$\text{Rel of Sys} = (\text{Rel of Comp \#1}) + (\text{Rel of Comp \#2}) - (\text{Rel Comp \#1}) \times (\text{Rel Comp \#2})$$

$$= 0.90 + 0.90 - (0.90) \times (0.90)$$

$$= 1.80 - .81$$

$$= 0.99.$$

The above calculations were based on the General Law of Addition which can be stated by saying that "if two events can occur, the probability that either one or both will occur is given by the sum of their individual probabilities less the product of their individuai probabilities." Our event again consisted of a single component working. The system was said to function as long as either one or both components were working.

An important point illustrated by the above examples was the fact that system configuration can have a major impact on the overall system reliability. Parallel (or redundant) systems such as the one illustrated in Figure 1-4 are often used where high reliability is required. For the anti-torque control system of Figure 1-3, this would mean that a back up system (or back up components) would be necessary to provide control in the event of any failure in the primary system.

One of the basic concepts of reliability analysis is the fact that all systems, no matter how complex, can be reduced to a simple series system. For example: The illustrations of Figure 1-2 and Figure 1-4 can be combined to yield the system in Figure 1-5.

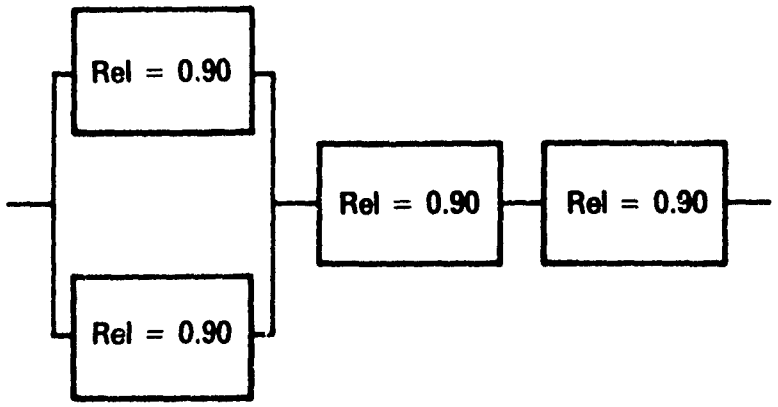
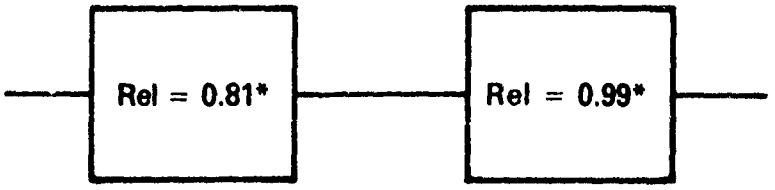


Figure 1-5 Combination Parallel and Series System

Using the results of our previous calculations we can reduce this system to the one of Figure 1-6 which has a system reliability of 0.80. The reliability of the parallel portion of Figure 1-5 was previously calculated to be 0.99 while the series portion of Figure 1-5 was previously calculated to be 0.81. These reliabilities are now used to represent the new two component series system of Figure 1-6.



*Results obtained from Figures 1-2 and 1-4

Figure 1-6 Condensed Series System

1.3 The Meaning of Chance Failures

Equipment can fail from a variety of causes. For the purpose of reliability studies, these causes have been put into three catagories which are:

1) Early or burn in failures
2) Chance failures
3) Wearout failures.

Early failures occur during the initial phases of an equipment's life and are normally the result of substandard materials being used or a malfunction in the manufacturing process. When these mistakes are not caught by quality control inspections, an early failure is likely to result. Early failures can be eliminated by a "burn in" period during which time the equipment is operated at stress levels closely approximating the intended actual operating conditions. The equipment is then released for actual use only when it has successfully passed through the "burn in" period.

Wearout failures occur at the end of an equipment's useful life and are a result of equipment deterioration due to age or use. For example; transmission bearings will eventually wear out and fail regardless of how well they are made. Early failures can be postponed and the useful life of equipment extended by good maintenance practices. The only way to prevent failure due to wearout is to replace or repair the deteriorating component before it fails.

Chance failures are those failures that result from strictly random or chance causes. They can not be eliminated by either lengthy burn in periods or good preventative maintenance practices. Equipment is designed to operate under certain conditions and up to certain stress levels. When these stress levels are exceeded due to random unforseen or unknown events, a chance failure will occur. While reliability theory and practice is concerned with all three types of failures, its primary concern is with chance

failures as these occur during the useful life of the equipment.

The time when a chance failure will occur can not be predicted; however, the likelihood or probability that one will occur during a given period of time within the useful life can be determined by analyzing the equipment design. If the probability of chance failure is too great, either design changes must be introduced or the operating environment made less severe.

The failure rates discussed above are further illustrated by the Life Characteristic Curve, commonly called the "Bathtub Curve" shown in Figure 1-7.

Figure 1-7 shows failure rate as a function of age and clearly indicates the time periods when the three types of failures can occur. The chance failure rate is approximated

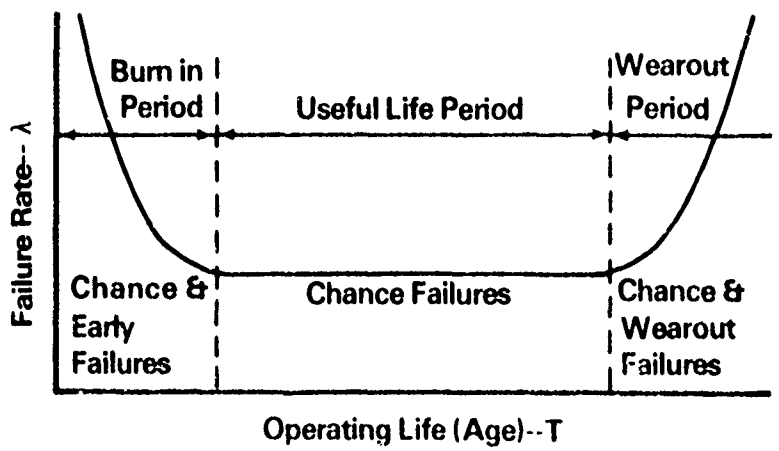


Figure 1-7 Failure Rate as a Function of Equipment Age

by the exponential distribution. Reliability during the useful life is calculated by using the equation:

$$R(t) = e^{-\lambda t}$$

where "t" is the elapsed time period during which the equipment must operate and "λ" is the chance failure rate during the useful life period. As a specific example let $\lambda = .001$ failures per hour and $t = 1$ hour, then

$$R(t) = e^{\lambda t}$$

$$= e^{-.001(1)}$$

$$= e^{-.001}$$

$$= .999.$$

In other words, there is a 99.9% chance that the above equipment will operate successfully for one hour.

1.4 Failure Definition Problems

According to MIL-STD-721B a failure is the inability of an item to perform within previously specified limits. This means that system performance levels must be clearly defined before failures can be identified. The task of defining system performance levels is not completely straight foreward. A complete analysis of complex systems such as Army helicopters normally requires that these limits be specified at multiple levels. For helicopters, these performance levels are those affecting.

A. Flight safety
B. Mission success
C. System unscheduled maintenance.

The reason for specification of multiple performance levels is that the consequences of a failure and thus the seriousness of the problem can be vastly different from level to level. For example, a failure affecting flight safety may well mean loss of lives while a failure affecting system unscheduled maintenance might mean only an inconvenient delay in a training flight.

Examination of the above catagories reveals that it is possible for a malfunctioning equipment item to be classified as an equipment failure in one of the lower ranking catagories above and yet not be classified as a failure in a higher ranking catagory. For example, carbon seal leakage can require an unscheduled maintenance action to replace the seal and thus be classified as a failure in terms of not meeting the operational limits specified in Catagory C above. However, a leaking seal may pose no immediate threat to flight safety and thus not be classified as a failure with respect to the performance limits imposed in Catagory A.

Proper failure definition is extremely important if accurate data is to be collected for use in monitoring system performance and determining system reliability. Design decisions affecting system safety can not be made correctly if the analysis is based on a mixture of reported failures affecting only mission success and unscheduled maintenance actions. Failure reporting forms must be designed in such a way as to remove the ambiguity associated with failure classification and reporting.

Section 2.0

Reliability of Single Component Versus Multiple Components

2.1 Series Components

The reliability of the total system is of prime importance in reliability analysis. A system usually consists of many different components. For example, large complex systems such as aircraft, helicopters and rockets have several thousands of components while a small appliance may have less than one hundred. Components can be structured in one of two ways, either in series or in parallel. Briefly, if components are in series, then all of the components must operate successfully for the system to operate. On the other hand, if components are in parallel only one of the components must operate for the system to continue functioning. This is also referred to as redundancy. We will study both of these configurations to see how they affect the system reliability.

System reliabilities are calculated by means of the laws of probability. To apply these laws to systems, we must have some knowledge of the reliabilities of its parts, since they affect the reliability of the system. Component reliabilities are derived from tests which yield information about failure rates. When a new component is designed and built, no measure of the electrical, mechanical, chemical, or structural properties can tell us the reliability of the component. This is obtained only through testing the system in a realistic environment.

The problem that is of interest is the manner in which the reliability changes as the number of components in series increases. Thus we must be able to compute the reliabilities of the components grouped together in a series manner.

Reliability calculations for a group of components are based on two important operations: (1) as precise as possible

a measurement of the reliability of the components used in the system environment, and (2) the calculation of the reliability of the series system. The rule that is used for a group of series components is the product of the reliabilities of the individual components.

As a specific example consider the rotor blade on a helicopter as one component. The rotor blade is part of the rotor subsystem which is part of the power transmission system. The power transmission system is just one of the systems that make up the helicopter. The reliability calculations for this hypothetical case are given in Figure 2-1. For illustrative purposes all components were considered to have an identical reliability of 0.99995. This is equivalent to having only 5 failures in 100,000 flights.

Note how the reliability decreases from a component reliability of 0.99995 to a reliability for the helicopter with 5000 components of 0.779. This is an average of 20 failures every 100 flights. This decrease is even more pronounced for lower component reliability. For example, with a component reliability of 0.99, the reliability of a helicopter with 5000 components would be practically zero.

Figure 2-2 is a graphic portrayal of how the reliability of series components changes with the reliability of the individual component when each component is identical. This concept illustrates exactly why highly reliable systems are much more expensive and require extensive testing to verify the high reliability. A low reliability can be verified with a small amount of testing; but, unfortunately, as has been illustrated, this is simply not possible for high reliability systems.

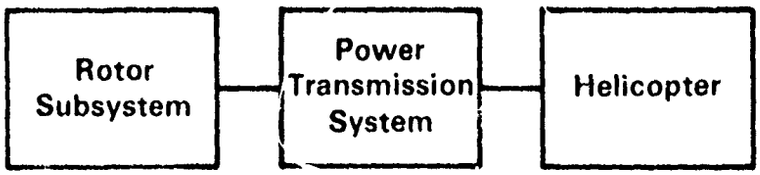
2.2 Redundancy

When very high system reliabilities are required, the designer must often duplicate components, and sometimes whole subsystems, to meet the reliability goals. When this

SERIES FORMULA: The reliability of a group of series components is the product of the reliabilities of the individual components. An example for a helicopter is shown below.

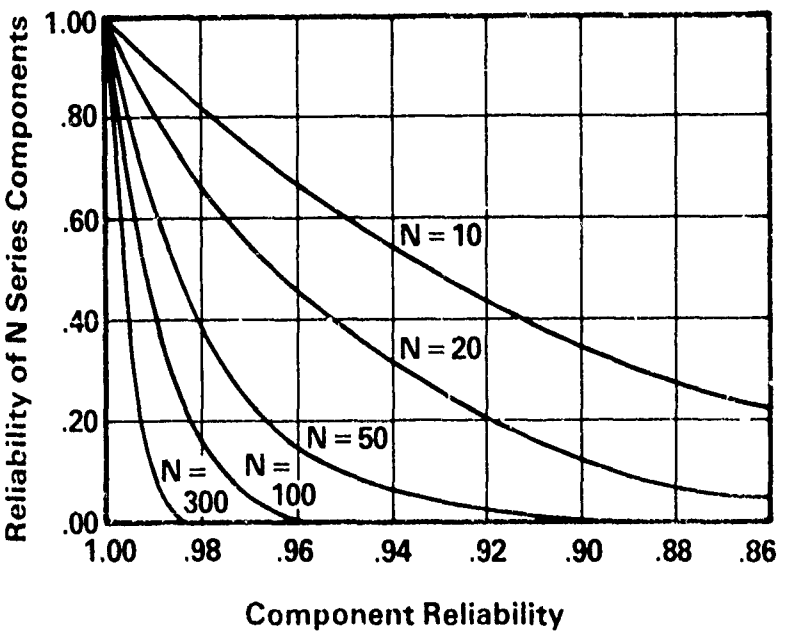
$$\text{Reliability} = (0.99995)^{5000}$$

$$= 0.779$$



	Group		
	Rotor Subsystem	Power Transmission System	Helicopter
Number of Components	10	500	5,000
Component Reliability	.99995	.99995	.99995
Group Reliability	.9995	.975	.779

Figure 2 1 Effect on the Reliability of Increasing the Number of Components in Series.



Typical Values from
the Above Curve

Component Reliability	Reliability of N Series Components				
	N = 10	20	50	100	300
.99	.904	.818	.605	.366	.049
.95	.599	.358	.077	.006	.000

Figure 2-2 Reliability of N Series Components When Each Component is Identical

type of design is used, the components are said to be redundant or parallel. Just as we saw earlier how the reliability for series components decreases as the number of components increases, the inverse is true here. Redundant components can dramatically increase the reliability of a system. However this increase in reliability is at the expense of such factors as cost, weight and space.

When redundant components are being analyzed, the term "unreliability" is frequently used. This is because the calculations are easier to perform using the unreliability of a component. The unreliability is simply defined as

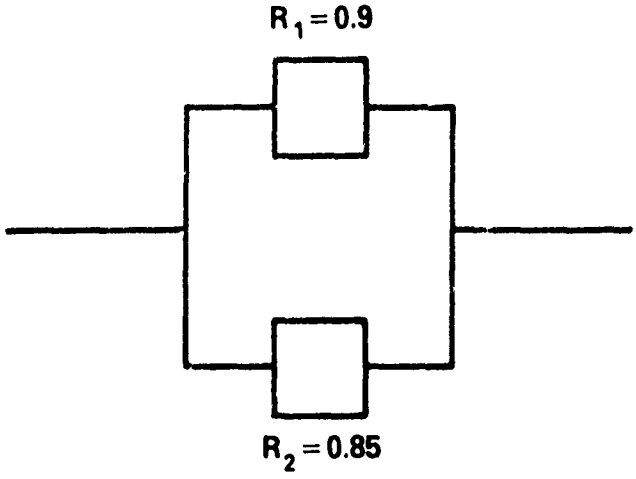
Unreliability = 1 - Reliability.

As a specific example consider the two parallel components shown in Figure 2-3. With reliabilities of only $R_1 = 0.9$ and $R_2 = 0.85$, we were able to obtain a combined reliability of 0.985 by putting them in parallel.

The improvement in reliability achieved by operating components in parallel can be further illustrated by referring to Figure 2-4. These curves show the reliability of the parallel group compared to the reliability of individual components for one, two, three, or five components in parallel. Each component in the parallel group is considered to be identical.

From this graph we see that a significant gain in reliability is obtained from redundancy. To cite a few examples from the curve, if the reliability of one component is 0.9, the reliability of two such components in parallel is 0.99. The reliability of 3 such components in parallel is 0.999. Which means that on an average only one time out of a thousand will all three components fail to operate.

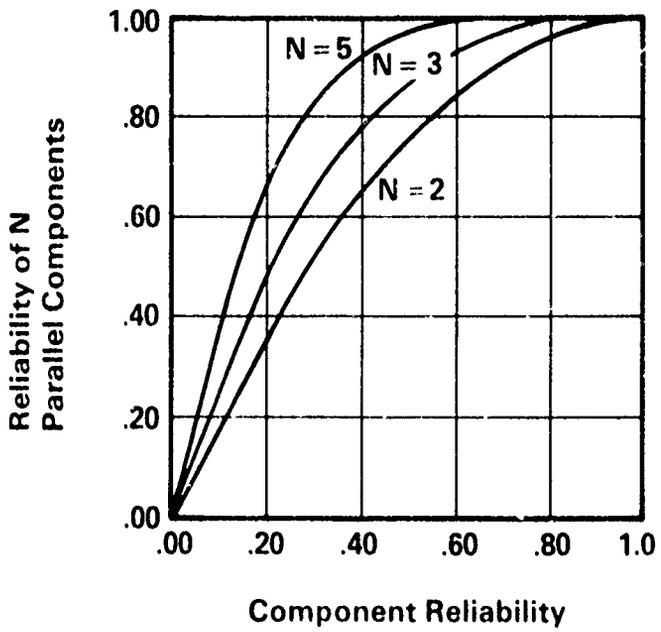
PARALLEL FORMULA: The reliability of a group of parallel components is one minus the unreliability of the group. The unreliability is the product of the unreliabilities of the individual components.



UNRELIABILITY OF GROUP $= (1 - R_1) \times (1 - R_2)$
$= (1 - .9) \times (1 - .85)$
$= (.1) \times (.15)$
$= 0.015$

RELIABILITY OF GROUP $=$ 1- UNRELIABILITY
$=$ 1- .015
$=$ 0.985

Figure 2-3 Reliability of Two Parallel Components



Component Reliability	Reliability of N Parallel Components		
	N = 2	3	5
.9	.990	.999	1.000
.7	.910	.973	.998

Figure 2-4 Reliability of N Parallel Components when Each Component is Identical

Section 3.0

The Exponential and Weibull Models

3.1 Introduction

The previous discussions on reliability analysis have assumed that the component reliability was known and we were only interested in using this component reliability to compute the system reliability. We are now going to see how these component reliabilities are computed. Specifically the important role that the exponential and Weibull density functions play in reliability analysis will be illustrated. Other density functions that are occasionally used in reliability analysis are also presented.

3.2 Exponential Distribution

When a component is subject only to failures which occur at random intervals, and the expected number of failures is the same for equally long operating periods, its reliability is defined by the exponential equation

$$R = e^{-t/\Theta}$$

or in terms of the failure rate λ

$$R = e^{-\lambda t}$$

where

e = the base of the natural logarithm (2.71828)

Θ = a parameter called the mean time between failure usually referred to as MTBF,

t = the operating time for which we want to know the reliability R of the component,

$\lambda = 1/\Theta =$ **failure rate for the component.**

This equation is applicable as long as the component is in its useful life. The useful life of a component is considered to be the time after which burn-in failures no longer exist and wear-out has not yet begun. In highly complex electro/mechanical systems, such as helicopters, the system failure rate is effectively constant over the useful life period regardless of the failure pattern of individual parts. This constant failure rate is the criteria for assuming the exponential density function for the helicopter system as a whole.

An important point to note about this equation is that the reliability R is a function of operating time for the component and the mean time between failures. This relationship can be seen by inspecting Figure 3-1. From this figure it can be seen that for small operating times, the reliability is high. The reliability rapidly decreases as the operating time increases. For an operating time equal to the MTBF the reliability is only .368. This can be proven as follows:

$$R = e^{-t/\Theta}$$

when $t = \Theta$

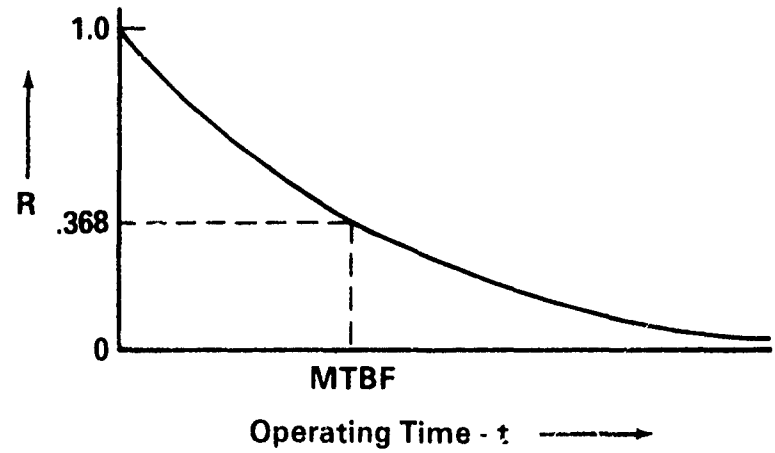
$$R = e^{-\Theta/\Theta}$$

$$= e^{-1}$$

$$= 1/e$$

$$= .368.$$

The MTBF is a measure of the average time until a component fails. If a large number of identical components



Typical values for the reliability are given below.

t	¼ MTBF	½ MTBF	¾ MTBF	MTBF
R	.778	.606	.472	.368

Figure 3-1 Relibility Versus Operating Time for Exponential Time Between Failures

were put on test and operated until they all failed, the MTBF would be obtained by adding up the operating times until failure and dividing by the number of items tested. As implied here, the MTBF is estimated from test data.

Note that the MTBF is not a function of the operating time. This is a unique and important property of the exponential equation. The end result of this property is that a component has an equal chance of survival for periods of equal length throughout its useful life.

To illustrate the important fact of an equal chance of survival for periods of equal operating time within the useful life, let us assume that a debugged component with a 500 hour useful life has a constant failure rate, λ, of 0.0001. Its

reliability for any 20 hours operation within these 500 hours is

$$R = e^{-\lambda t}$$

$$R = e^{-.0001(20)}$$

$$= e^{-.002}$$

$$= .998.$$

Thus the probability of the component surviving during the first 20 hours of the 500 hours is 0.998, and the probability of the component surviving during the last 20 hours of the 500 hours is also 0.998. However, if this component should continue operation beyond a total of 500 hours, wearout will begin to play a role and the reliability of the component will begin decreasing.

The reciprocal of the MTBF is an important and often used value. It is commonly referred to as the failure rate. By definition then, the failure rate is

$$\lambda = 1/MTBF.$$

Thus the reliability function, R, can also be written in the form

$$R = e^{-\lambda t}.$$

Thus far we have discussed when the exponential distribution applies in reliability analysis but we have really not stated that it is the time between failures that is exponentially distributed. The curve that describes how the time between failures is distributed is called a probability density function. This density function is shown pictorially in Figure 3-2. The reliability for an operating time of t_1, is then the probability that the component will not fail prior to time

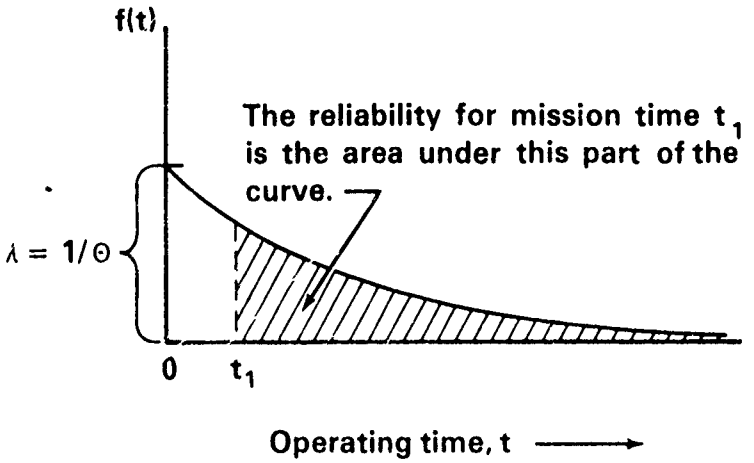


Figure 3-2 Exponential Density Function

t_1, which is the area under the curve between $t = t_1$ and $t = \infty$. As a specific example, let MTBF = 1000 hours and the mission time t_1 = 1 hour. Then

$$R = e^{-t/MTBF}$$

$$R = e^{-1/1000}$$

$$= e^{-.001}$$

$$= .999.$$

The complement of the reliability function, the probability of failing between t = 0 and t = 1 hour or the unreliability, Q is

$$Q = 1-R$$

$$= 1-.999$$

$$= .001.$$

When components are combined in a series manner, the calculations are not very complex, particularly if the operating times are the same for each component. Consider the block diagram in Figure 3-3, which shows 2 components in series. Let the operating time for each component be 1 hour and the failure rates are as shown on the block diagram. The reliability calculations for this system are shown on the figure. The resultant value for the reliability of the system is $R_S = 0.9987$ Note that during the intermediate steps of the calculations, we added the failure rates for each of the components to obtain a combined failure rate for the system. This can be done for any system when the components have exponential times to failure. Thus a general expression for the system failure rate is

$$\lambda_S = \lambda_1 + \lambda_2 + \ldots + \lambda_n .$$

When each component has identical operating time t, the reliability formula reduces to

$$R_S = e^{-\lambda_S t} .$$

When components are arranged in parallel for redundancy, the calculations are done using the unreliability of each component. The exact (or classical) equations used for obtaining the system unreliability for components arranged in parallel for redundancy are

$$Q_S = \prod_{i=1}^{n} (1-e^{-\lambda t_i})$$

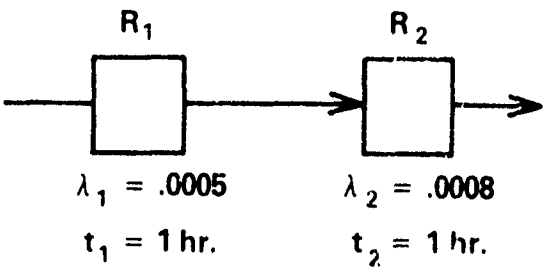
and

$$R_S = 1-Q_S .$$

The first equation states that for n units in parallel, the unreliability of the system is the product of the unreliabilities

Series Formula for Exponential Time Between Failures is

$$R_S = e^{-\sum_{i=1}^{n} \lambda_i t_i}$$



$$R_S = e^{-[\lambda_1 t_1 + \lambda_2 t_2]}$$

$$= e^{-[.0005(1) + .0008(1)]}$$

$$= e^{-[.0005 + .0008](1)}$$

$$= e^{.0013(1)}$$

$$= 0.9987$$

Figure 3-3 Reliability Calculations for Two Series Components with Exponential Times Between Failure

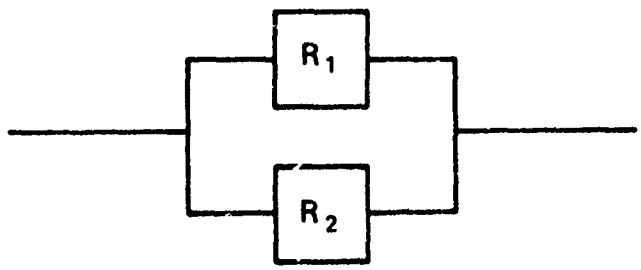
Parrallel Formula for Components with Exponential Time Between Failures Is

$$Q_s = \prod_{i=1}^{n} (1-e^{-\lambda_i t_i})$$

$$R_s = 1-Q_s$$

$\lambda_1 = .01$

$t_1 = 1$ hr.



$\lambda_2 = .05$

$t_2 = 1$ Hr.

$$Q_s = Q_1 Q_2$$
$$= (1-e^{-\lambda_1 t_1}) \quad (1-e^{-\lambda_2 t_2})$$
$$= (1-e^{-.01(1)}) \quad (1-e^{-.05(1)})$$
$$= (1-0.9900) \quad (1-0.9512)$$
$$= (.01) \quad (.0488)$$
$$= 0.000488$$
$$R_s = 1- Q_s$$
$$= 1-0.000488$$
$$= 0.999512$$

Figure 3-4 Reliability Calculations for Two Parallel Components with Exponential Times Between Failures

of each of the individual components. Knowing the unreliability of the system, the reliability is obtained from the second equation. Components operating in parallel are often identical. The above equations then simplify to

$$Q_s = Q^n$$

and

$$R_s = 1 - Q^n.$$

When the failure rates are small, say .01 or less, an approximation is sometimes used for two components in parallel. The system failure rate, λ_s, is simply the product of the two individual rates, or

$$\lambda_s = \lambda_1 \lambda_2$$

and again,

$$R_s = e^{-\lambda_s t}.$$

Sample calculations for two components in parallel using the exact equations are given in Figure 3-4. This example illustrates mathematically how paralleling two components of mediocre reliability produces a system of very high reliability. This assumes, of course, that when either of the components fail the other component takes over the function. For the system to fail, both components must fail.

3.3 Weibull Distribution

Although life testing of components during the period of useful life is generally based on the exponential density function, we have already pointed out that the failure rate of a component may not be constant throughout the period

under investigation. In some instances the main purpose of life testing may be that of determining the time to wearout failure rather than chance failure. In such cases the exponential density function generally does not apply, and it is necessary to substitute a more general density function such as the Weibull. The Weibull density function is particularly useful because it can be applied to all three of the phases of the life characteristic curve. The exponential density function is a special case of the Weibull.

The equation associated with the Weibull density function is given by

$$f(t) = a\beta t^{\beta-1} e^{-a t^{\beta}} \qquad t>0, a>0, \beta>0$$

where

f(t) = Weibull density function

t = the operating time for which we want to know the reliability R(t) of the component,

β = parameter of the density function usually referred to as the shape parameter,

a = parameter of the density function usually referred to as the scale parameter.

The range of shapes that a graph of the Weibull density function can take on is very broad, depending primarily on the value of the shape parameter β. Figure 3-5 shows three of these curves corresponding to $\beta = 1/2$, $\beta = 1$, and $\beta = 2$. For β less than 1, the Weibull curve is asymptotic to both axes and highly skewed to the right. For β equal to 1, it is identical to the exponential density function, and for β greater than 1, it is "bell-shaped" but skewed. The amount the curve is spread out along the abscissa depends on the

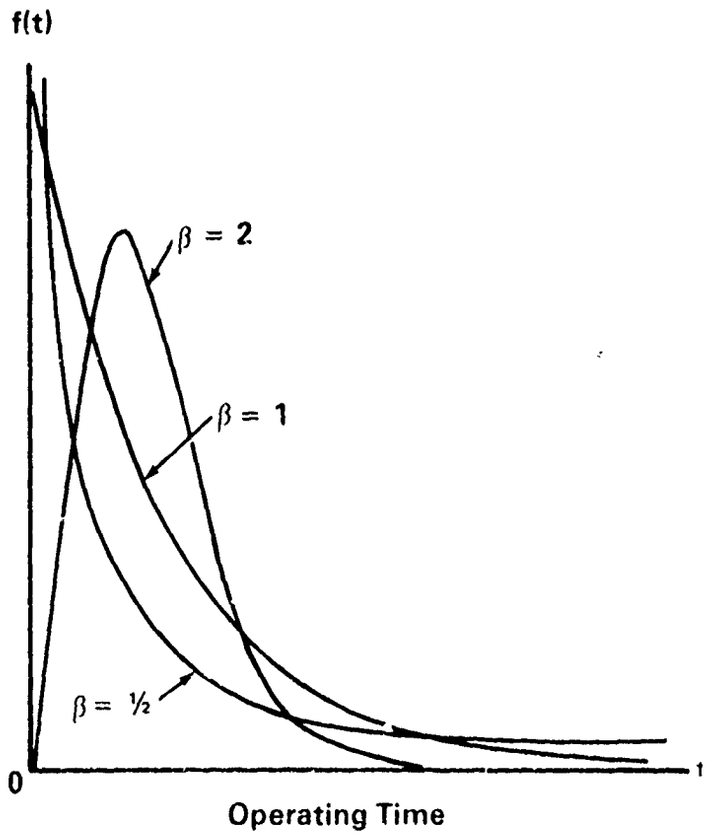


Figure 3-5 Weibull Density Function for $a = 1$

other parameter a, thus the reason for it being called the scale parameter.

The reliability formula associated with the Weibull density function is

$$R(t) = e^{-at^{\beta}}.$$

To use this formula, one must estimate the values of a and β from test results. Estimates of the parameters a and β

of the Weibull density function are computationally difficult to obtain without the use of a computer. There exist analytical methods for estimating these parameters, but they involve the solution of a system of transcendental equations. A more rapid and commonly used method is based on a graphical technique. Although none of these methods for estimating the parameters is presented here, an example problem is given that illustrates the use of the Weibull distribution.

For the example suppose that a sample of 50 components is put on life test for 500 hours. The times to failure of the 10 components that failed during the test are as follows: 10, 45, 140, 190, 220, 250, 320, 380, 440 and 480. Using the graphical method the estimates of a and β for this example are

$$a = .0016$$

$$\beta = .775.$$

Thus the reliability formula is

$$R = e^{-.0016\, t^{.775}}.$$

It should be mentioned that even though only ten of the fifty components failed, the fact that there were fifty components tested was used in computing a and β. None of the data is discarded.

Assume now that we want to know the reliability of this component for a 2 hour mission. Using the reliability equation

$$R = e^{-.0016(2)^{.775}}$$

$$= .997.$$

3.4 Other Distributions

Although the exponential and Weibull density functions are the most frequently used in reliability analysis, there are occasions when other density functions are required. This is particularly true when wearout failures are being analyzed. The normal and lognormal density functions are appropriate to use for these types of failures.

The normal density function is a perfectly symmetrical distribution. It is the most widely used density function in engineering applications. The time to wearout of mechanical equipment often follows the normal density functions. A graph of a normal density function with a mean of zero and a standard deviation of one is given in Figure 3-6. This density function is referred to as the standard normal density function. It is the one found tabulated in most statistical tables. Computations about any normal density function can be made using the standard normal with the appropriate transformation.

The log-normal density function occurs in practice whenever we encounter a random variable such that its logarithm has a normal density function. In reliability analysis it is most frequently used for analyzing wearout failures. A graph of a typical log-normal density function is shown in Figure 3-6. It can be seen from the figure that this density function is positively skewed.

Many reliability problems deal with situations referred to as "repeated trials." For example, if we want to know the probability that 9 of 10 rounds will hit their target, that 1 of 5 rivets will shear in a tensile test, we are in each case concerned with a number of "trials" and we are interested in the probability of getting a certain number of "successes." If we are conducting a test, the number of test specimans represents the number of trials, and the number of specimans that do not fail represents the number of successes. The reliability is then the number of successes divided by the number of trials.

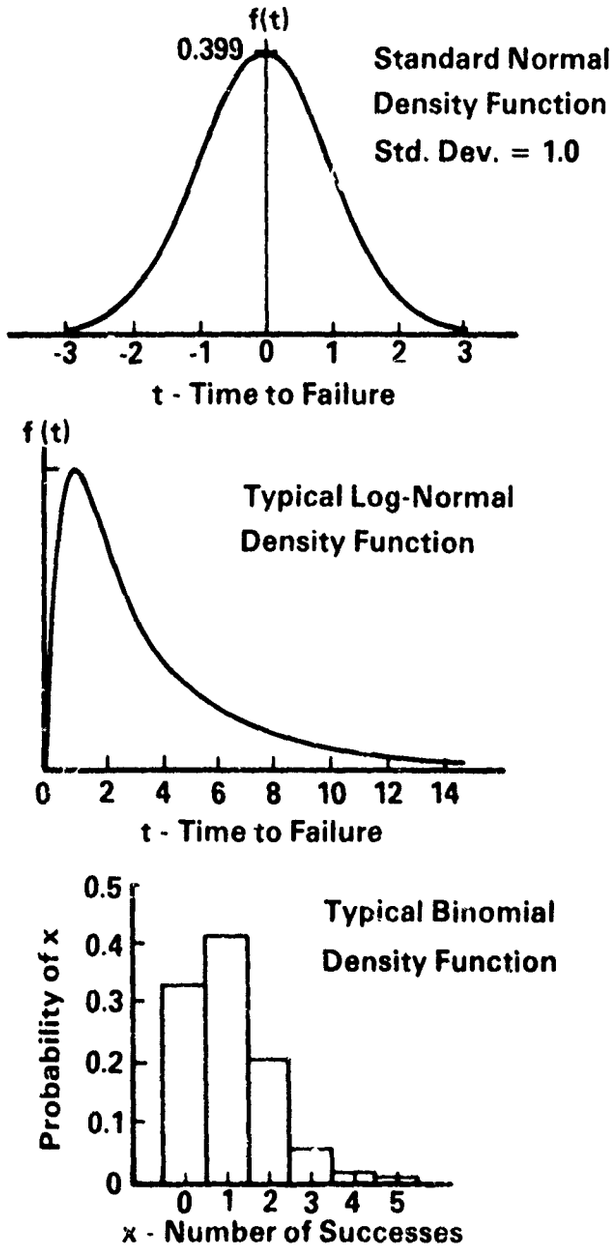


Figure 3-6 Graphs of Typical Distributions

One advantage that the binomial density function has over the other density functions used in reliability analysis is that no assumption is made about the time-to-failure density function. This is why it is referred to as distribution-free analysis.

A histogram of a typical binomial density function is given in Figure 3.6. A histogram is used because it is a discrete distribution, that is, the number of successes can only assume the non-negative integer values.

Section 4.0
Estimating Reliability Using Test Data

4.1 Determination of Failure Rates from Test Data

An effective and widely used method of handling reliability prediction problems is that of life testing. To conduct such tests, a random sample of n components is selected from a lot, put on test under specified environmental conditions, and the times to failure of the individual components are observed. If each component that fails is immediately replaced by a new one, the resulting life test is called a replacement test; otherwise, the life test is called a nonreplacement test. Whenever the mean lifetime of the components is so large that it is not practical, or economically feasible, to test each component to failure, the life test may be truncated, that is, it may be terminated after the first r failures have occurred or after a fixed period of time has been accumulated.

In tracking the reliability of Army weapon systems under development, it is normally the case that test costs are so high that life testing without truncation is not feasible, we will concentrate our discussion on truncated life testing. In what follows we will assume the failure time density function of each component is exponential, that n components are put on test, life testing is discontinued after a fixed number r with $r \leq n$, of components have failed, and that the observed failure times are $t_1 \leq t_2 \leq \cdots \leq t_r$.

We are interested in estimating the mean time between failures, the failure rate can be estimated using the equation

$$\hat{\lambda} = 1/\Theta .$$

From the results given by Epstein[2], it can be shown that an unbiased estimate of the mean time between failure is

$$\hat{\Theta} = T_r/r$$

where T_r is the total accumulated test time until r failures occurred. Thus for nonreplacement tests T_r can be obtained using

$$T_r = \sum_{i=1}^{r} t_i + (n-r)t_r .$$

Note that if the tests are continued until all items fail, T_t reduces to

$$T_r = \sum_{r=1}^{n} t_i$$

and the mean time between failures is simply the mean of the observed times to failure.

To illustrate the methods presented, let us consider the following example. Suppose that 10 rotor blades are placed on test without replacement and that the test is truncated after r = 3 of them have failed. Furthermore, suppose that the 3 failure times are 250, 410, and 480. Thus n = 10, r = 3, and

$$T_3 = (230 + 410 + 480) + (10-3)480$$

$$= 4480 \text{ hours}.$$

Hence we estimate the mean time between failure as

$$\hat{\Theta} = \frac{4480}{3}$$

$$= 1493 \text{ hours}.$$

The estimate of the failure rate is

$$\hat{\lambda} = 1/\hat{\Theta}$$

$$= \frac{1}{1493}$$

$$= .00067 \text{ failures/hour}.$$

This means that on an average one can expect seven failures to occur every 10,000 hours of operation.

4.2 Confidence Intervals

The estimates of the mean time between failures obtained from test data as described earlier, are called point estimates of the true unknown MTBF. We will now look at what confidence we can have in the point estimates.

Statistical estimates are more likely to be close to the true value as the sample size increases. Thus, there is a close correlation between the accuracy of an estimate and the size of the sample from which it was obtained. Only an infinitely large sample size could give us a 100 per cent confidence or certainty that a measured MTBF coincides with the true value. In testing of Army equipment the longer the test time and the greater the number of failures experienced, the greater the confidence one has in the point estimate MTBF.

When the estimate of the MTBF is obtained from a reasonably sized sample, we may logically assume that the true value of that MTBF will be somewhere in the neighborhood of the estimate, either greater or smaller. Therefore, it would be more meaningful to express the MTBF estimates in terms of an interval with an associated probability or confidence that the true value lies within such interval rather than to express them only as point estimates. This is exactly what we are doing when we assign confidence limits to point estimates obtained from test data.

Confidence intervals around the mean time between failures, or around any other point estimate, have a lower confidence limit and an upper confidence limit. For example, a 90 per cent confidence interval might read

$$P(500 \leq MTBF \leq 700) = 0.90.$$

The 500 hours is the lower confidence limit and the upper confidence limit is 700 hours. This confidence interval is structured so the probability that the interval between 500 and 700 hours includes the true value of the MTBF is equal to 0.90. Since the MTBF is estimated from test data, the true value of it is never known. However, the concept of a true value for the MTBF must be understood to correctly interpret confidence intervals.

There is another way of explaining the meaning of confidence intervals that relates it more to test data. The 500 and 700 hours in the above confidence interval were estimated based on the results of test data and the fact that the time between failures is exponentially distributed. If another test were conducted and a new set of confidence limits was calculated based on the results, the limits could be slightly different from 500 and 700 hours. If this same process was repeated many many times, realizing this is not practical to do so, then each time we would get slightly different confidence limits. However, on the average, 90 per cent of the confidence intervals thus computed would contain the true value of the MTBF. Also, on the average, 10 times out of 100 the interval thus computed would not contain the true value of the MTBF.

Confidence intervals are easily computed for the MTBF when the time-to failure distribution is the exponential. The information that is required is the number of failures that occurred during the test, the estimate of the MTBF obtained from the test results and the confidence level required. The point estimate MTBF is equal to the total accumulated test

hours divided by the total failures. Usually only the lower confidence limit for the MTBF is of interest. This is because we are primarily interested in the probability that the MTBF is greater than some minimum value. Or, stated another way, it is at least as great as the minimum value.

To illustrate how the lower confidence limit for the MTBF is computed, let us consider the following problem.

> Twenty items are tested for a period of 100 hours. Four of the items fail with failures occurring at 20, 48, 76, and 92 hours. Compute the lower 95% confidence limit for the MTBF.

In reliability terminology, this is a time terminated test without replacement of failed components. Time terminated means that the test is stopped after the predetermined time of 100 hours is reached. Without replacement means that as the failures occur, the failed items are not replaced. For this particular test, the equation for the lower confidence limit, $MTBF_L$, once the confidence level has been established, is as follows:

$$MTBF_L = \frac{2T}{\chi^2_{a,2r+2}}$$

where

T = total test hours, or item-hours

$\chi^2_{a,2r+2}$ = chi-square function

a = level of significance

r = number of failures that occurred during the test

The level of significance "a" is related to the confidence level as C.L. = 100(1 - a) in per cent. The numerical value of

the Chi-square function, $\chi^2_{a,2r+2}$ is obtained from a set of tables for the Chi-square distribution. In most of these tables the numbers in the vertical column on the left under ν, "nu", are the degrees of freedom and are equal to the quantity 2r+2 in the above equation. Across the top of the table are values of the levels of significance, a. Knowing "ν" and "a", values of χ^2 can be read from the table. For the above problem $\nu = 2r+2 = 10$. Under the $a = .05$ column, read 18.3 from a χ^2 table.

Therefore

$$\text{MTBF} = \frac{20+48+76+92+16(100)}{4}$$

$$= 459 \text{ hours (Approximate MTBF} = 460)$$

$$\text{MTBF}_L = \frac{2(1836 \text{ hrs.})}{18.3}$$

$$= 200.6 \text{ hours}.$$

What the above answer states is that one can be 95 percent confident that the MTBF will be at least as great as 200.6 hours.

The amount of confidence that can be placed in an estimate of the MTBF is a definite function of the number of hours tested and the number of items tested. That is we can raise our lower confidence limit for the MTBF by testing for a longer period of time, assuming that we don't run into wearout failures. If, in the above example, we test for 200 hours and have failures occur at the following times

20, 48, 76, 92, 105, 130, 155

the point estimate of the MTBF would be

$$MTBF = \frac{20 + 48 + 76 + 92 + 105 + 130 + 155 + 13\,(200)}{7}$$

$$= \frac{3226}{7}$$

$$= 461 \text{ (Approximate MTBF = 460)}$$

which is approximately the same estimate previously obtained. However the lower 95% confidence limit now is raised and becomes

$$MTBF_L = \frac{2(3226)}{26.29}$$

$$= 245$$

which is significantly higher than previously obtained.

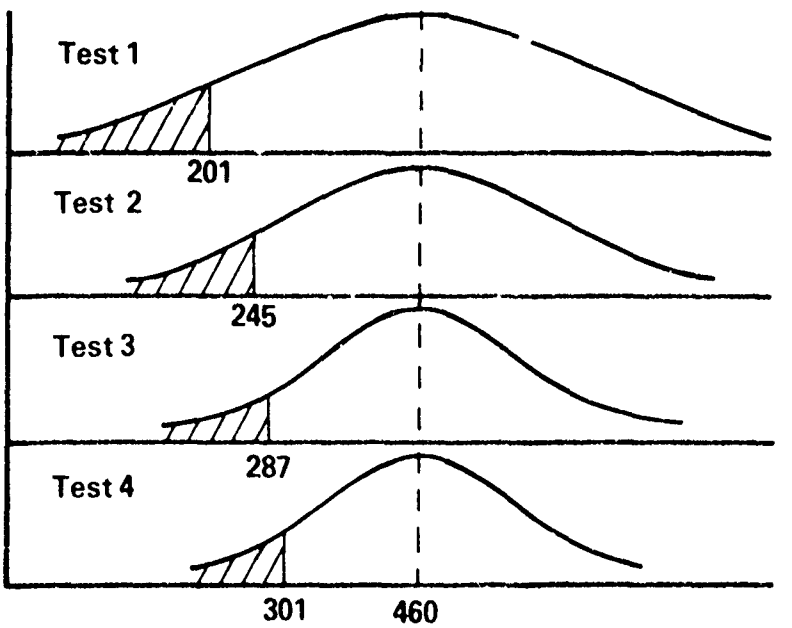
This same pattern continues as the test time is increased. Table IV summerized the results for four different cases, including the two presented above. Each of these cases resulted in about the same estimate for the MTBF, but due to the longer test time the lower confidence limit is higher. This same argument can be extended to the number of units tested.

The points made in this section can be summarized by stating that there must be a trade-off between the resources available for testing and the confidence levels required on reliability estimates. We can never achieve the ultimate of 100 percent confidence in our results. Neither can we afford not to have any testing. A balance between resources and confidence is our objective.

TABLE 4-1

ILLUSTRATION OF HOW A LONGER TEST TIME CAN INCREASE THE LOWER MTBF CONFIDENCE LIMIT

Test No.	Test Time	Number Failures	Approximate MTBF	Lower Confidence Limit
1	100	4	460	201
2	200	7	460	245
3	300	10	460	287
4	400	12	460	301



MEAN TIME BETWEEN FAILURES

4.3 Reliability Growth Curves

4.3.1 Introduction

The subject of reliability improvement by means of conscious efforts on the part of designers, test engineers, customers, etc., has been of interest since the beginnings of reliability analysis. The modeling of such growth processes has followed, for the most part, a common procedure: Formulas are presented that are intended to represent the growth of reliability as a function of time. These formulas contain unknown parameters, and it becomes an exercise in statistics to find appropriate estimates for these parameters as a function of observed failure data.

The central purpose of most reliability growth models includes one or both of the following objectives:

a) prediction of the current system reliability,
b) projection on the system reliability for some future development time.

As in any mathematical model, reliability growth models are idealizations. They are based on a number of assumptions that vary with the different models. If a program manager desires to use a growth model to help him plan a development program, he should choose a particular model based on prior experience with similar type systems. As the development program progresses, he can use the model along with test data to monitor and project the reliability of the system and make necessary decisions accordingly.

4.3.2 The Duane Model

One of the most widely used growth models used for Army systems is the Duane model presented in reference 6 by J.T. Duane of the General Electric Company. He analyzed test and operational data for programs with test times as high as 6 million hours on five divergent groups of products. The five

groups included two hydro-mechanical devices, two aircraft generators, and one jet engine. His analysis revealed that for these systems, the observed cumulative failure rate versus cumulative operating hours fell close to a straight line when plotted on log-log paper.

Crow [8] presents a formal mathematical development of the Duane model. He showed that when the above conditions hold, the failure rate during development follows the Weibull failure rate curve. The development given below and the notation are similar to that given by Crow.

Mathematically, this model may be expressed by the equation

$$F(t) = \lambda t^{-a}, \qquad \lambda > 0$$

$$0 < a < 1$$

where F(t) is the cumulative failure rate of the system at time t and λ and a are parameters. The cumulative failure rate is by definition

$$F(t) = \frac{E(t)}{t}$$

where E(t) is the expected number of failures experienced by the system during t time units of development testing. Thus from the above two equations

$$E(t) = \lambda t^{1-a}.$$

The instantaneous failure rate, r(t), is of most interest for applications. It is defined as the change in the expected number of failures per unit time. For a nonexponential system, it varies with time while for an exponential system the failure rate is constant.

Differentiating E(t) with respect to time gives the instantaneous failure rate r(t) as follows:

$$r(t) = \frac{dE(t)}{dt}$$

$$= (1-\alpha)\lambda t^{-\alpha}.$$

By substituting in the above equations

$$\beta = 1-\alpha$$

one gets

$$r(t) = \lambda \beta t^{\beta-1}$$

which is the Weibull failure rate function for a repairable system.

Thus if one plans to use the Duane model during a development program, the above expression can be used to determine the failure rate at a particular development time t. The values of λ and β are estimated from test data. Since λ is only a multiplier and β determines how much the failure rate changes with the development time, β is referred to as the growth parameter. For the systems studied by Duane, a β of approximately 0.5 was estimated.

A report published by J.D. Selby and S.G. Miller, also of G.E., gives some additional computational experience with the Duane model. An established and experience growth rate of 0.5 for an aggressive reliability program was reported. A maximum growth of .4 was estimated which had never been achieved. A minimum of .9 can be expected on those programs where no real specific consideration is given to reliability.

To gain further insight into the Duane model, consider Figure 4-1 which is a plot of the Weibull failure rate versus development time for $\beta = 0.5$ and $\lambda = 0.4$. During the early

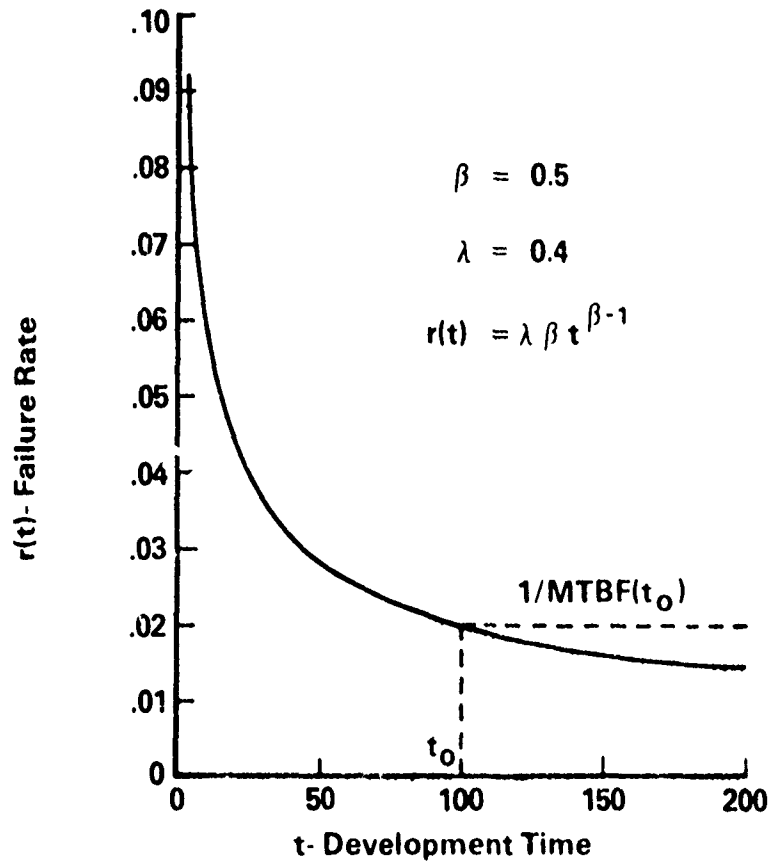


Figure 4-1 Failure rate versus development time for Weibull Failure Rate

stages of development the failure rate decreases rather rapidly due to more failures and more rework going on during this time. As the development progresses, the rate of decrease of the failure rate drops off considerably. The Duane model assumes that at some time t_0, which corresponds to about the time that development ends and production starts, the failure rate levels off to a fairly

constant value. At this point in time when the failure rate becomes constant, the time between failures can be described by the exponential distribution with a mean time between failure of

$$\text{MTBF}(t_o) = \frac{t_o^{1-\beta}}{\lambda \beta} .$$

Note the similarity between the failure rate curve described here and the life characteristic curve previously described.

Crow [8] has developed the maximum liklihood estimates (MLE) of λ and β and also a goodness of fit test to determine if the Duane model fits a particular set of data. The MLE estimate for β is

$$\hat{\beta} = \frac{N}{\sum_{r=1}^{K} \sum_{i=1}^{N_r(T)} \log \frac{T}{X_{ir}}}$$

where

K = number of different subsystems,

T = the operation time for each of the K subsystems,

N_r (T) = number of failures observed for the r-th subsystem during T time,

X_{ir} = the age of the r-th subsystem at the i-th failure, beginning of development being 0,

and

$$N = \sum_{i=1}^{K} N_r(t).$$

The above MLE estimate of β is biased. The unbiased estimate is obtained by using

$$\bar{\beta} = \frac{N-1}{N}\hat{\beta}.$$

The MLE of λ is

$$\hat{\lambda} = \frac{N}{KT^{\hat{\beta}}}.$$

The chi-square goodness of fit test can be used to determine if the observed data fits the Duane model. The chi-square statistic is calculated using

$$\chi_c^2 = \sum_{i=1}^{c} \frac{(O_i - E_i)^2}{E_i}.$$

To compute the statistic the development time is divided into c intervals. The observed number of failures in the i-th interval, O_i, is obtained from the observed data. The expected number of failures in the i-th interval, E_i, is obtained using

$$E_i = \frac{N(t_i^{\bar{\beta}} - t_{i-1}^{\bar{\beta}})}{T^{\bar{\beta}}}$$

where t_{i-1} and T_i are the beginning and ending times for the i-th interval. The χ_c^2 is compared with the tabled value of chi-square, χ_T^2, with degrees of freedom equal to c-1 and

the specified level of significance. If

$$\chi^2_C < \chi^2_T$$

then it can be concluded that the data fits the Duane model.

4.3.3 Application

An engine system was analyzed for reliability growth using the Duane model. The data available for analysis was based on 806.3 hours of development testing. During this time there were 40 failures and the times of each failure were recorded. The average rates for this system during each 1000 hour interval are shown in Figure 4-2.

Using the data the MLE's of λ and β were computed to be

$$\hat{\lambda} = 0.128$$

$$\hat{\beta} = 0.639.$$

The unbiased estimate of β is

$$\bar{\beta} = 0.623.$$

The chi-square goodness of fit statistic was calculated next using an interval width of 1500 hours. The result was

$$\chi^2_C = 1.343.$$

Using a 1% level of significance and a degrees of freedom of 6-1 5 , the tabled value of chi-square is

$$\chi^2_T = 15.086.$$

Thus it can be concluded that the Duane model fits the data

Using the formula on page 46, the estimated failure rate for the engine becomes

$$r(t) = .128(.623)\, t^{.623-1}$$

$$= .08\, t^{-.377}.$$

A plot of this failure rate curve is given in Figure 4-2. Notice how the curve is beginning to flatten out. In fact it would take 100,000 hours of development time to get the failure rate down to .001 failures/ hour.

Failure Times

1, 43, 43, 171, 234, 274, 377, 530, 533, 941, 1074, 1188, 1248, 2298, 2347, 2347, 2381, 2456, 2456, 2500, 2913. 3022, 3038, 3728, 3873, 4724, 5147, 5179, 5587, 5626, 6824, 6983, 7106, 7106, 7568, 7568, 7593, 7642, 7928, 8063

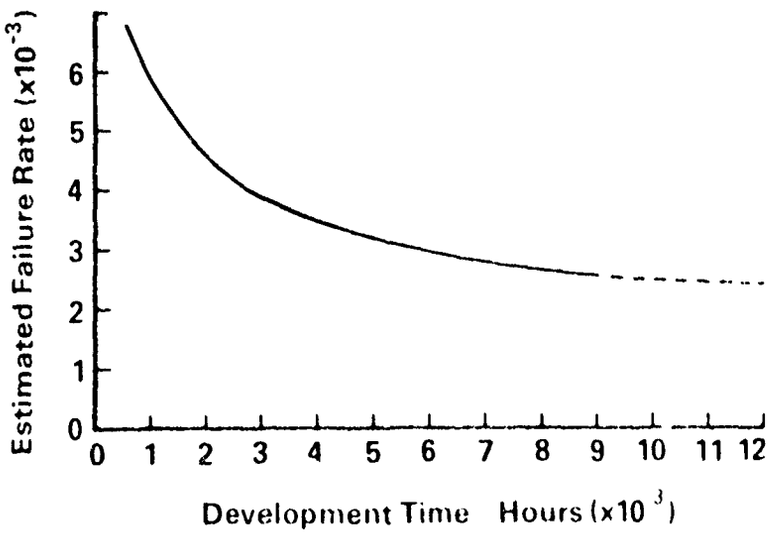


Figure 4 2 Failure Times and Estimated Failure Rate for Example

Section 5.0
Operating Characteristic Curves

5.1 Use and Interpretation of OC Curves

Conclusions concerning equipment reliability will of necessity be based on the results obtained from a test sample. The behavior of a sampling plan is established by specifying the lot size for which it is to be used, the sample size to be tested, and the number of faulty pieces which will cause acceptance or rejection of the lot. From this information it is possible to predict the results that will be obtained when the plan is used in practice. Specifically, it is possible to predict the risk that a wrong decision will be made by either accepting lots of poor quality or rejecting lots of good quality. Such predictions are made from the operating characteristic (OC) curve of the sampling plan. Each different sampling plan has its own OC curve.

The use of operating characteristic curves can greatly aid in the selection of reliability test plans. The purpose of using them is to control the cost of the testing program while at the same time assuring that reliability goals are met. This purpose is accomplished by allowing one to use a systematic, quantitative approach to evaluate and select the test plan rather than relying completely on subjective judgement to make the selection.

As we saw earlier in Section 4.1, test data can allow us to draw conclusions concerning the reliability or MTBF of the equipment items undergoing test. However, it should be emphasized again that these conclusions concerning the MTBF are based on samples and can, therefore, be in error. This is where OC curves provide valuable assistance. They allow both the customer and the manufacturer of the test items to predict the risk of error when basing their conclusions concerning product reliability on a sample of test results.

One of the important characteristics of a demonstration test plan is its probability of acceptance as a function of the test items actual MTBF. The probability of acceptance of a test plan is the probability of concluding that the test results are satisfactory and indicate an acceptable MTBF. For example: One of the test plans taker from the Department of Defense handbook, "Quality Control and Reliabiity Handbook (interim) H-108 - Sampling Procedures and Tables for Life and Reliability Testing (Based on Exponential Distribution)" has the following acceptance criteria.

> "Select 22 items at random from a lot and place these items on test. When an item fails, replace it with another item selected at random from the lot. If the test continues for 500 hours with not more than two failures, accept the lot. If 3 failures occur prior to 500 hours, reject the lot."

This plan could be used to demonstrate a desired MTBF of 10,000 hours by following the above accept, reject criteria. However, before deciding to use this plan (or any other) we should examine its OC Curve which is shown in Figure 5-1. This curve represents a sample size of M - 11000 item hours (i.e. 500 hours x 22 items = 11000) and an acceptance number of K -- 2.

The "α" of Figure 5-1 is designated the producer's risk and is defined as the probability of rejecting equipment with a true MTBF equal to the 10,000 hour MTBF. The term α is also called the level of significance. The "β" of Figure 5.1 is designated the consumers risk and is equal to the probability of accepting equipment with a true MTBF equal only to a minimum acceptable MTBF, such as 2000 hours. It should be pointed out that the minimum acceptable MTBF is not necessarily a desirable MTBF and is selected so that an associated and specified risk (β) of accepting equipment of this value is tolerable. Examination of Figure 5.1 shows the way in which the probability of accepting test results as valid

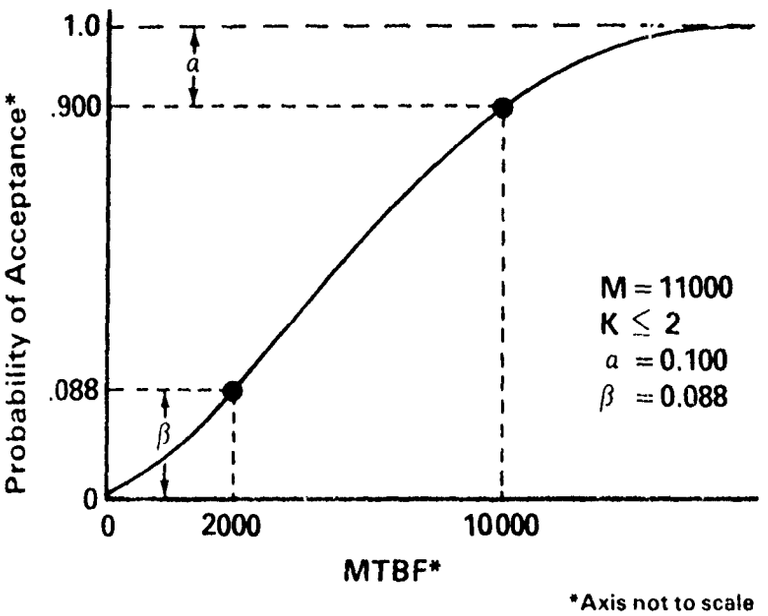


Figure 5-1 Operating Characteristic Curve

indicators of the desired MTBF increases as the true MTBF of the equipment on test approaches the 10,000 hour MTBF.

The selection of a sound test plan involves considerable experience and judgement to insure that a plan is chosen with proper values for a and β. It is these values that dictate the overall cost of a testing program. The three major costs associated with any sampling and testing plan are:

1. The costs associated with rejecting equipment as not meeting a reliability standard (such as the 10,000 hours MTBF given above) when in fact it does meet it. This is called a Type I error and the chance or probability of making this error is called the producers risk (a).
2. The cost associated with accepting equipment as meeting a reliability standard (such as the 10,000 hours MTBF) when in fact it is some specified value less than

the standard. This is called a Type II error and the chance of making this mistake is called the consumer risk (β).

3. The cost associated with the sampling and testing operation. This cost is generally assumed to be directly proportional to the number of units tested.

Examination of the curve in Figure 5-1 provides some insight concerning the test plan's ability to discriminate between good and bad values of the actual MTBF for a production lot. It shows that there is a 90% chance that 2 or fewer failures will occur when using this plan to test production items with a true MTBF of 10,000 hours. If this occurs, the previously stated acceptance criteria of our test plan will be met and we will correctly conclude that the items under test have a MTBF of at least 10,000 hours. The producer of the item under test should also be satisfied with this arrangement because there is only approximately a 10% chance (i.e. α = 1.00-0.90 = 0.10) that more than 2 failures will occur and cause the rejection of equipment that actually has an MTBF of the called for 10,000 hours.

On the other hand, when using this same test plan there is an 8.8% chance (i.e. β = .088 that two or fewer failures will occur and the acceptance criteria will be met when testing items with a true MTBF of only 2000 hours. Stated another way, if a contract called for an MTBF of 10,000 hours to be demonstrated by using the above plan, an 8.8% chance exists that the demonstration could be successfully made with equipment whose MTBF was only 2000 hours.

The consumer must now decide whether or not the risk of making an error and experiencing the associated costs of accepting equipment with an MTBF of only 2000 hours can be tolerated. If the decision is made that an 8.8% chance of accepting a production lot with a true MTBF of 2000 hours is too great a risk the situation can be remedied by making a change in the test plan.

These changes can take many forms but normally the intent is to change the shape of the OC curve. The following discussion will be limited to the impact on the OC curve of changing either sample size or acceptance number. The sample size can be changed by altering the number of items on test or changing the length of time the items remain on test. Of course, care must be taken to insure that wearout does not influence the failure rate and that the probability that an item will fail remains independent of the number of hours an item has been on test. Figure 5-2 illustrates how the probability of acceptance is reduced by increasing the sample size to 22,000 item hours. This increase could have been achieved either by doubling the number of items on test from 22 to 44, doubling the length of test time to 1000 hours for the original 22 items, or some other combination of increased test time and increased number of items on test.

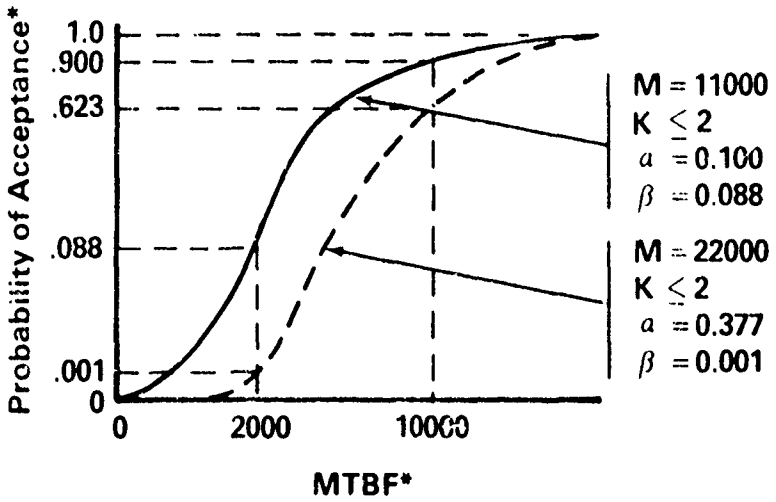


Figure 5-2 Operating Characteristic Curve

The above change in sample size reduces the probability (β) that items with a true MTBF of only 2000 hours will be judged acceptable to 0.001. However, while this change is desirable, two things have happened that are not desirable. First, our sampling and testing costs have increased. Second, the probability (a) of rejecting items with an acceptable MTBF of 10,000 hours has been increased to 0.377 and the producer will not be happy about this.

Another way to bring about the same general type of change in the shape of the OC curve is to reduce the acceptance number from 2 to 1. The results are shown in Figure 5-3.

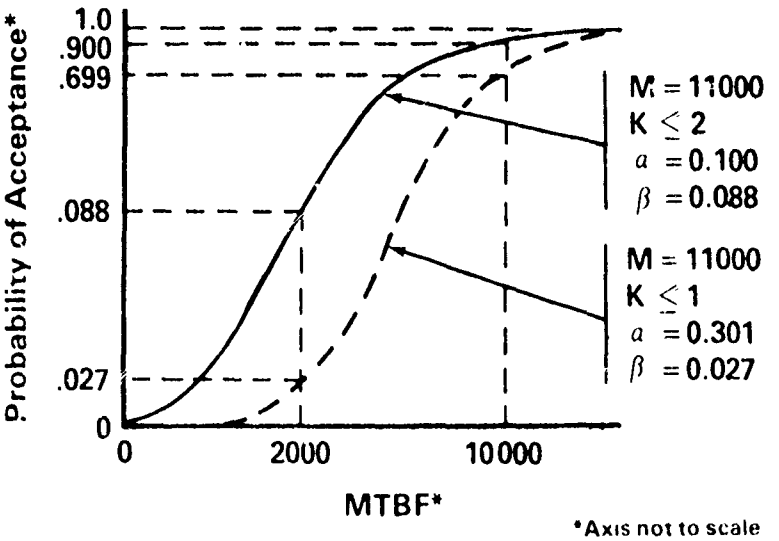


Figure 5-3 Operating Characteristic Curve

In this case the probability (β) of accepting a lot with an MTBF of 2000 hours is reduced to 0.027 while the probability (a) of rejecting a lot with the desired MTBF of

10,000 is increased to 0.301. The difference between this situation and the previous one is that we have not increased our sampling and testing cost and in fact may have reduced it slightly if a curtailed testing scheme is being used.

The ability of our test plan to discriminate better between an MTBF of 10,000 hours and an MTBF of 2,000 hours can be accomplished by increasing both the acceptance number and the sample size. The result is a decrease in both the producers risk (α) and the consumers risk (β). For example; if the sample size is increased to 31,500 hours and the acceptance number to 6, the result is an OC curve like that shown in Figure 5-4.

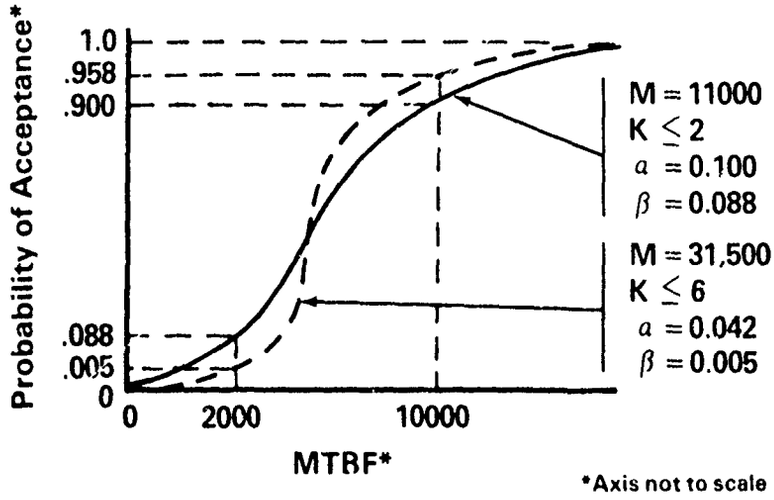


Figure 5-4 Operating Characteristic Curve

In this case, the probability (β) of accepting items with a true MTBF of 2000 hours has been reduced from 0.10 to 0.005 and, at the same time, the probability (α) of rejecting items with a true MTBF of 10,000 hours has also been reduced from 0.10 to 0.042. These are desirable results for

both producer and consumer but the penalty paid is an increase in sampling and testing costs.

It sometimes requires an extremely large sample size to provide both the producer and the consumer with the degree of protection they desire from the risk of making an error. For example; suppose that in our example of Figure 5-1 we and the producer are satisfied with the 90% chance of accepting a production lot whose true MTBF is 10,000 hours. However, we feel that the 8.8% risk of accepting a production lot whose true MTBF is only 2000 hours cannot be tolerated. Assume further that if the lower MTBF of 2000 hours were raised to 7000 hours, we could tolerate about a 10% chance of acceptance at this level. The sample size or number of item hours required to satisfy these new requirements and also keep a at 0.10 is equal to approximately 330,000 item hours with an acceptance number of K=4.0. This is a substantial increase over the 11,000 item hours required earlier and may not be possible because of excessive testing costs. If such large scale testing were not feasible a compromise would have to be reached and OC curves provide a mechanism for doing this intelligently by helping us evaluate the risks of error associated with different test plans.

5.2 Examples of Calculations

The OC Curves in the previous section were developed to show the probability that a specific testing and sampling plan would indicate an acceptable MTBF for various true values of equipment MTBF. Acceptable results were indicated when the number of failures during a test was found to be equal to or less than some specified number. In the example shown in Figure 5-1 of the previous section we concluded that an acceptable MTBF was demonstrated when two or fewer failures occurred during a 500 hour test of twenty-two items. The probability of two or fewer failures

during the test is determined by calculating the probability of exactly two failures, exactly one failure, and zero failures and adding these probabilities together.

Calculations of the specific values for the OC Curves in Section 5.1 were accomplished by using the Poisson distribution as an approximation to the Binomial Probability Law. This approximation is very good for small probabilities of failure and relatively large sample sizes. These conditions are easily met in most life testing situations.

The equation for the Poisson distribution is:

$$P(k) = \frac{(\lambda X)^k \quad e^{-\lambda X}}{k!} \qquad (5\text{-}1)$$

where:

$P(k)$ = The probability of k failures

k = Specified number of failures

λ = Failure rate per hour (i.e. $\lambda = \frac{1}{MTBF}$ where MTBF is given in hours.)

X = Test Time (i.e. No. of items being tested multiplied by the length of test in hours)

e = the constant 2.71828

The above equation will now be used to calculate the probability that exactly "k" failures will occur during a life test. The failure rate per hour is simply the reciprocal of the MTBF given in hours. The test time in "item hours" is determined by multiplying the number of items being tested by the length of the test in hours. This calculation assumes

the replacement of failed items during the duration of the test. For example: The probability of exactly two failures occurring when 22 items with a true MTBF of 10,000 hours are placed on test for 500 hours is determined as follows:

$$X = \text{Test Time} = \text{No. Items x Length of Test}$$

$$= 22 \times 500$$

$$= 11{,}000 \text{ item hours}$$

$$\lambda = \text{Failure Rate} = \frac{1}{\text{MTBF}}$$

$$= \frac{1}{10000}$$

$$= 0.0001 \text{ failures/hour}.$$

Since we wish to determine the probability that exactly two failures will occur (i.e. $k = 2$) we must now evaluate Equation 5-1.

$$P(k=2) = \frac{(0.0001 \times 11000)^2 \, (2.71828)^{-(.0001 \times 11000)}}{2!}$$

$$= 0.201.$$

This evaluation indicates that there is a probability of 0.201 that exactly two failures will occur during the test. To determine the probability of our acceptance criteria being met (i.e. $k \leq 2$) we must next find the probability of exactly one and also the probability of zero failures occurring and then add the three values together.

These last two values are:

$$P(k = 1) = \frac{(0.0001 \times 11000)^1 (2.71828)^{-(.0001 \times 11000)}}{1!}$$

$$= 0.366$$

$$P(k = 0) = \frac{(0.0001 \times 11000)^0 (2.71828)^{-(.0001 \times 11000)}}{0!}$$

$$= 0.333.$$

By adding these three values we get a probability of 0.201 + 0.366 + 0.333 = 0.900 as the probability of two or fewer failures during the test. This value gives us one point on the OC Curve for the test.

A second point on the curve for this test of 22 items can be calculated in a similar fashion. As before, we assume a 500 hour test and an acceptance number of $k \leq 2$. However, this time a true MTBFof only 2000 hours will be used. The probability of two or fewer failures is then

$$P(k = 2) = \frac{(0.0005 \times 11000)^2 (2.71828)^{-(.0005 \times 11000)}}{2}$$

$$= 0.062$$

$$P(k = 1) = 0.022$$

$$P(k = 0) = 0.004$$

$$P(k = 2) + P(k = 1) + P(k = 0) = 0.088.$$

This total indicates the probability of two or fewer failures to be equal to 0.088. These two points are shown on the OC Curve of Figure 5-1 and repeated below in Figure 5-5.

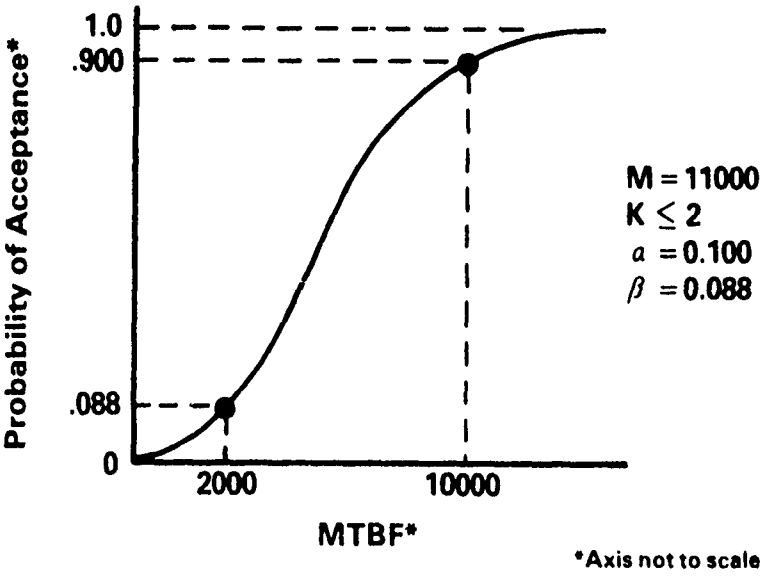


Figure 5-5 Operating Characteristic Curve

Other points on the curve can be calculated in a manner similar to that shown for MTBF of 2000 hours and 10,000 hours. The OC Curve itself can be thought of as showing the probability that our acceptance criteria will be met when the true MTBF varies from zero to over 10,000 hours.

Section 6.0

Bayesian Methods in Reliability Analysis

6.1 Introduction

A particular problem associated with reliability prediction and assessment is that there is normally a very long time period between the design phase and the time at which sufficient usable test data has been accumulated. Predictions during the early period of design and development are traditionally based on engineering judgement. Classical prediction methods are not adequate when only a small amount of data is available. Bayes' Theorem presents a method of assessing achieved reliability during the lengthy interim phase, when management decisions regarding design and development can have their biggest impact. Bayes' technique combines relevant operating experience early in the program phases with the prior prediction to form a new prediction. As more and more test data are accumulated, the prediction is continually updated.

Bayes' analysis begins by assigning an initial reliability on the basis of whatever evidence is currently available. The initial prediction may be based solely on engineering judgement or it may be based on data from other similar types of items. Then, when additional test data is subsequently obtained, the initial reliabilities are revised on the basis of this data by means of Bayes' Theorem. The initial reliabilities are known as prior reliabilities in that they are assigned before the acquisition of the additional data. The reliabilities which result from the revision process are known as posterior reliabilities.

6.2 Bayes' Theorem

A basic theorem in probability theory which relates joint and conditional probabilities is

$$P(A \text{ and } B) = P(A|B)P(B).$$

From the above equalities we can write

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

which is called Bayes' Theorem. With this expression we are attempting to estimate the posterior probability, $P(B|A)$, which is interpreted as the probability of getting outcome B given that outcome A has occurred. The probability P(B) is the prior probability which is being revised.

To illustrate Bayes' Theorem, let's consider the following specific example:

> There are three suppliers of a servo, B_1, B_2, and B_3. Company B_1 supplies .6 of the servos, B_2 supplies .3, while B_3 supplies .1. Past history indicates that 95% of the servos supplied by B_1 perform according to specifications, 80% of those supplied by B_2, and 65% of those supplied by B_3. Given that a servo performed according to specifications, compute the probability that it came from B_1, B_2, and B_3.

The prior information in this problem is the percent of servos provided by each supplier. The updated information is that a servo performs according to specifications. We want to compute the posterior probabilities based on the new information. The posterior probabilities will give us the updated probabilities that the servo came from B_1, B_2, or B_3. Note that without this new information and without Bayes' Theorem, the only conclusion we could draw is the same as that given in the prior information.

To solve this problem, we will write Bayes' formula as follows,

$$P(B_r|A) = \frac{P(A|B_r)P(B_r)}{P(A)}$$

where

$r = 1, 2, \text{and } 3$

B_r = event "servo came from supplier r"

A = event "servo performed according to specifications".

In the above equation, the denominator can be written:

$$P(A) = \sum_{i=1}^{3} P(B_i)P(A|B_i).$$

Expanding, this equation becomes

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3).$$

We can visualize this situation by constructing a tree diagram like the one shown in Figure 6-1, where the probability of the final outcome "A" is given by the sum of the products of the probabilities corresponding to each individual branch.

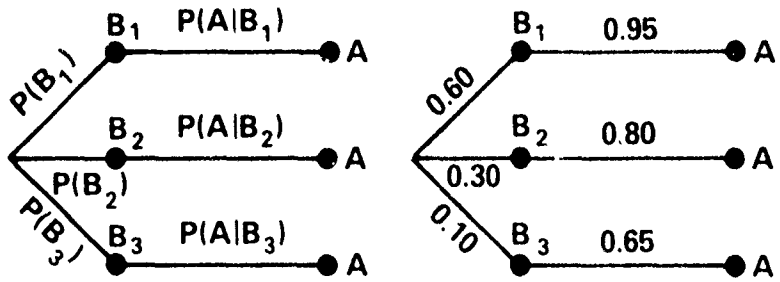


Figure 6-1 Tree Diagram for Example

Then the probability that any one servo received by the plant will perform according to specifications is

$$P(A) = .6(.95) + .3(.80) + .1(.65)$$

$$= 0.875.$$

The posterior probabilities can now be computed as,

$$P(B|A) = \frac{P(B_1)P(A\ B_1)}{P(A)}$$

$$= \frac{.6(.95)}{.875}$$

$$= .65$$

$$P(B_2|A) = .27$$

$$P(B_3|A) = .07.$$

Thus the new information has provided us revised estimates about who supplied the servo. Table 6-1 shows the prior and posterior probabilities for comparison purposes.

6.3 Bayes' Theorem Applied to Reliability Analysis[1]

As has been mentioned earlier, Bayes' Theorem is used in conjunction with prior information and current data to provide updated reliability estimates. Table 6-2 shows the results for several different cases when the prior MTBF was

[1] The discussion given here is based on that given in references 3 and 4

Table 6-1 Results for the Bayes Theorem Example

	Probability	
Supplier	Prior	Posterior
B_1	.6	.65
B_2	.3	.27
B_3	.1	.07

Table 6-2 Results of Bayesian Analysis

Prior MTBF (Hrs)	Operating Time (Hrs)	Failures	Posterior MTBF (Hrs)
10,000	0	0	10,000
10,000	1000	0	10,547
10,000	5000	0	11,010
10,000	1000	1	8,517
10,000	5000	1	9,675

Table 6-3 Typical Discrete Prior Distribution

Cell No.	Cell Value (R_i)	Prior Cell Probability $P(R_i)$
1	1.	.002
2	.999975	.002
3	.999925	.496
4	.999875	.408
5	.99975	.002
6	.999625	002
.	.	.
.	.	.
.	.	.

10,000 hours. For each case the prior MTBF, operating time, and number of failures during this operating time were used to obtain posterior MTBF.

The discussion given here is confined to the exponential failure distribution, since this is the one most commonly applied to reliability analysis. The method given in references 3 and 4 uses a discrete prior rather than a continuous prior. When using continuous priors, difficulties arise due to the mathematical functions involved.

The given discrete prior is divided into a number of discrete cells. Each cell consists of a cell reliability, R_i, and a probability associated with this reliability, $P(R_i)$. The first six cells for a typical prior distribution is given in Table 6-3. Each cell also has a lower and upper boundary for R; however, for simplicity they are not shown. Note how the distribution is heavily biased toward the lower cells. This bias is made in order that any test data which differs significantly from the predicted prior failure rates will readily "wash out" the prior. Test data producing results similar to the prior will not significantly alter the value when forming the posterior. This is because cells 3 and 4 which have probabilities associated with them contain the prior reliability value which was .9999 for this particular problem.

After the prior distribution is established and some test data becomes available, the posterior reliability is computed. This is done by first computing $P(B|R_i)$ for each cell where

B = event "x failures in t total units of time".

To compute this probability, one uses the number of failures of the item, number of current hours accumulated on the item, and the cell failure rate that is computed using formulas given in references 3 and 4.

Next the P(B) is calculated by summing individual cell values of $P(B|R_i)$, times the cell probabilities $P(R_i)$. The

posterior probability for each cell $P(R_i|B)$ is calculated as the ratio

$$P(R_i|B) = \frac{\{P(B|R_i)\}\ \{P(R_i)\}}{P(B)}$$

where

$$P(B) = \sum_{i=1}^{n} P(B_i)P(A\,B_i)$$

and summed over each cell. When this sum equals to .5, then the corresponding reliability value is interpolated. This median reliability value is the posterior reliability prediction.

6.4 Example

Suppose that a system consists of two subsystems. Based on past experience, it is known that the time to failure for each subsystem is exponentially distributed. It is assumed that a discrete prior similar to the one given in Table 6-3 describes the prior distribution of the reliability for each subsystem.

The prior failure rates obtained from the discrete prior distributions are .00010 and .00050 failures per hour for subsystems 1 and 2 respectively. This yields prior reliabilities for 1 hour mission times of

$$R_1 = e^{-.0001} = 0.99990$$

$$R_2 = e^{-.0005} = 0.99950.$$

Using the above data and the procedure previously outlined, the posterior reliability for each subsystem can be

obtained. The results obtained for the updated failure rates for each subsystem are

$$\lambda_1 = .00889 \text{ failures/hour}$$

$$\lambda_2 = .00048 \text{ failures/hour}$$

and for the system

$$\lambda_S = .00889 + .00048$$

$$= .00937 \text{ failures/hour}.$$

The updated reliability predictions for 1 hour mission times are

$$R_1 = e^{-.00889} = 0.99115$$

$$R_2 = e^{-.00048} = 0.99952$$

$$R_3 = e^{-.00937} = 0.99067.$$

Although this is a relatively small example the calculations are similarly done even when applied to large army weapon systems. Using the prior distribution and current test data, a new updated distribution is obtained, referred to as the posterior distribution. The reliability of the total system is then obtained in the usual manner.

Section 7.0

Summary

The foregoing material explains and illustrates the need for a quantitative approach to reliability analysis. The determination of overall system reliability was discussed in:

Section 2.0, "Reliability of a Single Component Versus Multiple Components",

Section 3.0, "The Exponential and Weibull Models" and

Section 6.0, "Bayesian Methods in Reliability Analysis".

Use of test data as input to system reliability models and as an aid in controlling program costs was discussed in:

Section 4.0, "Estimating Reliability Using Test Data" and

Section 5.0, "OC Curves in Reliability Analysis".

It was pointed out that during the analysis of system reliability that it was desirable to evaluate the impact of different system configurations and changes in component reliabilities on overall system reliability. This type of evaluation can contribute greatly to the overall design effort and insure that reliablity goals are met by using the most economically efficient design.

Test programs are necessary to insure that various components meet their design specifications. The material on estimating reliability from test data and the use of OC Curves can be used to insure that test programs yield adequate data and supply this data at reasonable costs. The danger of establishing a test program without using the proper statistical tools is that either the results will have cost an excessive amount to obtain or they will not be meaningful in nature.

References

1. Lloyd, David K. and Lipow, Myron, Reliability: Management, Methods, and Mathematics, Englewood Cliffs, N.J., Prentice-Hall, 1962.
2. Epstein, B., "Statistical Life Test Acceptance Procedures," Technometrics, Vol. 2, November, 1960.
3. Saunders, D.E., and Monahan, M.H., "GE12 Reliability Model," General Electric Report No. RMAEG11, May 20, 1971.
4. MacFarland, W.J., "Use of Bayes Theorem in its Discrete Formulation for Reliability Estimation Purposes," Seventh Reliability and Maintainability Conference.
5. Crow, L.H., "Reliability Growth Modeling," AMSAA Technical Report No. 55, August, 1972.
6. Duane, J.T, "Learning Curve Approach to Reliability Monitoring," IEEE Transactions on Aerospace, Vol. 2, No. 2, 1964.
7. Selby, J.T. and Miller, S.G., "Reliability Planning and Management-RPM," AES Report SI-471, September 26, 1970, ASQC/SRE Seminar, Niagara Falls, New York.
8. Crow, L.H., "On Reliability Growth Modeling," AMSAA unnumbered report, October, 1973.
9. Grant, I.G. and Leavenworth, R.S., "Statistical Quality Control, Fourth Edition, McGraw Hill Book Company. New York, N.Y.