

AD/A-007 093

A RELIABILITY TEST PROCEDURE FOR
PROTOTYPE SMALL CALIBER ARMS

Carl Samuel Royer

Army Materiel Command

Prepared for:

Texas A and M University Graduate Center

April 1974

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER USAMC-ITC-02-08-73-012	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER ADIA-007093
4. TITLE (and Subtitle) A RELIABILITY TEST PROCEDURE FOR PROTOTYPE SMALL CALIBER ARMS		5. TYPE OF REPORT & PERIOD COVERED Final
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Carl Samuel Royer		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Dept of Maintenance Effectiveness Engineering USAMC Intern Training Center Red River Army Depot, Texarkana, Texas 75501		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Dept of Maintenance Effectiveness Engineering and Texas A&M University Graduate Center USAMC Intern Training Center - USALMC		12. REPORT DATE April 1974
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Directorate of Research, Development and Engineering U.S. Army Materiel Command Alexandria, Virginia 22304		13. NUMBER OF PAGES 100
		13. SECURITY CLASS. (of this report)
		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 Research performed by Carl Samuel Royer under the supervision of Dr. Ronald S. Morris, Assistant Professor, Industrial Engineering Department, Texas A&M University		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Reliability Testing, Small Caliber Arms		
<div style="text-align: right;"> Reproduced by NATIONAL TECHNICAL INFORMATION SERVICE US Department of Commerce Springfield, VA 22151 </div>		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The purpose of this research report is to find a suitable test procedure for determining the reliability of prototype small caliber arms submitted to the U.S. Army by private contractors for engineering design tests. In finding a solution, present reliability testing procedures are researched, and a suitable procedure is developed.		

PRICES SUBJECT TO CHANGE

The developed procedure is limited by its underlying assumptions that there are three independent modes of failure, and the number of rounds between each type of failure has an underlying exponential distribution. Areas of future research and development are suggested, and recommended changes in present Army prototype small caliber arms testing techniques are made.

FOREWORD

The research discussed in this report was accomplished as part of the Maintenance Effectiveness Engineering Graduate Program conducted jointly by the USAMC Intern Training Center and Texas A&M University. As such, the ideas, concepts and results herein presented are those of the author and do not necessarily reflect approval or acceptance by the Department of the Army.

This report has been reviewed and is approved for release. For further information on this project contact Mr. R.D. Cook, Intern Training Center, Red River Army Depot, Texarkana, Texas 75501.

R. D. Cook

R.D. Cook
Maintenance Effectiveness Engineering Program

For the Commander

ACCESSION for	
NTIS	<input checked="" type="checkbox"/>
DDC	<input type="checkbox"/>
USAI	<input type="checkbox"/>
JU	<input type="checkbox"/>
BY	
L. J. C. AVAILABILITY CODE	
A. AIL and/or SPEC.	
A	

James Z. Arnett
J. Arnett, Director, ITC.

ABSTRACT

Research Performed by Carl Royer

Under the Supervision of Dr. Ronald S. Morris

The purpose of this research report is to find a suitable test procedure for determining the reliability of prototype small caliber arms submitted to the U.S. Army by private contractors for engineering design tests.

In finding a solution, present reliability testing procedures are researched, and a suitable procedure is developed.

The developed procedure is limited by its underlying assumptions that there are three independent modes of failure, and the number of rounds between each type of failure has an underlying exponential distribution. Areas of future research and development are suggested, and recommended changes in present Army prototype small caliber arms testing techniques are made.

ACKNOWLEDGMENTS

The author wishes to express his appreciation to Dr. Ronald S. Morris for his guidance and proofreading of this paper, as well as his help in initiating this project. Appreciation is also expressed to Dr. Roger J. McNichols and Dr. Gerald A. Elliff for serving on my committee. A special thanks goes to Mr. Tom Brown for initiating the project as well as providing information needed to complete the report.

During the course of this work, the author was employed by the U.S. Army as a career intern in the AMC Maintainability Engineering Graduate Program. This author is grateful to the U.S. Army for the opportunity to participate in this program.

The ideas, concepts, and results herein presented are those of the author and do not necessarily reflect approval or acceptance by the Department of the Army.

CONTENTS

Chapter		Page
I	INTRODUCTION	1
II	BASIC ASPECTS OF RELIABILITY TESTING	4
III	LITERATURE REVIEW OF RELIABILITY	
	TESTING PROCEDURES	11
	AGREE Task Group III Procedure	11
	H-108 Procedures	14
	TR3 Procedures	16
	TR4 Procedures	18
	Procedures Based on the Gamma Distribution	21
	Attributes Sampling Plans	24
	Summary	25
IV	EVALUATION OF PRESENT ARMY PROTOTYPE TESTING	
	METHODS	26
V	TRUNCATED AND SEQUENTIAL LIFE TESTING	35
	Truncated Life Tests	36
	Sequential Life Tests	42
	Comparison of Sequential and Truncated Life	
	Testing	47
VI	PROPOSED TESTING PROCEDURE	51
VII	CONCLUSIONS AND RECOMMENDATIONS	57

FIGURES

Figure		Page
1	GENERAL FORM OF THE OPERATING CHARACTERISTIC CURVES	9
2	$E_{\theta}(r)$ VERSUS θ FOR SEQUENTIAL AND TRUNCATED TESTS . . .	48
3	$E_{\theta}(R)$ VERSUS θ FOR SEQUENTIAL AND TRUNCATED TESTS . . .	49
4	ILLUSTRATION OF EFFECT OF MODIFIED TEST PROCEDURE . . .	53

TABLES

Table		Page
1	PROJECT SAW COST INFORMATION	30
2	XM207E2 EDT WEAPON AND ROUNDS USAGE	31
3	STONER EDT WEAPON AND ROUNDS USAGE	32
4	TABLE FOR TRUNCATED TESTS	41
5	TABLE OF $E_{\theta}(r)$ FOR SEQUENTIAL TESTS	46

CHAPTER I

INTRODUCTION

During the past few decades the rate of technological development has increased enormously, causing the development and use of increasingly complex equipment. This increased complexity of equipment, along with its resulting greater costs, has spurred interest and developments in the field of reliability, not only in developing methods of establishing reliability requirements in the design stage of an item, but also in developing reliability testing techniques used to establish that the item meets its reliability requirements.

Reliability testing procedures can be expensive: environmental requirements, personnel, test instruments, and the value of the items put on test all contribute to the expense. An expense of testing which is sometimes neglected is the cost of the decision resulting from the test. As an example, if the military accepts a weapon with reliability somewhat less than what is indicated by a reliability test, then the cost of this incorrect decision includes increased maintenance time, repair time, downtime, and a decreased useful life. In developing an economic testing procedure, the direct testing costs should be balanced against the cost of an incorrect decision resulting from the test.

The purpose of this report is to develop a reliability testing procedure for use by the Army on prototype small caliber arms to determine whether they meet the specified reliability requirements.

The testing procedure developed is intended to be cost effective; that is, it seeks to minimize the total cost of testing, both direct testing costs and the costs resulting from the decision made on the basis of the test. The prototype arms being considered are submitted to the Army by several private contractors. These weapons then undergo engineering design tests by the Army, and, depending on the results of these tests, one or two of the prototypes are approved for further development and eventually production. Reliability is one of several criteria used to determine which weapons will continue development and which ones will be rejected. It can be easily seen that the decisions made at this stage of a weapon's development will have a tremendous effect throughout the remainder of its life cycle. The discussion following in Chapter II is a brief summary of the general theory of reliability applicable to this problem. Chapter III presents a broad consideration of the various types of reliability testing procedures found in the literature. Chapter IV describes and evaluates the present procedures used by the Army in their engineering design tests on prototype small caliber arms, and indicates why the present methods used to test reliability are unsatisfactory. Chapter V further discusses two types of reliability testing, truncated and sequential testing, and describes how they both may be applicable to the problem considered here. Chapter VI presents the recommended testing procedure, along with a numerical example of how it can be applied. Finally, Chapter VII contains some concluding remarks, as well as, a discussion of difficulties which may be encountered in implementing the

proposed procedure. Some recommendations for areas of further research pertaining to reliability testing of prototype armament are also suggested.

CHAPTER II

BASIC ASPECTS OF RELIABILITY TESTING

An essential element in the general growth of reliability activities in the past few years has been the development of various statistical reliability testing plans. Standard plans have been developed from essentially the same mathematical framework with different underlying assumptions and are basically similar procedurally, but have been cataloged in many different ways. A specific plan can be selected by selecting various sets of test parameters such as total testing time, number of items on test, maximum number of failures, etc. This chapter will present a brief summary of the general theory applicable to reliability tests.

Reliability is defined in the literature in various ways, but it is always expressed in some manner as the probability of performing without failure for a specified time under given conditions. In the statistical development of sampling plans the essential element in the definition is the fact that reliability is expressed as the probability of successful operation. What constitutes a failure and what the given conditions are must be specified. Plans have been developed for both cases in which time is and is not a criterion of successful operation.

Mathematically, then, the reliability function R can be expressed as

$$\begin{aligned}
 R &= P\{X_1 \leq \hat{X} \leq X_2\} = F(X_2; \theta_1, \theta_2, \dots) - F(X_1; \theta_1, \theta_2, \dots) \\
 &= R(\theta_1, \theta_2, \dots)
 \end{aligned}
 \tag{2.1}$$

where $F(X; \theta_1, \theta_2, \dots)$ is the cumulative distribution function of the random variable of interest X , X_1 , and X_2 are the limits which define success, and $\theta_1, \theta_2, \dots$ are the parameters of the distribution of X .

In the case where success is a simple attribute, the parameter of the reliability function is simply the probability of success and

$$R = R(\theta) = \theta = \text{probability of success.} \tag{2.2}$$

Most reliability testing plans are based on a definition of success which uses a time period of successful operation. In these cases reliability is a function of the required time of operation, t , and the parameters $\theta_1, \theta_2, \dots$ of the distribution of the time to failure, T . If the distribution of the random variable, time to failure, is known then the reliability function can be expressed precisely. Most reliability testing plans are based on the Weibull, gamma, or a special case of both, the exponential distribution. The reliability function for the Weibull distribution is given by

$$\begin{aligned}
 R &= R(t; \alpha, \theta) = P(T > t) = \int_t^{\infty} \alpha x^{\alpha-1} e^{-x^\alpha/\theta} dx \quad t, \alpha, \theta > 0 \\
 &= 0 \text{ elsewhere;}
 \end{aligned}
 \tag{2.3}$$

for the gamma distribution by

$$\begin{aligned}
 R &= R(t; \alpha, \theta) = P(T > t) = \int_t^{\infty} \frac{\theta^{-\alpha} x^{\alpha-1} e^{-(x/\theta)}}{\Gamma(\alpha)} dx \quad t, \alpha, \theta > 0 \\
 &= 0 \text{ elsewhere;}
 \end{aligned}
 \tag{2.4}$$

where α is equal to the shape parameter of the distribution,

and for the exponential distribution by

$$R = R(t; \theta) = P\{T > t\} = \int_t^{\infty} \frac{e^{-x/\theta}}{\theta} dx = e^{-t/\theta} \quad t, \theta > 0 \quad (2.5)$$

$$= 0 \text{ elsewhere.}$$

Because of the random nature of the variable of interest, reliability cannot be measured directly but rather can only be estimated. An estimate which is good in a statistical sense will have such properties as being maximum likelihood, unbiased, minimum variance, efficient, and sufficient. For example, let n items be placed on test, where all are tested to failure. Let t_i be the observed operating time to failure for the i^{th} item. The maximum likelihood estimate of the parameter θ is, for the exponential case,

$$\hat{\theta} = (1/n) \sum_{i=1}^n t_i, \quad (2.6)$$

for the one parameter Weibull distribution

$$\hat{\theta} = (1/n) \sum_{i=1}^n t_i^{\alpha}, \quad (2.7)$$

and for the gamma distribution

$$\hat{\theta} = (1/\alpha n) \sum_{i=1}^n t_i, \quad (2.8)$$

where α is equal to the shape parameter of the distribution.

Estimates can also be derived for the case when the test is terminated at the r^{th} failure, where $r \leq n$, and for a test with replacement; that is, when each item which fails is either replaced or repaired, and the test is terminated at either the r^{th} failure or else at a preassigned time T . For practical use a more meaningful

indication of the reliability of an item is the confidence interval estimate. This form of an estimate provides a range of values which contains the actual reliability, $\{R_1 \leq R \leq R_u\}$, with a specified degree of confidence that R lies in this range. Usually a confidence statement is of the form $P(R_1 \leq R \leq R_u) = \gamma$, where γ is the confidence level of the statement. Many times it will specify only the lower, one-sided confidence interval, i.e., $P(R_1 \leq R) = \gamma$. Lower confidence interval estimates are relatively easy to compute when only one parameter is unknown and the reliability is a function of only the unknown parameter, as is the case with the exponential and one parameter Weibull and gamma distributions. In the cases where more than one parameter is unknown exact confidence limits are not readily obtained and one must normally resort to approximations. Lloyd and Lipow (15)*, et al., indicate methods for computing approximate limits for the gamma and Weibull distributions where two parameters are unknown.

In practice, reliability tests are usually conducted to determine whether the true reliability exceeds some specified lower value rather than to make a specific estimate of the value of the reliability. In these cases it is more appropriate to use hypothesis testing procedures rather than confidence interval methods to obtain the desired results.

Hypothesis tests can usually be attained directly from a knowledge of the form of the confidence interval. Suppose it is desired

* Numbers in parentheses refer to numbered references in the List of References.

to test the hypothesis that the actual reliability R is at least some specified value R_s at a level of significance α , where α is defined by $P\{\text{concluding that } R < R_s \text{ when } R \text{ actually equals } R_s\} = \alpha$. A suitable procedure is to conduct the test, compute the $(1-\alpha)$ upper confidence limit R_u , compare with the specified minimum value R_s , and make a decision according to the following rule: If $R_s \leq R_u$, conclude $R_s \leq R$; but if $R_u < R_s$, reject the hypothesis that $R_s = R$ and conclude that $R < R_s$. In general, a hypothesis is accepted if the hypothesized value lies in the confidence interval, as demonstrated by Dellinger (5).

In statistical sampling the differentiating ability of a test is of paramount importance and is indicated by the operating characteristic curves, usually called O.C. curves. These curves are plots of the probability of accepting a hypothesis versus the actual value of the hypothesized quantity of interest. Figure 1 shows a typical set of O.C. curves plotted for various sample sizes n where the quantity of interest is reliability.

The O.C. curves can be a very useful tool in selection of or design of a test. Not only does it indicate the differentiating ability of a particular test, but also helps to determine the risk of drawing erroneous conclusions. As indicated by Figure 1 (page 9), increased sample size also increases the differentiating ability of a test, but the increased sample size also increases the testing cost. There must be a trade-off between the ability of a test to differentiate against an undesired level of reliability and the cost of

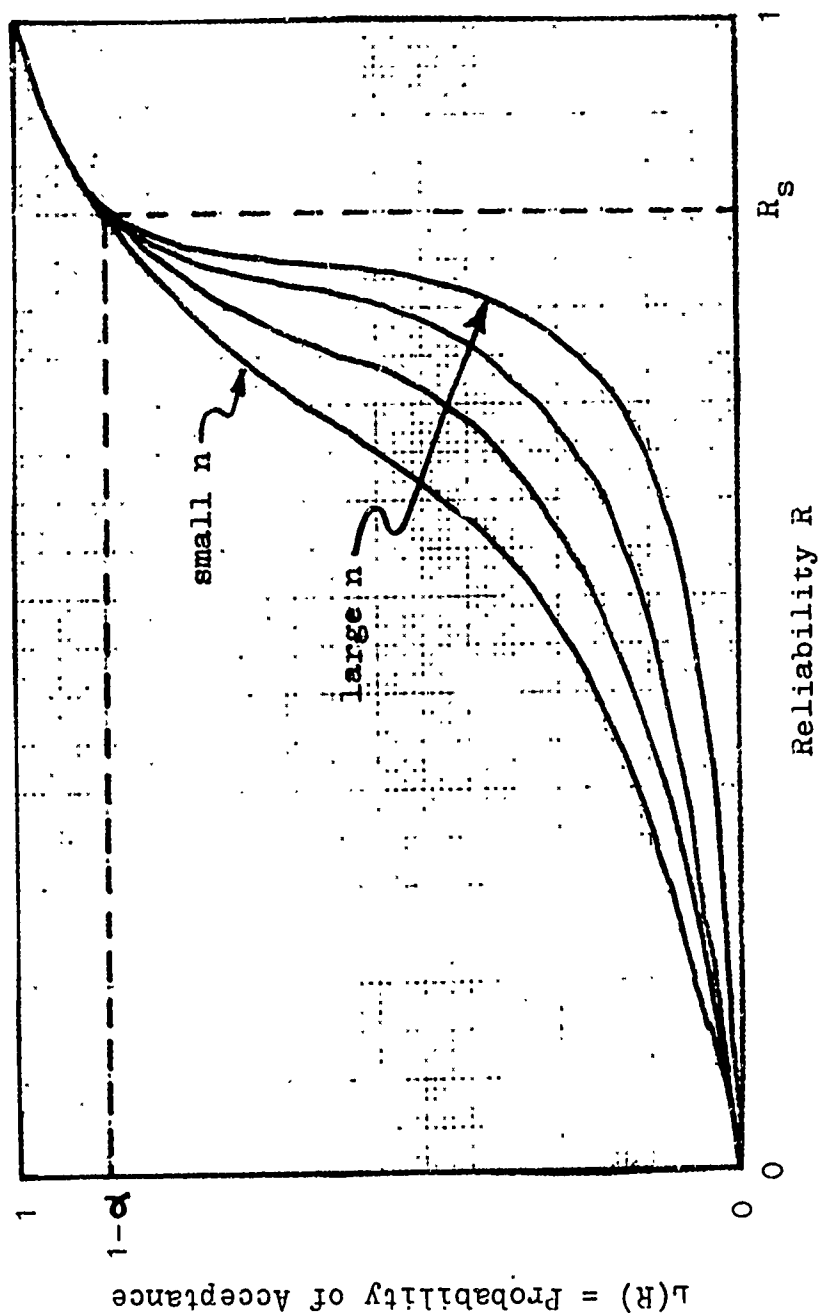


Figure 1 - General form of the operating characteristic curves.

testing.

A test can be completely specified by the selection of one point on the O.C. curve and the sample size, or by selecting two points on the O.C. curve. Dellinger (5), et al., recommends determining the test by selection of two points on the O.C. curve. One point is defined by assigning a level of the parameter which it is desired to accept with probability $1-\alpha$, giving $(R_s, 1-\alpha)$ in Figure 1 (page 9). The other point is selected for an undesirable level of the parameter for which it is desired to have a low probability (β) of acceptance. Figure 1 (page 9) illustrates the fact that even though all the tests represented by the O.C. curves pass through the point $(R_s, 1-\alpha)$, there is a large difference in their ability to reject values of R close to R_s .

Building on these results, the following chapter presents a literature review of reliability testing procedures, along with their underlying assumptions, and describes their methods of application.

CHAPTER III

LITERATURE REVIEW OF RELIABILITY TESTING PROCEDURES

As discussed previously, reliability testing plans are actually tests of hypothesis in which decision criterion are tabulated for a variety of test parameters. Underlying any such set of testing plans is a model which completely describes the conditions under which the tests are valid. Assumptions included in the test are the distribution of the random variable in the test, the method of test sampling, and sometimes limiting assumptions which allow for the use of approximations in designing the test. For example, a typical test may assume that the random variable time to failure is exponentially distributed, that samples are taken randomly from the population being considered, and that the random variable under consideration is homogeneous throughout the population. In this chapter various testing plans are discussed, and their underlying assumptions are noted. Any misapplication of these plans would result in invalid results.

Some models may complicate the derivation of the mathematical relationships to the degree that esoteric mathematical methods of approximations must be used. However, any test presented in this chapter can be represented by a standard Operating Characteristic curve, which, in many cases, is published as a part of the testing procedure.

AGREE Task Group III Procedure

The Advisory Group on Reliability of Electronic Equipment (AGREE)

was established by the Department of Defense in 1952. In 1957 this group published a report which contained procedures for testing for reliability, (17). These procedures were designed for testing components which could be repaired or replaced in service after failure and can be only partially applied to the problem of testing for the reliability of parts which are normally disposed of after failure. Since their publication, these procedures have been included in many contracts for military electronic equipment.

The procedure uses a sequential method for testing the hypothesis that the true mean time to failure θ of the items under test is greater than a specified value θ_1 . The tests are based on the work of Epstein and Sobel (10), in which n items are placed on test, with a failing item either replaced or repaired. The test is continued until a decision can be made by the sequential decision rules. The test assumes that the times between failures for each item on test is from an exponential distribution with parameter θ . The test procedure is completely specified by the selection of the contract specification of time to failure θ_1 . The operating characteristics of the test are such that α , the probability of a reject decision when $\theta = \theta_1$, is .10, and β , the probability of an accept decision for $\theta = 2/3\theta_1$, is also .10. The test results are independent of the number of items n put on test originally, so that the procedure provides a means of determining n so as to minimize the testing time needed to reach a decision given the constraints of production rates and the specified value of θ_1 .

The procedure also provides a means of relating the results of

the test to other untested items using an attributes sampling plan. Having identified each item tested as either satisfactory or unsatisfactory as far as reliability is concerned, a multi-level attribute sampling plan is applied to make decisions pertaining to the remainder of the month's production which has not been tested. If an accept decision is obtained from the multi-level attribute sampling plan, the remainder of the month's production is accepted without testing. Otherwise, every item will be tested prior to acceptance. The rules of the multi-level attributes sampling plan are as follows:

(1) If the first 22 items are found to be satisfactory, the remainder of the month's production is accepted without further testing.

(2) If a single unsatisfactory item is discovered, the next 37 items must be found satisfactory in order to accept the remainder of the month's production without further testing.

(3) If more than one unsatisfactory item is discovered, the next 50 items must be found satisfactory in order for the remainder of the month's production to be accepted without testing.

Once a process qualifies for acceptance without testing, the items produced by that process during the remainder of the month are accepted without testing. At the beginning of the following month, however, a small number of items must be tested. The exact number is not specified, but the document indicates that the number would be on the order of from four to ten items. If no unsatisfactory items are detected, all items will be tested according to the original

multi-level attributes sampling rules until the process again qualifies for acceptance without testing.

This procedure is not without criticisms. As noted by Dellinger (5), although the method refers to itself as a continuous attributes sampling plan, it does not provide a known level of protection as do most continuous sampling plans. Since the procedure does not require the continuous selection of items to be tested on a probabilistic basis to detect changes in quality level, overall risk cannot be determined statistically. The procedure is also designed only for use on a regular production process basis, and is clearly not applicable to the testing of pre-production prototype items. However, the basics by Epstein and Sobel, on which the procedure was developed, does contain ideas which can be applicable, as discussed later.

H-108 Procedures

Quality Control and Reliability Handbook H-108, Sampling Procedures and Tables for Life and Reliability Testing (18), was published by the Office of the Assistant Secretary of Defense in 1960 to provide standard sampling plans for life and reliability testing relative to established reliability requirements in government contracts. It has been distributed widely among government agencies and contractors, and is designed to be used as a standard reference by reliability and quality control engineers. The plans are derived from the results of Epstein and Sobel (10) and Epstein (8).

All of the plans are based on the assumption of an underlying

exponential distribution of time to failure

$$f(t) = \frac{1}{\theta} e^{-t/\theta} \quad (3.1)$$

where θ is the mean time to failure. A number of items n are randomly selected and placed on test, and observations are made of times to failure or the number of failures in a given time. It is assumed that the lot is homogeneous; that is, that all items in the lot have the same distribution of time to failure. As a result of the sampling plan a decision is made whether the lot meets specified reliability requirements. The plans presented are applicable to testing with and without replacement. For testing with replacement, failed items are either replaced or repaired, and the plan assumes that times between failures can be approximated by an exponential distribution with mean θ .

Three different types of procedures are presented in this handbook. The first type is made up of tests which are terminated upon occurrence of a preassigned number of failures. The second type consists of tests which are terminated at preassigned times, and the third type is a group of sequential test plans. An important feature of both the first and third types is that the operating characteristics of the plans are independent of the number of items placed on test. Consequently, for any particular O.C. curve there are a large number of plans each with its own sample size. The handbook provides a cost model which allows determination of the optimal sample size by balancing the cost of placing items on test with the cost of waiting for the decision resulting from the test. The selection of sample size does

not determine the amount of testing, only how the testing will be spread over time.

The plans of the second type mentioned above, where testing is terminated at a predetermined time, may be terminated earlier if a specified number of failures occurs. For these plans the selection of the O.C. curve and the termination time completely specify the plan, including sample size. This type plan is useful when practical considerations require that a test cannot exceed a set deadline.

All of the plans described in this handbook, along with their underlying theory, can be applied, with some modification, to the testing of prototype items. The greatest limitation of these plans is the assumption of an exponential distribution of times to failure. Although Davis (4), et al., has given evidence that many failures do occur randomly in time, and it has been assumed largely that failures will follow such a distribution, there is evidence that such an assumption is many times invalid. In the case of prototype small caliber arms, the assumption would be that the number of rounds between failures can be approximated by an exponential distribution. Further discussion of this follows in later chapters. In any case, any testing plan except for non-parametric tests can be subject to the criticism of assuming a distribution which may not be valid.

TR3 Procedures

Quality Control and Reliability Technical Report TR3, Sampling Procedures and Tables for Life and Reliability Testing Based on the

Weibull Distribution, (12), was published by the Department of Defense in September, 1961. Since the exponential distribution is but a special case of the Weibull distribution, this report can be considered a generalization of one form of the testing procedure presented for the exponential case in H-108. The testing plans are for the case where n items are placed on test and the test terminated when $r(r \leq n)$ failures occur or at time T_1 , whichever occurs first. The plans provide only for the case of non-replacement of failed items, and the underlying theory is further described by Goode and Kao (11). The criterion for the test is the mean time to failure. A three parameter model is used, however it is assumed that two of the parameters, γ and β , are known. The density function of the three parameter Weibull distribution is given by

$$f(t; \gamma, \beta, \eta) = (\beta/\eta) ((t-\gamma)/\eta)^{\beta-1} e^{-((t-\gamma)/\eta)^\beta} \quad (3.2)$$

where

γ = location parameter,

β = shape parameter, and

η = scale parameter.

The plans in TR3 are based on the assumption that $\gamma=0$; however, procedures are included for making a simple adjustment in the test plan if γ is not zero.

The test procedure given in TR3 is as follows:

- (1) Select at random a sample of n items from a lot size of N .
- (2) Place the sample items on life test for some preassigned period of time T .

(3) Denote by y the number of failures observed prior to time T .

(4) Accept the lot if $y \leq C$, where C is an acceptance number specified by the plan, but if $y > C$, reject the lot.

O.C. curves are not provided in TR3 for the available plan. However, tables are provided which give values of C and the minimum sample size n for various levels of protection. Also included are graphical procedures for estimating the parameters of a Weibull distribution, as well as determining if a given set of data came from a Weibull distribution.

The advantage of these plans are that they provide a reliability testing procedure which allows for a more general assumption of the underlying distribution of the parameter of interest. However, the plans are limited to testing without replacement only, which limits their efficiency.

TR4 Procedures

Quality Control and Reliability Technical Report TR4 Sampling Procedures and Tables for Life and Reliability Testing Based on the Weibull Distribution (13) was developed to supplement the plans presented in TR3. The only significant difference between the test plans in TR3 and those in TR4 lies in the parameter used to define reliability. In TR3, the mean life (mean time to failure) was used. In TR4, the hazard (failure) rate $z(t)$ is used.

In both TR3 and TR4 the plans are attribute sampling plans based on the probability p of a single item failing prior to the test termination time T . In TR4 the relationship between the hazard rate $z(t)$

and p is established as follows. The hazard rate is defined by

$$z(t) = f(t)/R(t) \quad (3.3)$$

where, in the case of the Weibull distribution,

$$f(t) = \beta/\eta (t/\eta)^{\beta-1} e^{-(t/\eta)^\beta} \quad \beta, \eta, t \geq 0 \quad (3.4)$$

and

$$R(t) = 1 - F(t) = e^{-(t/\eta)^\beta} \quad (3.5)$$

giving

$$z(t) = (\beta/\eta) (t/\eta)^{\beta-1}. \quad (3.6)$$

Multiplying by (t/β) gives

$$(tz(t)/\beta) = (t/\eta)^\beta \quad (3.7)$$

and it is seen that

$$p = F(t) = 1 - R(t) = 1 - e^{-(t/\eta)^\beta} \quad (3.8)$$

which, upon substitution, gives

$$p = 1 - e^{-tz(t)/\beta}. \quad (3.9)$$

This gives the desired relationship between $z(t)$ and the probability p of a single item failing prior to time t . This relationship can be used to obtain the value of p to use in the attributes plan by simply setting $t = T$, the test termination time.

The testing plans in TR4 are of the same general form as in TR3. There is one difference, however, in that it is assumed by all plans in TR4 that the test time will be the same as the value of t for which $z(t)$ is specified. When this is not the case, a table of conversion factors is provided which allows the conversion of the specified hazard rate for time t into an equivalent hazard rate for the desired test time. In addition to the plans based on a specified hazard rate

the report also contains instructions for sampling when the average hazard $m(t)$ is specified. The average hazard is defined as the cumulative hazard $M(t)$ divided by t . That is,

$$m(t) = M(t)/t = (1/t) \int_{-\infty}^t z(y) dy. \quad (3.10)$$

In the appendix to TR4 it is shown that for any suitable hazard function $z(t)$,

$$R(t) = e^{-\int_{-\infty}^t z(x) dx}. \quad (3.11)$$

Therefore, if $m(t)$ is specified, p can be determined without determining the form of $z(t)$,

$$p = 1 - e^{-tm(t)}. \quad (3.12)$$

From this relationship nonparametric tests can be devised using any suitable attributes sampling plan with p as the percent defective.

In practice, as noted by Dellinger (5), this form of test can be very desirable since no assumption need be made concerning the distribution of time to failure. Specifying reliability in terms of percent failure for some given time period would allow use of this plan as an attributes sampling plan to obtain a nonparametric reliability test. However, to accomplish this, the test time must equal the value of t for which $m(t)$ is specified. If one desires to test for a period of time other than the time for which $m(t)$ is specified it is necessary to use conversion tables in TR4 which results in an assumption of a Weibull distribution and a value for the parameter β . Also, if the time t for which $m(t)$ is specified is other than the required time of operation t' for which $R(t')$ is desired, the test results can be

applied to the calculation of $R(t')$ only after a form for the distribution of time to failure has been assumed. In other words, one can utilize a truly nonparametric test only if the time t for which $m(t)$ is specified is equal to both the test time and the time t' for which $R(t')$ is desired.

It is doubtful that reliability would be specified in terms of $m(t)$ when there is knowledge of the form of the distribution of time to failure. It would be more logical to specify reliability in terms of the unknown parameter of the distribution. Similarly, there is little reason to specify reliability in terms of the hazard rate ($z(t)$) if the distribution of times to failure is known. It is difficult to visualize any practical situation where the specification of reliability in terms of the hazard rate or average hazard would make much sense. Such specifications only confuse the basic issue of establishing the reliability of an item. For these reasons, in addition to those given for the TR3 report, this method shows little promise for the problem under consideration by this report.

Procedures Based on the Gamma Distribution

Gupta and Groll (14) discuss sampling plans which are a generalization to the gamma distribution of the truncated life tests without replacement for the exponential distribution in H-108. It is similar to the generalization of the same plans to the Weibull distribution in TR3 and TR4. A relationship is developed between the single unknown parameter of the distribution and the probability p of a single item failing prior to the end of the test termination time. Standard

attribute sampling procedures are then applied using p as the parameter of a binomial distribution. Testing procedures are identical to those in TR3.

It is assumed in these plans that times to failure are approximated as a random variable with a gamma distribution density function

$$f(t) = 1/\Gamma(\alpha) \{x/\theta\}^{\alpha-1} e^{-x/\theta} \text{ for } t, \alpha, \theta \geq 0. \quad (3.13)$$

The cumulative distribution function is then

$$F(t) = \int_0^t 1/\Gamma(\alpha) \{x/\theta\}^{\alpha-1} e^{-x/\theta} dx, \quad (3.14)$$

and the reliability function thus is

$$R(t) = 1 - F(t). \quad (3.15)$$

It is assumed that the parameter α is known, so that the probability p of a single item failing prior to time t , in terms of t and the unknown parameter θ , is

$$p = F(t, \theta) = (1/\Gamma(\alpha)) \int_0^t (x/\theta)^{\alpha-1} e^{-(x/\theta)} dx. \quad (3.16)$$

The mean of the gamma distribution, $E(t) = \alpha\theta$, is used as the acceptance criterion in these plans. The sampling plans are indexed by the ratio $t/\alpha\theta_0$, in which t is equal to the test termination time, α is equal to the known shape parameter of the distribution, and $\alpha\theta_0$ is a specified value of the mean time to failure $\alpha\theta$. The specified value of $\alpha\theta_0$ is a minimum value for which it is desired to reject the lot with a high probability, which places emphasis on the consumer's risk rather than the more common practice of emphasizing the producer's risk.

Several methods of selecting sampling plans are available,

including the method of choosing an O.C. curve and finding a matching plan. However, only a limited number of O.C. curves are presented.

The suggested method is as follows:

(1) Determine the value of the ratio $t/\alpha\theta_0$ and for p^* such that $P(\text{Rejecting the lot when } \alpha\theta = \alpha\theta_0) = p^*$. The time to terminate the test is arbitrary, and selecting $\alpha\theta_0$ and p^* results in choosing a point on the O.C. curve of the test corresponding to the consumer's risk.

(2) A table is provided which gives a family of plans corresponding to the ratio $t/\alpha\theta_0$ and p^* . Each plan is specified by the numbers C is equal to the acceptance number and n is equal to the minimum sample size.

(3) A value for $\alpha\theta$ is chosen for which it is desired to accept the lot with a probability of .95. Tabulated values are given for acceptance numbers C which have this desired property. The desired plan will then be the number of the family of plans selected in step 2 having the acceptance number found in step 3.

A rather broad range of plans is presented, with procedures included for the selection of plans based on their operating characteristics. The report doesn't give any guidance relative to the selection of an economic test termination time which minimizes test costs, as given in H-108. Neither is it indicated how to determine whether a set of failure data can be approximated by a gamma distribution. However, the procedures are useful in that they specify how to test for reliability when it is assumed that failure times are approximated by a rather general density function, the gamma. However, the tests

are again limited to testing without replacement.

Attributes Sampling Plans

In the area of quality control there are many attribute sampling plans which can be applied to the problem of reliability testing. It is necessary only that the test be conducted in which success or failure are possible outcomes; for example, when there is no required time of operation, such a test would be the operation of the device in the specified environment. In those cases where the definition of success requires operation for a specified time, an attribute test would be one in which the device is tested for the specified operation time and each trial judged success or failure depending on whether the device operated properly for the test period. Such tests are usually called nonparametric or distribution-free tests.

Where such tests can be applied, it is assumed that each device in the lot has the same probability of failure p and that reliability is simply $R = 1 - p$. The actual distribution of the number of failures x in a sample size n drawn from a lot size N is known to be hypergeometric, Duncan (6). However, when n is small compared to N , as is usually the case, a binomial distribution can be assumed with little error. Most attribute sampling plans are based on the binomial distribution.

The principal advantage of attribute sampling plans lies in the fact that one does not need to know anything about the underlying distribution of time to failure. The disadvantage is that more time is required to make a decision using an attributes plan than if one

used a plan based on a known distribution of time to failure with equivalent operating characteristics.

Summary

The documents discussed in this chapter are representative of the standard reliability testing techniques generally available today. A sample of other books and articles on the subject failed to present any techniques which are basically different from those considered in this chapter, although many minor variations of the techniques discussed can be found.

The general form of all plans were similar, patterned after standard statistical quality control plans developed in the last decade. The emphasis is on the mechanism of application, while basic assumptions are somewhat neglected and the statistical basis for each usually relegated to footnotes. However, they are easily applicable in many practical situations.

None of the plans are directly applicable to the problem of testing prototype small-caliber arms for reliability. All assume the random variable of interest is time to failure, and many assume a regular production process is ongoing. Except for the plans given in H-108, which will be further discussed in Chapter V, all plans are for tests without replacement, which makes them relatively useless for arms testing. The following chapter describes the testing procedure presently used by the Army to test prototype arms.

CHAPTER IV

EVALUATION OF PRESENT ARMY PROTOTYPE TESTING METHODS

This chapter presents a general discussion of the procedures presently used by the U.S. Army for testing prototype small caliber arms. Also discussed are the deficiencies of the present procedures.

In the early development of a small caliber weapon the Army receives sealed bids for development of a specific number of prototypes from private contractors. Of these bidders, as many as four or five are awarded contracts to develop prototypes. When delivered by the contractors these weapons undergo engineering design tests, which include testing for reliability, one of several factors used to determine which contractor will continue development of the weapon. Occasionally, two contractors continue developing their prototype models because each has certain desirable attributes not found in the other.

Experience has shown that three failure modes occur in the weaponry, Brown (1):

- (1) part breakage - the actual breakage of a part of the gun;
- (2) non-immediately clearable - gun jam which takes some period of time to clear;
- (3) immediately clearable - gun jam which can be cleared immediately.

When any assumptions at all are made about the basic underlying distribution of the number of rounds between failures, it has been assumed that each failure mode can be approximated by independent

exponential distributions, with each having its own mean number of rounds to failure θ . The use of the exponential distribution to approximate the number of rounds to failure for each type of failure assumes that each type of failure occurs independently only when an infrequent random environmental excursion, greater than some magnitude E_c , occurs in the environment E in which the weapon is operating. It is difficult to determine what, if any, data analysis had been performed to justify the assumption of the exponential distribution for each failure mode.

The Army contract usually contains a reliability requirement which specifies a minimum number of rounds between failures for each failure mode, although there is quite a bit of variation in the form of the specifications. The following example from the Light Weight Individual Weapon Systems, Final Report (2) illustrates one form used:

"These requirements establish reliability limits as follows:

- (1) No more than two stoppages clearable by immediate action in 1000 rounds (2000 rounds desirable)
- (2) No more than one stoppage not clearable by immediate action, but not requiring parts replacement in 6000 rounds (10,000 rounds desirable)
- (3) Not more than two malfunctions requiring parts replacement by operator in 6000 rounds (10,000 rounds desirable)."

The exact form of the contract requirement for reliability may vary rather extensively from contract to contract and between different weapon programs, but these basic malfunction modes are usually given

a required number of rounds between failures.

It is important to note that all of these specifications stipulate only a point estimate requirement and ignore that the number of rounds between failures is a random variable which may vary quite a bit in a manner described by the form of its actual underlying probability density function. Values for the α and β levels are not specified, and in fact the entire statistical nature of the random variable appears to have been ignored. A requirement that "10,000 rounds is desirable" is clearly of little use - as discussed previously, only a lower limit is usually specified practically, and any number exceeding this value is desirable.

Project SAW (Squad Automatic Weapon), which is underway at the time of this writing, provides an example of an on-going Army project in this area. Project SAW is the development by ARMCOM (U.S. Army Armament Command) of a 6 millimeter light machine gun, weighing less than 20 pounds, with a 200 round clip, capable of firing 500 rounds per minute, fully automatic. The present contract calls for each of 3 contractors to produce 12 guns. Forty thousand experimental rounds are available at a purchase cost of \$1 per round, with TECOM (U.S. Army Test and Evaluation Command) performing the actual testing at an additional cost to ARMCOM. Testing costs are determined for each individual weapon system. For small arms it usually ranges around \$50,000 for the use of ranges, personnel and other overhead. As mentioned, the experimental ammunition costs \$1 per round presently; this will be reduced about 50% in one year as the production process becomes more

efficient, and eventually will range around \$.10-\$.20 per round as full production commences. The present contract dictates delivery of 10 new usable weapons and two weapons which undergo a contractor demonstration test. Table 1 (page 30) gives the actual cost data associated with this stage of the program.

Table 2 (page 31) gives a summary of weapon and rounds usage during a typical engineering design test - in this case, from tests of the XM207E2 machine gun, Stoner EDT (19). As indicated, out of 10 weapons only two are used for testing for reliability during the accuracy and endurance test, firing a total of 72,667 rounds. The table indicates the various tests the armament is subjected to in addition to the reliability test. Many of the tests, such as the mud test, dust test or icing test are controlled condition tests where the operation of the weapon is tested under the extreme conditions it may be subjected to. Other tests are accuracy tests, and a third group of tests are such tests as sound level, smoke, and flash tests where it is desirable to minimize such environmental conditions resulting from firing the weapon. All of the tests are designed to check if the arms under test meet specified criteria. For example, in the cookoff test it may be desired to determine the number of rounds fired continuously from a cool gun which will cause a chambered round to "cook off" within 30 minutes. The requirement may be that the weapon will not cook off after firing less rounds than the determined critical cook off level (assuming stoppage on a chambered round).

Table 3 presents a summary of weapon and rounds usage during

	<u>Maremont</u>	<u>Philco-Ford</u>	<u>Rodman</u>
Base Line	\$427K	\$460K	*
First Overrun	\$51K	\$108K	*
Extension	\$99K	\$98K	*
Magazine	\$13K	\$17K	*
Second Overrun	\$0K?	\$68K	*
Total	<hr/> \$590K	<hr/> \$751K	<hr/> *

* No actual contract exists on Rodman (Rock Island Developed Fixture) but it will probably equal if not exceed the figures given.

Table 1 - Project SAW Cost Information.

TEST	Weapon Serial Number											
	Total Rounds on Weapon		003336	003337	003338	003339	003340	003341	003342	003343	003344	003333
2.3 Cookoff			X									
2.4 Water Spray Test			X	X								
2.5 Range Dispersion Test							X					
2.6 Attitude Test			X	X								
2.7 Dust Test			X	X								
2.8 Mud Test			X	X								
2.9 Sound Level Test									X	X		
2.10 Smoke Test										X		
2.11 Flash Test									X	X		
2.12 Belt Pull Test									X	X		
2.13 Icing Test												
2.14 Temperature-Humidity Test									X	X		
2.15 Low Temperature Test												
2.16 Unlubricated Weapon Test									X	X		
2.17 Accuracy and Endurance Test				X	X							
2.18 USMC Special Endurance Test											X	
2.19 Time-Displacement Curves												X

Table 2 - XM207E2 EDT Weapon and Rounds Usage.

WEAPON SERIAL NUMBER	TOTAL ROUNDS FIRED	APPLICABLE SUBTEST FROM TPR SAL-71-1-009																						
		CONTROLLED CONDITION TESTS												CAMP MCCOY TESTS										
		FUNCTIONAL CHECK	ACCURACY	WEAPON PERFORMANCE	RANGE LIFE	RELIABILITY & MAINTAINABILITY	STATIC DUST	RAIN	MUDDY WATER	SUNBURN TEMP.	HIGH TEMP.	HUMIDITY	ICING	UNLUBRICATED	SALT WATER IMERSION	SUSTAINED FIRING	RAYONET	GRAINDE LAUNCHING	ROUGH HANDLING	SAND	DEPRESSED & FLATTED FIRING	LONG RANGE ACCURACY	EFFECTIVE RANGE	MAXIMUM ORIGINATE
CARBINE 003200	35,160	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○					○	○	○	○
CARBINE 003201	4,443	○	○			○	○	○	○	○	○	○	○	○	○				○		○	○	○	○
CARBINE 003202	16,557	○	○	○			○	○	○	○	○	○	○	○	○	○	○		○	○		○	○	○
CARBINE 003203	1,570	○	○																	○				
CARBINE 003204	21,896	○	○	○	○																			
CARBINE 003249	2,263	○	○														○		○					
RIFLE 003205	4,810	○	○			○	○	○	○	○	○	○	○	○	○	○		○	○		○	○	○	○
RIFLE 003206	13,291	○	○	○			○	○	○	○	○	○	○	○	○	○	○			○				
RIFLE 003207	2,386	○	○																			○	○	○
RIFLE 003208	20,880	○	○	○	○																			
RIFLE 003209	31,867	○	○	○	○																			
RIFLE 003248	12,159	○	○	○		○		○										○	○					
M16A1 4392305	2,727	○	○				○	○	○	○	○	○	○	○	○	○			○	○		○	○	○
M16A1 4394839	35,230	○	○	○	○																			
M16A1 4395871	4,827	○	○			○	○	○														○	○	○

Table 3 - Stoner EDT Weapon and Rounds Usage.

another engineering design test on another weapon system, from the Stoner EDT (19). Comparison of Table 2 (page 31) with Table 3 (page 32) demonstrates the wide disparity between tests of similar systems. In this set of tests 15 weapons are used, with five weapons used for reliability and accuracy testing, with a total of 121,967 rounds fired. In these tests a maximum mean life of 36,000 rounds between failures was being tested. The variation between the tests summarized in Table 2 and those in Table 3 is found in all engineering design tests, as not only is there no set procedures for these tests, but the objectives of the tests vary among weapon systems due to the special requirements of different weapons.

The reliability tests used in these tests are of a very arbitrary nature. The number of both guns and rounds fired are not selected on a theoretical basis, but rather on a "rule of thumb" that a few weapons firing close to the maximum number of rounds mentioned in the specification should indicate whether or not the reliability of the weapon meets contracted values. The simple arithmetic average of the number of rounds between failures is used as the estimate of the parameter θ , mean rounds between failure, for each failure mode, with no consideration given to the distribution of the random variable.

As will be shown, ignoring the distribution of θ in these tests results in a much less efficient test, and may increase β , the probability of deciding that θ is greater than or equal to the contracted value when it is actually less than the specified value. Such an erroneous decision would greatly increase the overall life cycle cost of

the weapon.

CHAPTER V

TRUNCATED AND SEQUENTIAL LIFE TESTING

This chapter further considers two testing methods mentioned previously in terms applicable to testing of prototype small caliber arms. Throughout the chapter it is assumed that the number of rounds between failures for each failure mode follows an underlying exponential probability density function, with each failure mode having its own mean number of rounds between failures θ .

All of the testing procedures which were discussed previously were concerned with testing whether the mean time between failures met or exceeded a specified value. In the case of testing small caliber arms the test is not concerned with times between failures, but rather tests whether the mean number of rounds between failures meets or exceeds a specified value. The procedures for the two types of tests are similar, however. Testing small-caliber arms can be viewed as putting n guns on test, recording the number of rounds to the first failure, second failure, and so on, while repairing each weapon which fails (clear, unjam, or replace part, depending on which type of failure occurs) and returning the gun to test. This is analogous to putting n items on test, and testing with replacement, as discussed previously, except the testing is independent of time and concerned rather with the number of rounds fired.

This chapter presents the development of two feasible tests. The first test discussed is a truncated life test, where the test is

terminated either on the r^{th} failure or on a pre-assigned number of rounds R_0 . The other test is a sequential life test in which the test continues indefinitely until an accept or reject decision is reached on the basis of the number of failures within a number of rounds fired. Note that in both cases the term "life" refers to the number of rounds to failure rather than the number of time units.

Truncated Life Tests

A truncated life test proceeds as follows. With n guns placed on test, it is decided in advance that the experiment will be terminated at $\min(X_{r_0,n}, R_0)$, where $X_{r_0,n}$ is a random variable equal to the number of rounds fired before the r_0^{th} failure occurs, and R_0 is the truncation number of rounds, beyond which the experiment will not run. Both r_0 and R_0 are assigned before experimentation starts. If the experiment is terminated at $X_{r_0,n}$ (that is, if r_0 failures occur before R_0 rounds are fired), then the action in terms of hypothesis testing is the rejection of some specified null hypothesis. If the experiment is terminated at R_0 (that is, if the r_0^{th} failure does not occur before round number R_0 is fired), then the action in terms of hypothesis testing is the acceptance of some specified null hypothesis.

A previous development by Epstein (7) is applicable here, with a simple substitution of variable. The test is of the replacement type where the test is started with n guns and any gun that fails is replaced or repaired to like new condition, so that the number of guns under test is always n . Since the underlying distribution is

exponential with θ equal to the mean number of rounds between failures, the replacement or repair of failed items makes the life test a Poisson process with occurrence rate $\lambda_\theta = n/\theta$.

Thus, the probability of reaching a decision requiring exactly k failures is

$$P\{r=k|\theta\} = p(k;\lambda_\theta R_0) = (1/k!)e^{-nR_0/\theta} (nR_0/\theta)^k, \\ k = 0, 1, 2, \dots, r_0-1, \quad (5.1)$$

and

$$P\{r=r_0|\theta\} = 1 - \sum_{k=0}^{r_0-1} p(k;\lambda_\theta R_0). \quad (5.2)$$

The expected number of observations to reach a decision is

$$E_\theta(r) = \sum_{k=0}^{r_0} kP(r=k|\theta). \quad (5.3)$$

It can readily be shown that (5.3) can be written as

$$E_\theta(r) = \lambda_\theta R_0 \left\{ \sum_{k=0}^{r_0-2} p(k;\lambda_\theta R_0) \right\} + r_0 \left\{ 1 - \sum_{k=0}^{r_0-1} p(k;\lambda_\theta R_0) \right\}. \quad (5.4)$$

Equation (5.4) is a convenient form for calculation. For any pre-assigned n , R_0 and r_0 , $E_\theta(r)$ can be found easily from Molina's tables of the Poisson distribution (16) or other similar tables.

The expected number of rounds to reach a decision is given by

$$E_\theta(R) = (\theta/n)E_\theta(r). \quad (5.5)$$

The proof of (5.5) is as follows. It can be shown that

$$E_\theta(R) = E_\theta(X_{r_0,n}) + \sum_{k=0}^{r_0-1} p(k;\lambda_\theta R_0) \{R_0 - E_\theta(X_{r_0,n}|r=k)\} \quad (5.6)$$

and

$$E_{\theta}(X_{r_0,n}|r=k) = R_0 + E_{\theta}(X_{r_0-k,n}), \quad k=1,2,\dots,r_0-1. \quad (5.7)$$

Furthermore,

$$E_{\theta}(X_{r_0-k,n}) = E_{\theta}(X_{r_0,n}) - E_{\theta}(X_{k,n}) = (r_0 - k)\theta/n, \quad 1 \leq k \leq r_0, \quad (5.8)$$

since the unconditional expected number of rounds fired to get exactly s failures in a replacement situation is

$$E_{\theta}(X_{s,n}) = s\theta/n \quad (5.9)$$

for any integer s . Substituting (5.7) and (5.8) into (5.6) gives (5.5).

It can be shown that (5.5) can be rewritten as

$$E_{\theta}(R) = \sum_{k=1}^{r_0} P(r=k|\theta) E_{\theta}(X_{k,n}) \quad (5.10)$$

and the unconditional expected number of rounds fired is

$$E_{\theta}(X_{k,n}) = k\theta/n. \quad (5.11)$$

Suppose the truncation rule is such that the hypothesis $H_0: \theta = \theta_0$ is accepted if $\min(X_{r_0,n}, R_0) = R_0$, that is, the number of rounds fired required to observe $X_{r_0,n}$ is more than R_0 . Then if $L(\theta)$ is defined as the probability of accepting $\theta = \theta_0$ when θ is true, it follows that

$$L(\theta) = 1 - P\{r=r_0|\theta\} \quad (5.12)$$

where $P\{r=r_0|\theta\}$ is given by Equation (5.2).

Let $X_{k,n}$ be the number of rounds to the k^{th} failure (whether it be an original item or a replacement item) measured from the beginning of the experiment. It has been shown by Epstein and Sobel (9), et al., in the replacement case, that if one starts with n items, then the

"best" region of acceptance, in the Neyman-Pearson sense, for testing a hypothesis H_0 that $\theta = \theta_0$ against alternatives of the form $\theta = \theta_1$ ($\theta_1 < \theta_0$), based on the first r number of rounds to failure $X_{1,n}, X_{2,n}, \dots, X_{r,n}$, is of the form $\hat{\theta}_{r,n} > C$, where

$$\hat{\theta}_{r,n} = nX_{r,n}/r. \quad (5.13)$$

It follows that the region of acceptance for H_0 is of the form $X_{r,n} > C^* = rC/n$. Use of $X_{r,n} > C^*$ as a region of acceptance means in words that the test is terminated at $\min(X_{r,n}, C^*)$ with acceptance of H_0 if truncation occurs at $X_{r,n}$. This is precisely the test treated above with $r=r_0$ and $C^* = R_0$.

So far formulae have been given for the O.C. curve, expected number of rounds used and expected number of failures encountered in the course of reaching a decision for any preassigned n , R_0 , and r_0 . The following is a procedure for finding the appropriate truncated test (that is, for finding r_0 and n) when the truncation number of rounds is preassigned and the O.C. curve is required (for preassigned type I error, α , and type II error, β) to be such that $L(\theta_0) \geq 1-\alpha$ and $L(\theta_1) \leq \beta$. Here θ_0 and θ_1 are preassigned with $\theta_0 > \theta_1$.

Epstein and Scbel (9) have shown that the best acceptance region of size α for the hypothesis $\theta = \theta_0$ (against any alternative $\theta_1 < \theta_0$), based on the first r out of n failures, for preassigned r and n , is

$$\hat{\theta}_{r,n} > C = \theta_0 X_{1-\alpha}^2 (2r)/2r \quad (5.14)$$

where $\chi^2_{1-\alpha}(2r)$ is a chi-square variable with $2r$ degrees of freedom.

In order that the test have an O.C. curve for which $L(\theta_0) = 1-\alpha$ and $L(\theta_1) \leq \beta$, we need to choose r suitably. The appropriate values of r for certain values of α, β , and θ_0/θ_1 are given in Table 4. For values of α, β , and θ_0/θ_1 not given in the table, the appropriate r to use is the smallest integer r such that $\chi^2_{1-\alpha}(2r)/\chi^2_{\beta}(2r) \geq \theta_1/\theta_0$.

It is now an easy matter, in the replacement case, to find a truncated test meeting the conditions prescribed above. The appropriate r_0 is given by the values in Table 4, and

$$R_0 = C^* = rC/n = \theta_0 \chi^2_{1-\alpha}(2r)/2n. \quad (5.15)$$

Since n must be an integer, the equality can be satisfied only approximately. For all practical purposes n can be chosen as

$$n = \{\theta_0 \chi^2_{1-\alpha}(2r_0)/2R_0\} \quad (5.16)$$

where $\{x\}$ means the greatest integer $\leq x$. It is important to note the appropriate number of guns on test n , for fixed α and β , is inversely proportional to the truncated number of rounds R_0 fired by each gun. Thus, for example, to reduce the truncation number of rounds fired by each weapon on test by a factor of 2 requires doubling n , the number of guns on test. This has clear implications in the case of prototype testing, where a maximum n is set by the number of prototypes available, requiring enlarging R_0 in order to achieve the desired α and β levels. It is clear from Equation (5.16) that tabulated values of

θ_0/θ_1	$\alpha = .01$			$\alpha = .05$			$\alpha = .10$		
	$\beta = .01$	$\beta = .05$	$\beta = .1$	$\beta = .01$	$\beta = .05$	$\beta = .1$	$\beta = .01$	$\beta = .05$	$\beta = .1$
3/2	136 110.4	101 79.1	83 63.3	95 79.6	67 54.1	55 43.4	77 66.0	52 43.0	41 33.0
2	46 31.7	35 22.7	30 18.7	33 24.2	23 15.7	19 12.4	26 19.7	18 12.8	15 10.3
5/2	27 16.4	21 11.8	18 9.62	19 12.4	14 8.46	11 6.17	15 10.3	11 7.02	9 5.43
3	19 10.3	15 7.48	13 6.1	13 7.7	10 5.43	8 3.98	11 7.02	8 4.66	6 3.15
4	12 5.43	10 4.13	9 3.51	9 4.7	7 3.29	6 2.61	7 3.9	5 2.43	4 1.75
5	9 3.51	8 2.91	7 2.33	7 3.29	5 1.97	4 1.37	5 2.43	4 1.75	3 1.1
10	5 1.28	4 .823	4 .823	4 1.37	3 .818	3 .818	3 1.10	2 .532	2 .532

Table 4 - Values of r (upper numbers) and of $\sqrt{1-d(2r)/2}$ (lower numbers) for truncated tests, from Epstein (7).

$\chi^2_{1-\alpha} (2r_0)/2$ would be useful. These are given below the associated r_0 in Table 4. The appendix contains alternate tabular forms of these results. Appendix A contains tabulated values of R_0/θ_0 and r for values of $\theta_1/\theta_0, \alpha$ and β . Appendix B contains tabulated values of R_0/θ_0 for various O.C. curves and r_0 . Both appendixes are modified from H-108 (18).

As noted by Epstein (7), the O.C. curve of the test $\min\{X_{r_0, n}, R_0\}$, where r_0 is given by Table 4 and n by Equation (5.16), is such that $L(\theta_0) \geq 1-\alpha$, but in some cases $L(\theta_1)$ may be slightly greater than β . This can be avoided in either of two ways. One way is to give the experimenter the freedom to use, instead of R_0 , the slightly larger truncation number $R_0' = \theta_0 \chi^2_{1-\alpha} (2r)/n$; the test based on $\min\{X_{r_0, n}, R_0'\}$ will have $L(\theta_0) = 1-\alpha$ and $L(\theta_1) \leq \beta$. The other way is to use $n+1$ items throughout the test, and to use, instead of R_0 , the slightly smaller truncation number $R_0'' = \theta_0 \chi^2_{1-\alpha} (2r_0)/(n+1)$. The test based on $\min\{X_{r_0, n+1}, R_0''\}$ will have $L(\theta_0) = 1-\alpha$ and $L(\theta_1) \leq \beta$. In most cases it will make no difference which procedure is used.

Sequential Life Tests

As with the truncated life test discussed above, the aim of the sequential life test is to test the simple hypothesis $H_0: \theta = \theta_0$, against the alternative $H_1: \theta = \theta_1$, where $\theta_1 < \theta_0$, with Type I and II errors equal

to preassigned values α and β , respectively. Again, n items are put on test, with failed items repaired or replaced. The test can be terminated either at failure times with rejection of H_0 , or at any time between failures with acceptance of H_0 . Since abnormally long intervals furnish "information" in favor of H_0 and abnormally short intervals furnish "information" in favor of H_1 , these features are not only reasonable but actually desirable.

A previous development by Epstein and Sobel (10) is useful in this case, utilizing a transformation of the random variable from time to rounds fired by the weapon. In their report they make use of Wald's results on sequential analysis (20) where decisions are made continuously.

During the life test, information is available continuously, so that a continuous analogue of the sequential probability ratio test of Wald can be used. The decision as the test is run depends on

$$B < (\theta_0/\theta_1)^r \exp\{-(1/\theta_1 - 1/\theta_0)V(R)\} < A \quad (5.17)$$

where B and A are constants, depending on α and β , such that $B < 1 < A$. The decision to continue experimentation is made as long as the inequality (5.17) holds. At the time the experiment is stopped, if the first inequality in (5.17) is violated, H_0 is accepted; if the second inequality is violated, H_1 is accepted. As in Wald's case, the test obtained by setting $B = \beta/(1-\alpha)$ and $A = (1-\beta)/\alpha$ is a satisfactory solution of the problem from a practical point of view.

In Equation (5.17), $V(R)$ is a statistic which can be interpreted as the total number of rounds fired by all n guns on test at round

number R , with R measured from the beginning of the test. Clearly, in the replacement case,

$$V(R) = nR. \quad (5.18)$$

To graph the data continuously versus the number of rounds fired, it is convenient to write Equation (5.17) in the form

$$-h_1 + rs < V(R) < h_0 + rs \quad (5.19)$$

where h_0, h_1 , and s are positive constants given by

$$h_0 = -\log B / \{(1/\theta_1) - (1/\theta_0)\}, \quad (5.20)$$

$$h_1 = \log A / \{(1/\theta_1) - (1/\theta_0)\}, \quad (5.21)$$

$$s = \log(\theta_0/\theta_1) / \{(1/\theta_1) - (1/\theta_0)\}. \quad (5.22)$$

Further, as shown by Wald, the O.C. curve is given approximately by a pair of parametric equations

$$L(\theta) = (A^h - 1) / (A^h - B^h) \quad (5.23)$$

and

$$\theta = \{(\theta_0/\theta_1)^h - 1\} / \{h(1/\theta_1 - 1/\theta_0)\}, \quad (5.24)$$

by letting the parameter h run through all real values. The values of $L(\theta)$ at the five points $\theta=0, \theta_1, s, \theta_0$, and ∞ enable one to sketch the entire curve. These values are respectively $0, \beta, \log A / (\log A - \log B), 1-\alpha$, and 1 .

Let $E_\theta(r)$ equal the expected number of observations required to reach a decision when θ is the true parameter value. Since the logarithm of the middle expression in Equation (5.17) is either $\log B$ or $\log A$ at the time experimentation stops, we have, neglecting only the

excess over $\log A$,

$$E_{\theta}(r) \log(\theta_0/\theta_1) - E_{\theta}(V(R)) \{1/\theta_1 - 1/\theta_0\} \sim L(\theta) \log B + \{1 - L(\theta)\} \log A \quad (5.25)$$

It can be shown that

$$E_{\theta}(V(R)) = \theta E_{\theta}(r). \quad (5.26)$$

Therefore, from Equations (5.25) and (5.26),

$$E_{\theta}(r) \sim \frac{L(\theta) \log B + \{1 - L(\theta)\} \log A}{\log(\theta_0/\theta_1) - \theta(1/\theta_1 - 1/\theta_0)} = \frac{h_1 - L(\theta)(h_0 + h_1)}{s - \theta}, \quad \theta \neq s, \quad (5.27)$$

and

$$E_{\theta}(r) \sim \frac{-\log A \log B}{\{\log(\theta_0/\theta_1)\}^2} = \frac{h_0 h_1}{s^2}, \quad \theta = s. \quad (5.28)$$

If $k = \theta_0/\theta_1$, the approximate values of $E_{\theta}(r)$ are simplified for $\theta = \theta_1$, s , and θ_0 . They are

$$E_{\theta_1}(r) \sim \{\beta \log B + (1 - \beta) \log A\} / \{\log k - (k - 1)/k\}, \quad (5.29)$$

$$E_s(r) \sim -\log A \log B / (\log k)^2, \quad (5.30)$$

$$E_{\theta_0}(r) \sim \{(1 - \alpha) \log B + \alpha \log A\} / \{\log k - (k - 1)\}. \quad (5.31)$$

In Table 5, (page 46), $E_{\theta}(r)$ is given for $\theta = 0, \theta_1, s, \theta_0, \infty$, for $k = 3/2, 2, 5/2, 3$, and for $\alpha = .01, .05$, and $\beta = .01, .05$. The formula for $L(\theta)$ and $E_{\theta}(r)$, given by Equations (5.24), (5.27) and (5.28) respectively, are approximations to the actual $L(\theta)$ and actual $E_{\theta}(r)$ arising from the use of the semi-continuous sequential decision rule specified by the inequalities (5.17). Epstein and Sobel (10) discuss the accuracy of the approximations and give formulae for the bounds of $L(\theta)$ and

$k = \theta_0/\theta_1$		3/2		2		5/2		3	
α		.01	.05	.01	.05	.01	.05	.01	.05
θ	β								
0	.01	11	7	7	4	5	3	4	3
	.05	11	7	7	4	5	3	4	3
θ_1	.01	62.4	40.3	23.3	15.1	14.2	9.2	10.4	6.74
	.05	60.4	36.7	22.6	13.7	13.8	8.38	10.1	6.14
s	.01	128	82.7	43.9	28.3	25.1	16.2	17.5	11.3
	.05	82.7	52.7	28.3	18.0	16.2	10.3	11.3	7.18
θ_0	.01	47.6	44.2	14.7	13.6	7.71	7.16	5.0	4.63
	.05	30.8	28.0	9.48	8.64	4.99	4.54	3.23	2.94
∞	any	0	0	0	0	0	0	0	0

Table 5 - Approximate values of $E_{\theta}(r)$ for sequential tests for various $\theta_0/\theta_1, \alpha, \beta$, from Epstein and Sobel (10).

$E_{\theta}(r)$.

Comparison of Sequential and Truncated Life Testing

From a purely statistical point of view, the sequential test is a more efficient test, in that a decision is reached in less time with usually a smaller total number of rounds fired and with fewer failures occurring. Figure 2 (page 48) shows an example of the expected number of failures versus θ for both the truncated and sequential test. As is shown, the sequential test on the average will reach a decision with less failures. Figure 3 (page 49) shows an example of the expected number of rounds fired per gun for the sequential and truncated tests. Again, the sequential test requires fewer rounds than the truncated test, with the truncated test requiring considerably more rounds for larger values of θ . Since, in prototype testing, experimental rounds are used which can cost on the magnitude of \$1 per round, the use of less rounds can represent a considerable cash savings. It also reduces the length of time needed to conduct the test, assuming the firing rate is the same in both tests.

However, while statistically the sequential test may appear the better of the two, practical considerations make it difficult to utilize. First, the number of rounds required for the test is not known in advance, since it terminates only when one of the inequalities in Equation (5.17) is satisfied. This problem can be avoided by use of a truncated sequential test as discussed in H-108 (18), where if the test continues to a certain point, a decision is made on the basis of

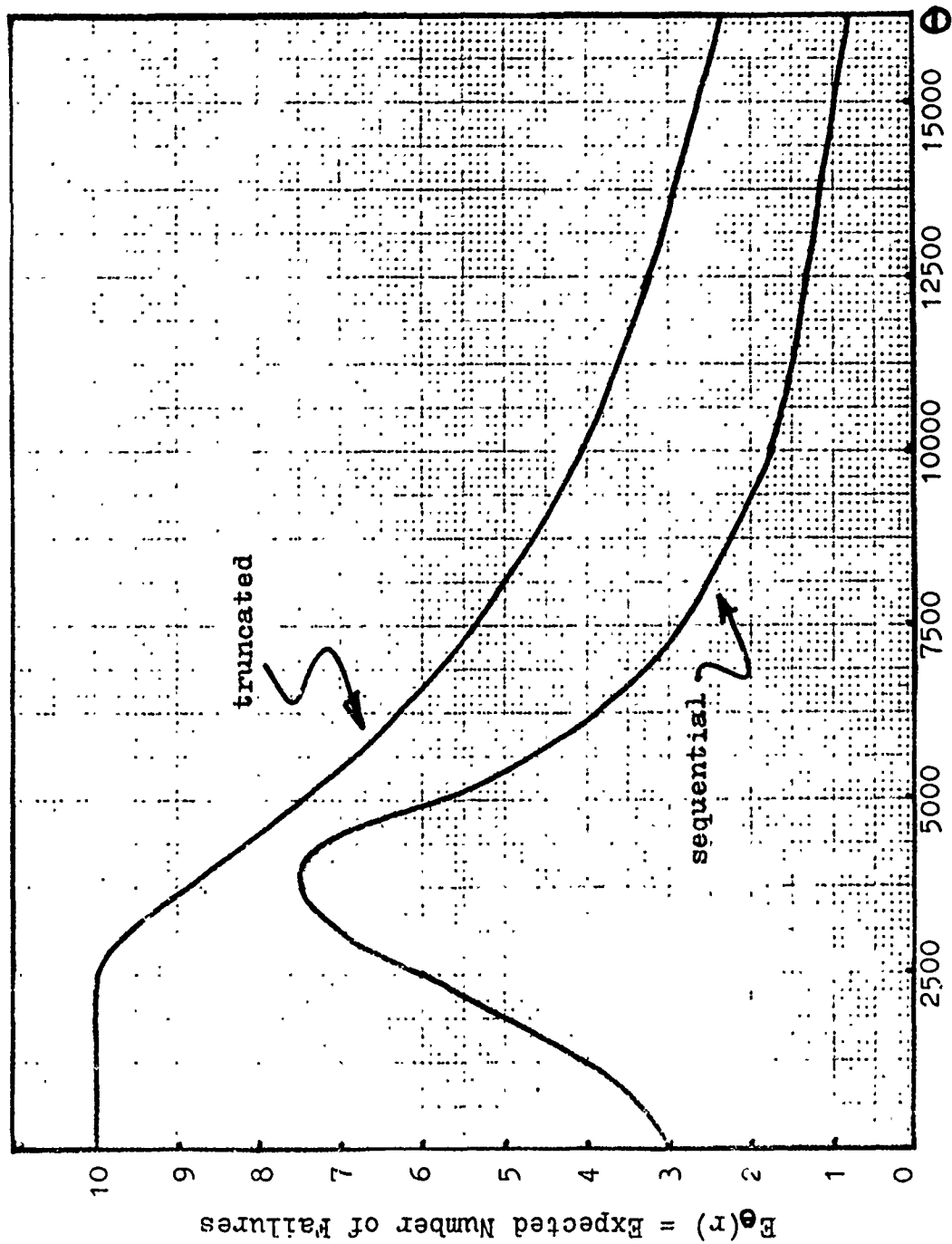


Figure 2 - $E_{\Theta}(r)$ vs. Θ for sequential and truncated tests with replacement, for $\alpha = \beta = .05, \Theta_0 = 7500, \Theta_1 = 2500, n = 100$.

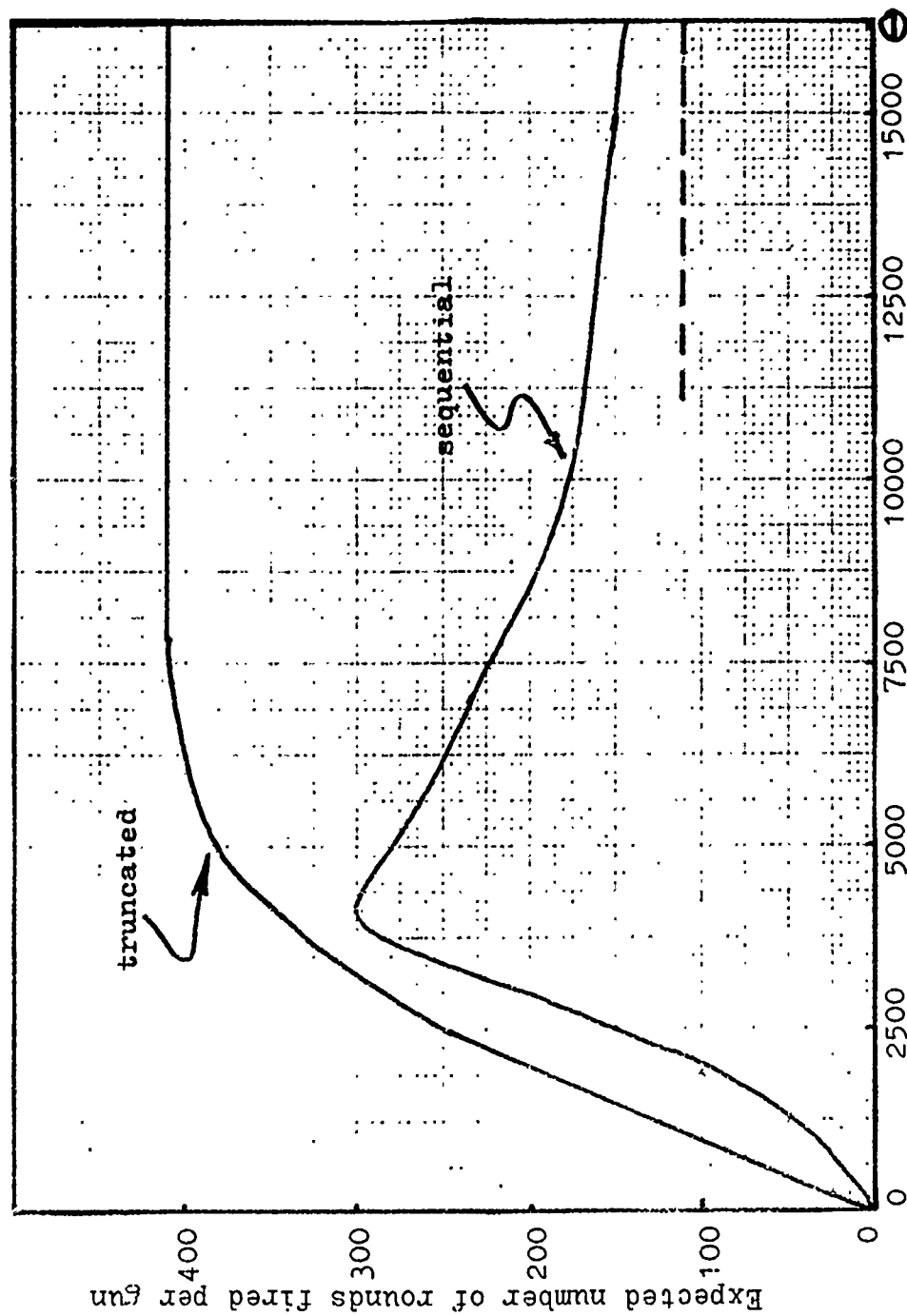


Figure 3 - $E_{\theta}(R)$ vs. θ for sequential and truncated tests with replacement, for $\alpha = \beta = .05, \theta_0 = 7500, \theta_1 = 2500, n = 100$. Dashed line gives value approached asymptotically as $\theta \rightarrow \infty$.

the number of failures which have occurred up to that point. However, the number of rounds to this truncation point is considerably greater than the expected value, so that if the truncation number of rounds was ordered by the testing facility, it would be expected that many of the rounds would not be used, wasting the money spent on the unused rounds.

A second difficulty of implementing this method is that all guns would have to be tested simultaneously and fired at the same rate. This would mean that, for n guns on test, n ranges and n riflemen would be necessary, plus failure recording personnel. Firing n guns at exactly the same rate would clearly be difficult, plus would mean that, when a failure occurs in any of the guns, all weapons would have to halt firing until the failed gun was repaired.

A third difficulty in implementing the procedure is the constant test monitoring necessary, where the results must be constantly recorded until one of the inequalities in Equation (5.17) is met and a decision made.

Although the sequential testing method would, on the average, use fewer rounds to reach a decision, it is felt that additional administrative and personnel costs would more than cancel any savings resulting from using fewer rounds. Because the truncated life test does not have these difficulties, it is felt a modified form of it would be the best testing method. The proposed procedure is demonstrated with a numerical example in the following chapter.

CHAPTER VI

PROPOSED TESTING PROCEDURE

The reliability testing procedure recommended by this report is described in this chapter, with an example presented to clarify its use. The recommended procedure is basically the truncated life testing procedure discussed in the previous chapter, modified as follows.

Rather than having all weapons firing simultaneously, it would be more practical to have the weapons fired one at a time for R_0 rounds. Assuming the environment in which each weapon is tested is the same, then it should make no difference whether the weapons are fired simultaneously or one at a time. The reason such a method of firing is desirable is that there are a limited number of test ranges and personnel available, with usually 2-3 personnel performing the actual testing, and simultaneous testing would be quite difficult. However, individual testing of the weapons results in the additional test cost of additional rounds used if the actual mean number of rounds is less than the specified value. In other words, the cost of a reject decision will be slightly higher. This is because ideally all n weapons would halt fire when the r_0^{th} failure occurs, with each weapon saving the number of unfired rounds equal to the difference between the round number R on which the r_0^{th} failure occurs and R_0 , the truncation number of rounds at which the test would have otherwise stopped. However, when tested individually each weapon must be fired R_0 rounds each until the r_0^{th} failure occurs which, if it does occur, will

probably occur with the n^{th} weapon, due to the small n on test. In other words, $(n-1)(R_0-R)$ rounds will be wasted. However, if the mean number of rounds between failures meets or exceeds the specified number being tested for, no rounds are wasted, since all n guns would have fired R_0 rounds even if tested simultaneously. Figure 4 shows this graphically for a particular example.

Testing small caliber arms amounts to testing three separate hypotheses. Each of the three failure modes is tested to determine whether the mean number of rounds between failures meets the specified value. In other words, it is hypothesized that $H_0: \theta = \theta_0$ against an alternative $H_1: \theta = \theta_1$, where $\theta_1 < \theta_0$, for each type of failure.

Let θ_{ic} equal the mean number of rounds between immediately clearable failures, θ_{nc} equal the mean number of rounds between non-immediately clearable failures, and θ_{pb} equal the mean number of rounds between part breakages. When these values are specified by the Army, θ_{ic} is always less than θ_{nc} and θ_{pb} , while θ_{pb} usually equals to or is greater than θ_{nc} . When testing the three hypotheses simultaneously, the largest θ being tested, probably θ_{pb} , will determine the value n and R_0 used in the test, since these numbers would be smaller for test-for a smaller θ .

In conducting the test, when testing for the largest θ for α and β fixed, more rounds than necessary are fired to test a lower θ at the same levels of α and β . If the test for the lower value of θ uses all R_0 rounds fired in testing for the larger θ , then the α and β levels for the test of the smaller θ will be larger than those set for testing

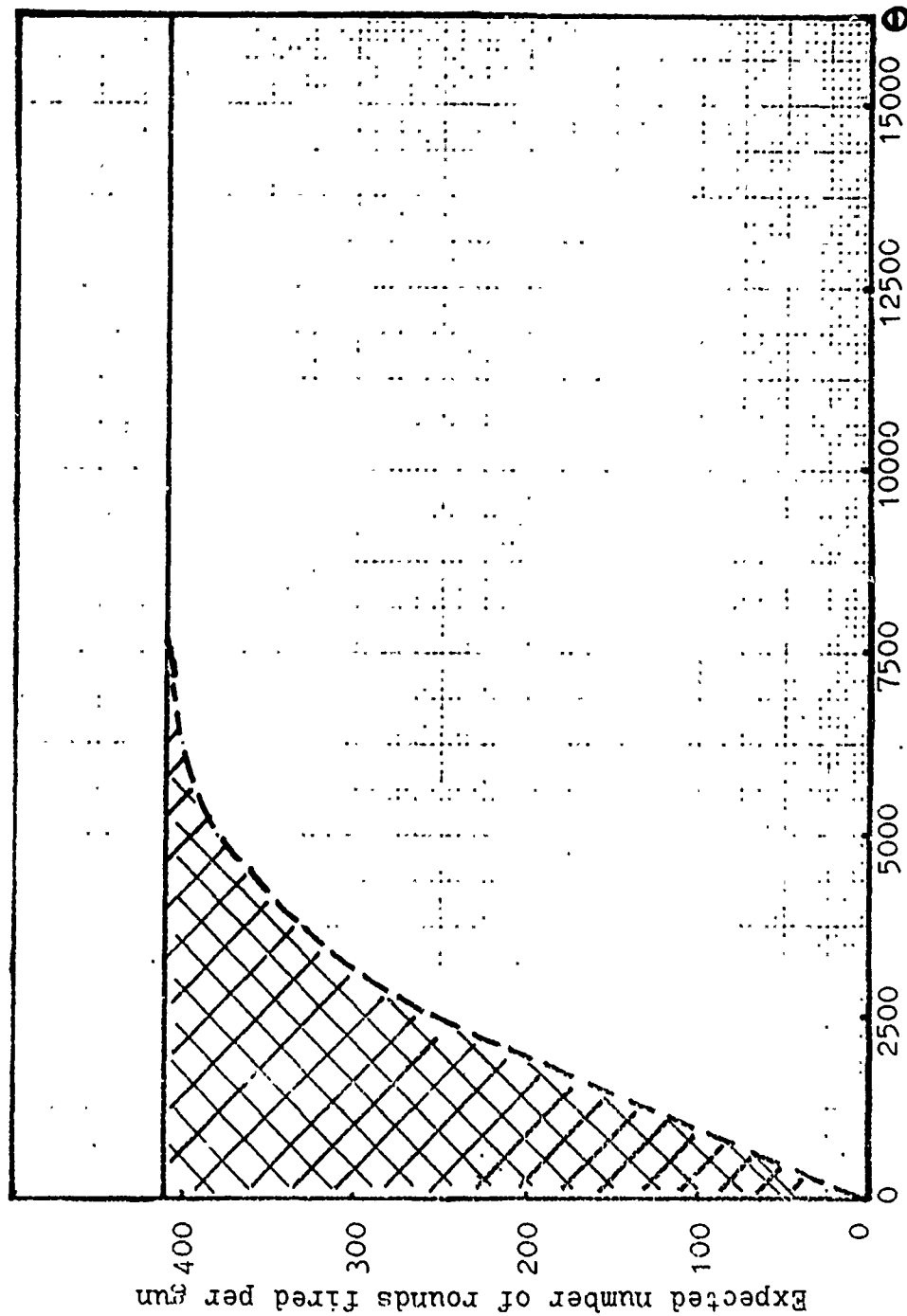


Figure 4 - Illustration of effect of modified testing procedure, for $\alpha = .05$, $\Theta_0 = 7500$, $\Theta_1 = 2500$, $n = 100$. Shaded area indicates possible number of wasted rounds per weapon under test.

the larger θ . The following numerical example should clarify these comments.

Suppose an Army specification on a prototype small caliber arm is as follows:

(1) Malfunction clearable by immediate action not to exceed 1 per 1000 rounds, with $\alpha \leq .10$, $\beta \leq .10$.

(2) Malfunction not clearable by immediate action but not requiring parts replacement not to exceed 1 per 6000 rounds, with $\alpha = .10$, $\beta = .10$, $\theta_0/\theta_1 = 3$.

(3) Malfunction requiring parts replacement not to exceed 1 per 6000 rounds, with $\alpha = .10$, $\beta = .10$, $\theta_0/\theta_1 = 3$.

This is the general form of specification suggested by this report. It is noticeably different from the form presently used due to the inclusion of α, β , and θ_0/θ_1 values.

The specification stipulates $\theta_{ic} = 1000$ rounds, $\theta_{nc} = \theta_{pb} = 6000$ rounds, for $\alpha \leq .10$, $\beta \leq .10$, $\theta_0/\theta_1 = 3$. Suppose 10 prototypes are delivered from each contractor and let all 10 be used in the test, fixing $n=10$.

To find r_0 , find the smallest r such that $\chi^2_{1-\alpha}(2r)/\chi^2_{\beta}(2r) \geq \theta_1/\theta_0$.

In this case, this is

$$\frac{\chi^2_{.9}(2r)}{\chi^2_{.1}(2r)} \geq 1/3. \quad (6.1)$$

Using tables of the chi-square distribution gives

r	$\frac{\chi^2_{.9}(2r)/\chi^2_{.1}(2r)}{2}$
1	.0458
3	.207
5	.3043
6	.3398

which gives $r_0=6$. With $n=10$, $\theta=6000$, and using Equation (5.16) gives

$$10 = \{18912/R_0\}$$

which sets $R_0=1891$ rounds. Therefore, for testing both θ_{pb} and θ_{nc} , this gives $r_0=6$, $n=10$, $R_0=1891$ rounds for the α, β , and θ_0/θ_1 specified. However, it should be understood that the $r_0=6$ is for testing θ_{pb} and θ_{nc} only, not θ_{ic} . Each θ has its own associated r_0 , and r_0 is the same for both θ_{pb} and θ_{nc} only because $\theta_{pb}=\theta_{nc}$. If it is desired that $\alpha=.1$ and $\beta=.1$ for testing for θ_{ic} as well, then use of Equation (5.16) gives.

$$10 = \{1000 (\chi^2_{.9}(2r_0)/(2)(1891))\}$$

giving

$$\chi^2_{.9}(2r_0) \leq 37.82$$

which gives $r_0=25$. However, use of such a large r_0 "wastes" information, since a lower value could be used in exchange for either smaller α and β levels, or else a smaller value of θ_0/θ_1 for the same α and β levels, which increases the discriminating ability of the test. However, using this value for r_0 , the test procedure would be as follows:

- (1) Each of the $n=10$ guns is fired one at a time for $R_0=1891$

rounds. When a failure occurs, the round number, failure number, and type of failure is recorded.

(2) Continue testing until 25 immediately clearable failures, 6 non-immediately clearable failures, and 6 parts replacements occur, or else until all weapons have fired $R_0=1891$ rounds.

(3) If all the weapons fired 1891 rounds without the above number of failures occurring, the three hypotheses are accepted at the specified levels of α, β , and θ_0/θ_1 . That is, the weapon meets its specifications. If any of the failure numbers given above is reached before all weapons have fired 1891 rounds, the hypothesis associated with the particular failure mode is rejected, and the weapon does not completely meet the specifications. If all of the above failure numbers are exceeded before all guns have fired 1891 rounds, all three hypotheses are rejected, and the weapon doesn't meet any of the specifications.

This example illustrates the application of this type of test. The testing procedure has both its advantages and disadvantages, as discussed in the following conclusion.

CHAPTER VII

CONCLUSIONS AND RECOMMENDATIONS

This report has attempted to improve on present techniques used by the U.S. Army to test for the reliability of prototype small caliber arms submitted for engineering design tests by several private contractors. The present methods used by the Army are rather arbitrary, varying in method from test to test, and are not built on a sound statistical basis. Present techniques fail to quantify in statistical terms either the specifications nor the testing for the actual values.

Essentially, the random nature of the specified parameters is ignored and no measure is provided as to whether the desired minimum failure rates are actually being achieved. Failure to achieve a desired mean life in weaponry of this type will have a tremendous impact on the overall life cycle costs of the weapon in the form of increased repair and maintenance, availability, downtime, and overall effectiveness. For this reason, it is felt that stringent test procedures in the prototype stage of development are of extreme importance.

The proposed reliability test procedure developed in this report is of much greater efficiency than past methods used in the area of armament testing, and is an attempt to standardize testing techniques in this area as well as provide a measure of how much confidence can be placed in the results of the reliability test. Among the proposals presented in this report are:

(1) More complete specifications of the reliability requirement of a proposed prototype gun in the contract stage of development. Besides stating what failure rate is desired for each failure mode, such quantities as α, β , and θ_0/θ_1 levels should be specified to help insure that what is specified is the same as what is achieved.

(2) Changing the testing procedure to the procedure described in this report. This would determine whether the weapons actually meet the specifications, and help to minimize overall life cycle costs which result from acceptance of substandard items.

The testing procedure described in this report does have one major drawback. The assumption of an underlying exponential distribution of the number of rounds between failures for each failure mode could prove to be a major restriction. Analysis of past firing data is needed to determine the actual underlying distribution of the number of rounds to failure for each failure mode. Also needed is a testing plan which allows for testing with replacement when the assumed underlying probability distribution is Weibull or gamma, which would probably have greater accuracy and flexibility than the exponential distribution, which is a special case of both. Any application of the proposed testing plan presented in this report would yield invalid results if the underlying assumptions are not met.

The effect of using a test which assumes one type of distribution when actually it is of a different type has been researched to a degree. Statistical procedures which are relatively insensitive to departures from the assumed form of the distribution are called robust.

Dannemiller and Zelen (3) have investigated the effect of utilizing testing procedures which were based on the assumption of an exponential distribution when the underlying distribution was actually Weibull with shape parameter 2. They conclude that none of the test procedures are insensitive to departures from the assumed distribution. However, they also conclude that a procedure of the type proposed here is on the conservative side; that is, the β level indicated may be a conservative figure. However, from the view of the producer, this may not be desirable.

It is felt, however, that the proposed testing plan is at the least an effort in the direction of better measurement of the reliability of a prototype weapon, and makes better use of available information than those methods now used.

SELECTED REFERENCES

1. Brown, Tom, Small Caliber Arms Division, Armament Command (ARMCOM), Rock Island, Illinois, Personal Communication With, November, 1973.
2. Buchanan, John D., "Lightweight Individual Weapon Systems (LIWS)," Technical Report No. 72-05, U.S. Army Land Warfare Laboratory, Aberdeen Proving Grounds, Maryland, March, 1972.
3. Dannemiller, M.C., and Zelen, M., "Are Life Testing Procedures Robust?", Proceedings of the Sixth National Symposium on Reliability and Quality Control in Electronics, Washington, D.C., January, 1960, pp. 185-189.
4. Davis, D.J., "An Analysis of Some Failure Data," Journal of the American Statistics Association, Vol. 47, 1952, pp. 113-150.
5. Dellinger, David C., Some Economic Aspects of Reliability and Project Management, Joint Army, Navy, and Air Force Technical Report, No. 67, Applied Mathematics and Statistics Laboratories, Stanford University, California, May, 1963.
6. Duncan, Acheson J., Quality Control and Industrial Statistics, Richard Irwin, Inc., Homewood, Illinois, 1965.
7. Epstein, Benjamin, "Truncated Life Tests in the Exponential Case," Annals of Mathematical Statistics, Vol. 25, pp. 555-564, 1954.
8. Epstein, B., "Life Test Acceptance Sampling Plans When the Underlying Distribution of Life is Exponential," Proceedings of the Sixth National Symposium on Reliability and Quality Control in Electronics, Washington, D.C., January, 1960, pp. 353-360.
9. Epstein, B., and Sobel, M., "Life Testing," Journal of the American Statistics Association, Vol. 48, 1953, pp. 486-502.
10. Epstein, B., and Sobel, M., "Sequential Life Tests in the Exponential Case," Annals of Mathematical Statistics, March, 1955, pp. 82-93.
11. Goode, H.P. and Kao, John H.K., "Sampling Plans Based on the Weibull Distribution," Proceedings of the Seventh National Symposium on Reliability and Quality Control, Washington, D.C., 1961, pp. 24-37.
12. Goode, H.P., and Kao, John H.K., Sampling Procedure and Tables for Life and Reliability Testing Based on the Weibull Distribution, Quality Control and Reliability Technical Report TR3, Office of the Assistant Secretary of Defense (Installations and Logistics),

Washington, D.C., 1961.

13. Goode, H.P., and Kao, John H.K., Sampling Procedures and Tables for Life and Reliability Testing Based on the Weibull Distribution, Quality Control and Reliability Technical Report TR4, Office of the Assistant Secretary of Defense (Logistics and Installation), Washington, D.C., 1962.
14. Gupta, S.S., and Groll, P.A., "Gamma Distribution in Acceptance Sampling Based on Life Tests," Journal of the American Statistics Association, Vol. 56, December, 1961, pp. 942-970.
15. Lloyd, David K., and Myron Lipow, Reliability; Management, Methods and Mathematics, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1962.
16. Molina, E.C., Poisson's Exponential Binomial Limit, D. Van Nostrand, 1949.
17. Reliability of Military Electronic Equipment, Report by Advisory Group on Reliability of Electronic Equipment (AGREE), Office of the Assistant Secretary of Defense (Research and Engineering), Washington, D.C., 1957.
18. Sampling Procedures and Tables for Life and Reliability Testing (Based on the Exponential Distribution), Quality Control and Reliability Handbook (Interim) H-108, Office of the Assistant Secretary of Defense (Supply and Logistics), Washington, D.C., 1960.
19. "Technical Report SWERR-TR72-13", Final Report of Engineering Design Tests Stoner Family of Weapons, Small Arms Weapon Systems Directorate, Weapons Laboratory, Rock Island, Illinois, March, 1972.
20. Wald, A., Sequential Analysis, John Wiley & Sons, New York, 1947.

A P P E N D I X E S

APPENDIX A

Tables for Truncated Life Testing

The following tables give values of n and r for values of $R_0/\theta_0, \theta_1/\theta_0, \alpha$, and β for a truncated testing procedure with replacement. Tables are modified from H-108 (18).

θ/θ_0	r	R_0/θ_0				r	R_0/θ_0			
		1/3	1/6	1/10	1/20		1/3	1/6	1/10	1/20
		α	β	γ	δ		α	β	γ	δ
		$\alpha=0.01$		$\beta=0.01$			$\alpha=0.05$		$\beta=0.01$	
2/3	136	331	551	1103	2207	95	238	397	795	1591
1/2	46	95	158	317	634	33	72	120	241	483
1/3	19	31	51	103	206	13	23	38	76	153
1/5	9	10	17	35	70	7	9	16	32	65
1/10	5	4	6	12	25	4	4	6	13	27
		$\alpha=0.01$		$\beta=0.05$			$\alpha=0.05$		$\beta=0.05$	
2/3	101	237	395	790	1581	67	162	270	541	1082
1/2	35	68	113	227	454	23	47	78	157	314
1/3	15	22	37	74	149	10	16	27	54	108
1/5	8	8	14	29	58	5	6	10	19	39
1/10	4	3	4	8	16	3	3	4	8	16
		$\alpha=0.01$		$\beta=0.10$			$\alpha=0.05$		$\beta=0.10$	
2/3	83	189	316	632	1265	55	130	216	433	867
1/2	30	56	93	187	374	19	37	62	124	248
1/3	13	18	30	60	121	8	11	19	39	79
1/5	7	7	11	23	46	4	4	7	13	27
1/10	4	2	4	8	16	3	3	4	8	16
		$\alpha=0.01$		$\beta=0.25$			$\alpha=0.05$		$\beta=0.25$	
2/3	60	130	217	434	869	35	77	129	258	517
1/2	22	37	62	125	251	13	23	38	76	153
1/3	10	12	20	41	82	6	7	13	26	52
1/5	5	4	7	13	25	3	3	4	8	16
1/10	3	2	2	4	8	2	1	2	3	7

α/β	r	R_0/Θ_0				r	R_0/Θ_0			
		1/3	1/6	1/10	1/20		1/3	1/6	1/10	1/20
		α	β	α	β		α	β	α	β
		$\alpha=0.10$		$\beta=0.01$			$\alpha=0.25$		$\beta=0.01$	
2/3	77	197	329	659	1319	52	140	234	469	939
1/2	26	59	98	197	394	17	42	70	140	281
1/3	11	21	35	70	140	7	15	25	50	101
1/5	5	7	12	24	48	3	5	8	17	34
1/10	3	3	5	11	22	2	2	4	9	19
		$\alpha=0.10$		$\beta=0.05$			$\alpha=0.25$		$\beta=0.05$	
2/3	52	128	214	429	859	32	84	140	280	560
1/2	18	38	64	128	256	11	25	43	86	172
1/3	8	13	23	46	93	5	10	16	33	67
1/5	4	5	8	17	34	2	3	5	10	19
1/10	2	2	3	5	10	2	2	4	9	19
		$\alpha=0.10$		$\beta=0.10$			$\alpha=0.25$		$\beta=0.10$	
2/3	41	99	165	330	660	23	58	98	196	392
1/2	15	30	51	102	205	8	17	29	59	119
1/3	6	9	15	31	63	4	7	12	25	50
1/5	3	4	6	11	22	2	3	4	9	19
1/10	2	2	2	5	10	1	1	2	3	5
		$\alpha=0.10$		$\beta=0.25$			$\alpha=0.25$		$\beta=0.25$	
2/3	25	56	94	188	376	12	28	47	95	190
1/2	9	16	27	54	108	5	10	16	33	67
1/3	4	5	8	17	34	2	2	4	9	19
1/5	3	3	5	11	22	1	1	2	3	6
1/10	2	1	2	5	10	1	1	1	2	5

APPENDIX B

Tables for Truncated Life Testing

The following tables give values of R_0/θ_0 for the various O.C. curves and r_0 given. The code number given in the tables refers to the corresponding O.C. curve having the same code number. Tables and figures are modified from H-108 (18).

Values of R_0/θ_0 for $\alpha = .01$

Code	r	Sample size									
		2r	3r	4r	5r	6r	7r	8r	9r	10r	20r
A-1	1	0.005	0.003	0.003	0.002	0.002	0.001	0.001	0.001	0.001	0.0005
A-2	2	.037	.025	.019	.015	.012	.011	.009	.008	.007	.004
A-3	3	.073	.048	.036	.029	.024	.021	.018	.016	.015	.007
A-4	4	.103	.069	.051	.041	.034	.029	.026	.023	.021	.010
A-5	5	.128	.085	.064	.051	.043	.037	.032	.028	.026	.013
A-6	6	.149	.099	.074	.060	.050	.043	.037	.033	.030	.015
A-7	7	.166	.111	.083	.067	.055	.048	.042	.037	.033	.017
A-8	8	.182	.121	.091	.073	.061	.052	.045	.040	.036	.018
A-9	9	.195	.130	.097	.078	.065	.056	.049	.043	.039	.019
A-10	10	.207	.138	.103	.083	.069	.059	.052	.046	.041	.021
A-11	15	.249	.166	.125	.100	.083	.071	.062	.055	.050	.025
A-12	20	.277	.185	.139	.111	.092	.079	.069	.062	.055	.028
A-13	25	.297	.198	.149	.119	.099	.085	.074	.066	.059	.030
A-14	30	.312	.208	.156	.125	.104	.089	.078	.069	.062	.031
A-15	40	.335	.223	.167	.134	.112	.096	.084	.074	.067	.033
A-16	50	.350	.234	.175	.140	.117	.100	.088	.078	.070	.035
A-17	75	.376	.250	.188	.150	.125	.107	.094	.083	.075	.038
A-18	100	.391	.261	.196	.156	.130	.112	.098	.087	.078	.039

Values of R_0/θ_0 for $\alpha = .05$

Code	r	Sample size									
		2r	3r	4r	5r	6r	7r	8r	9r	10r	20r
B-1	1	0.026	0.017	0.013	0.010	0.009	0.007	0.006	0.006	0.005	0.003
B-2	2	.089	.059	.044	.036	.030	.025	.022	.020	.018	.009
B-3	3	.136	.091	.068	.055	.045	.039	.034	.030	.027	.014
B-4	4	.171	.114	.085	.068	.057	.049	.043	.038	.034	.017
B-5	5	.197	.131	.099	.079	.066	.056	.049	.044	.039	.020
B-6	6	.218	.145	.109	.087	.073	.062	.054	.048	.044	.022
B-7	7	.235	.158	.117	.094	.078	.067	.059	.052	.047	.023
B-8	8	.249	.166	.124	.100	.083	.071	.062	.055	.050	.025
B-9	9	.261	.174	.130	.104	.087	.075	.065	.058	.052	.026
B-10	10	.271	.181	.136	.109	.090	.078	.068	.060	.054	.027
B-11	15	.308	.205	.154	.123	.103	.088	.077	.068	.062	.031
B-12	20	.331	.221	.166	.133	.110	.095	.083	.074	.066	.033
B-13	25	.348	.232	.174	.139	.116	.099	.087	.077	.070	.035
B-14	30	.360	.240	.180	.144	.120	.103	.090	.080	.072	.036
B-15	40	.377	.252	.189	.151	.126	.108	.094	.084	.075	.038
B-16	50	.390	.260	.195	.156	.130	.111	.097	.087	.078	.039
B-17	75	.409	.273	.204	.164	.136	.117	.102	.091	.082	.041
B-18	100	.421	.280	.210	.168	.140	.120	.105	.093	.084	.042

Values of R_0/θ_0 for $\alpha = .10$

Code	r	Sample size									
		2r	3r	4r	5r	6r	7r	8r	9r	10r	20r
C-1	1	0.053	0.035	0.026	0.021	0.018	0.015	0.013	0.012	0.011	0.005
C-2	2	.133	.089	.066	.053	.044	.038	.033	.030	.027	.013
C-3	3	.184	.122	.092	.073	.061	.052	.046	.041	.037	.018
C-4	4	.218	.145	.109	.087	.073	.062	.055	.048	.044	.022
C-5	5	.243	.162	.122	.097	.081	.070	.061	.054	.049	.024
C-6	6	.263	.175	.131	.105	.088	.075	.066	.058	.053	.026
C-7	7	.278	.185	.139	.111	.093	.079	.070	.062	.056	.028
C-8	8	.291	.194	.146	.116	.097	.083	.073	.065	.058	.029
C-9	9	.302	.201	.151	.121	.101	.086	.075	.067	.060	.030
C-10	10	.311	.207	.156	.124	.104	.089	.078	.069	.062	.031
C-11	15	.343	.229	.172	.137	.114	.098	.086	.076	.069	.034
C-12	20	.363	.242	.182	.145	.121	.104	.091	.081	.073	.036
C-13	25	.377	.251	.188	.151	.126	.108	.094	.084	.075	.038
C-14	30	.387	.258	.194	.155	.129	.111	.097	.086	.077	.039
C-15	40	.402	.268	.201	.161	.134	.115	.100	.089	.080	.040
C-16	50	.412	.275	.206	.165	.137	.118	.103	.092	.082	.041
C-17	75	.428	.285	.214	.171	.143	.122	.107	.095	.086	.043
C-1	100	.437	.291	.219	.175	.146	.125	.109	.097	.087	.044

Values of R_0/θ_0 for $\alpha = .25$

Code	r	Sample size									
		2r	3r	4r	5r	6r	7r	8r	9r	10r	20r
D-1	1	0.144	0.096	0.072	0.058	0.048	0.041	0.036	0.032	0.029	0.014
D-2	2	.240	.160	.120	.096	.080	.069	.060	.053	.048	.024
D-3	3	.288	.192	.144	.115	.096	.082	.072	.064	.058	.029
D-4	4	.317	.211	.158	.127	.106	.091	.079	.070	.063	.032
D-5	5	.337	.225	.168	.135	.112	.096	.084	.075	.067	.034
D-6	6	.352	.234	.176	.141	.117	.100	.088	.078	.070	.035
D-7	7	.363	.242	.182	.145	.121	.104	.091	.081	.073	.036
D-8	8	.372	.248	.186	.149	.124	.106	.093	.083	.074	.037
D-9	9	.380	.253	.190	.152	.127	.109	.095	.084	.076	.038
D-10	10	.386	.258	.193	.155	.129	.110	.097	.086	.077	.039
D-11	15	.408	.272	.204	.163	.136	.117	.102	.091	.082	.041
D-12	20	.421	.281	.210	.168	.140	.120	.105	.094	.084	.042
D-13	25	.429	.286	.215	.172	.143	.123	.107	.095	.086	.043
D-14	30	.436	.291	.218	.174	.145	.125	.109	.097	.087	.044
D-15	40	.445	.296	.222	.178	.148	.127	.111	.099	.089	.044
D-16	50	.451	.300	.225	.180	.150	.129	.113	.100	.090	.045
D-17	75	.460	.307	.230	.184	.153	.131	.115	.102	.092	.046
D-18	100	.465	.310	.233	.186	.155	.133	.116	.103	.093	.047

Values of R_0/θ_0 for $\alpha = .50$

Code	r	Sample size									
		2r	3r	4r	5r	6r	7r	8r	9r	10r	20r
E-1	1	0.347	0.231	0.173	0.139	0.116	0.099	0.087	0.077	0.069	0.035
E-2	2	.420	.280	.210	.168	.140	.120	.105	.093	.084	.042
E-3	3	.446	.297	.223	.178	.149	.127	.111	.099	.089	.045
E-4	4	.459	.306	.230	.184	.153	.131	.115	.102	.092	.046
E-5	5	.467	.311	.234	.187	.156	.133	.117	.104	.093	.047
E-6	6	.473	.315	.236	.189	.158	.135	.118	.105	.095	.047
E-7	7	.476	.318	.238	.191	.159	.136	.119	.106	.095	.048
E-8	8	.479	.320	.240	.192	.160	.137	.120	.107	.096	.048
E-9	9	.482	.321	.241	.193	.161	.138	.120	.107	.096	.048
E-10	10	.483	.322	.242	.193	.161	.138	.121	.107	.097	.048
E-11	15	.489	.326	.244	.196	.163	.140	.122	.109	.098	.049
E-12	20	.492	.328	.246	.197	.164	.140	.123	.109	.098	.049
E-13	25	.493	.329	.247	.197	.164	.141	.123	.110	.099	.049
E-14	30	.494	.330	.247	.198	.165	.141	.124	.110	.099	.049
E-15	40	.496	.331	.248	.198	.165	.142	.124	.110	.099	.050
E-16	50	.497	.331	.248	.199	.166	.142	.124	.110	.099	.050
E-17	75	.498	.332	.249	.199	.166	.142	.124	.111	.100	.050
E-18	100	.498	.332	.249	.199	.166	.142	.125	.111	.100	.050

