AFFDL-TR-65-95

RANDOM FATIGUE TEST SAMPLING REQUIREMENTS

S. C. CHOI
L. D. ENOCHSON

TECHNICAL REPORT AFFDL-TR-65-95

AUGUST 1965

AIR FORCE FLIGHT DYNAMICS LABORATORY
RESEARCH AND TECHNOLOGY DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Copies of this report should not be returned to the Research and Technology Division unless return is required by security considerations, contractual obligations, or notice on a specific document.

500 - September 1965 - AF(615)773-6-138

RANDOM FATIGUE TEST SAMPLING REQUIREMENTS

S. C. CHOI
L. D. ENOCHSON

FOREWORD

This report was prepared by the Measurement Analysis Corporation Los Angeles, California, for the Aero-Acoustics Branch, Vehicle Dynamics Division, AF Flight Dynamics Laboratory, Wright-Patterson Air Force Base, Ohio, under Contract AF 33(615)-1314. This study covers random fatigue test sampling requirements in conjunction with optimum utilization of the Sonic Fatigue Facility Data Analysis System. The Project No. is 4437, "High Intensity Sound Environment Simulation" and Task No. 443706, "Advanced Instrumentation Study for Sonic Fatigue Experimental Work." Mr. W. K. Shilling, III was the Project Engineer. Measurement Analysis Corporation report number is MAC 402-02A. This report covers work conducted from February 1964 to April 1965.

The manuscript was released by the authors in May 1965 for publication as an RTD technical report.

This technical report has been reviewed and is approved.

Walter J. Mykytow
Asst. for Research and Technology
Vehicle Dynamics Division
AF Flight Dynamics Laboratory

ABSTRACT

Basic considerations are discussed for determining sample sizes and record lengths for various statistical tests and estimates which are important to random fatigue testing. Methods for determining minimum sample sizes when comparing means and variances of normally (Gaussian) distributed random variables are described. Procedures for reducing a relatively large sample to a smaller sample are presented. Elimination of outliers and systematic resampling are two methods given.

An explanation is presented of the requirements and problems involved in the determination of record lengths necessary for an estimate of a given accuracy for autocorrelation functions, ordinary power spectral density functions, cross-correlation functions, cross-spectral density functions, frequency response functions, and probability density functions.

Due to its importance in random fatigue testing applications, the basic properties of the Weibull distribution in terms of its parameters and the failure rate are summarized. A presentation is given of estimation and statistical testing problems related to the Weibull distribution. The best available methods of estimating the parameters are described. Methods of determining sample sizes needed for various analyses are developed. Some problems of reliability analysis applicable in fatigue testing are discussed. New methods of decision techniques for comparing two or more systems are proposed in terms of reliability. The report concludes with an example of the application of the Weibull distribution to actual fatigue test data.

RANDOM FATIGUE TEST SAMPLING REQUIREMENTS

CONTENTS

ILLUSTRATIONS

RANDOM FATIGUE TEST SAMPLING REQUIREMENTS

1. INTRODUCTION

In random fatigue testing, many different statistical parameters are of interest in the various facets of a given test. This report discusses some of the sample size requirements which are necessary to estimate these parameters. Of central importance to fatigue life testing is the Weibull distribution. This is so because the probability distribution of parameters such as time to failure and cycles to failure are usually reasonably modeled by the Weibull distribution. Various statistical aspects of the Weibull distribution are discussed herein.

In many experimental design situations, an assumption of a Gaussian distribution of various sample statistics is invoked in order to allow the estimation of required sample sizes or record lengths. This assumption is almost always justified if the sample sizes one is concerned with are large, say greater than thirty. The central limit theorem guarantees Gaussian distributions for large N. In other situations, one is forced into Gaussian assumptions for small sample sizes due to the lack of an available exact theory. Any requirements based on Gaussian assumptions in this case become questionable but are at least reasonably proper guidelines for experiment planning purposes. Hence, although most of the results in this report are based on Gaussian assumptions, they are in practice usefully applied to most practical problems.

The two most fundamental quantities in statistics are mean values and variances. The first section following, therefore, discusses sample size requirements for means and variances assuming a Gaussian distribution. Often, a large sample of data will be collected which must be reduced to a smaller more tractable size. Some of the ideas involved in eliminating unwanted data points (outliers) and methods for over-all reduction of sample size are presented. Later sections discuss correlation function, power

spectral density function, and frequency response function estimates, and are presented in terms of an allowable percentage normalized standard error. The final section describes fatigue life testing applications of the Weibull distribution.

Two different approaches are used for determining sample size requirements. In Section 2, it is assumed that some reason exists for hypothesizing a specific value for a population parameter. One may then calculate the sample size necessary to detect a specified deviation from this hypothesized value with a given probability. This is the method to use when one has:

i) a specific value predicted by theory against which to test (for example, the theoretical expected number of runs in a sample of N independent observations is $\left[N/2\right] + 1$).

ii) a measured known value, possibly from previous experiments, and one is hypothesizing the new data to be significantly different (i.e., testing a supposedly improved product [new structural panel] against a former product).

The other approach is that of computing the sample size necessary to estimate a parameter with a given percentage error as opposed to specifying a specific value. The normalized standard error is employed. This is the square root of variance (the standard error) of the estimate divided by its expected value (normalized) to give variability in a percentage form. The pitfall in this concept lies in attaching undue importance to a deviation of one standard deviation (rms value). Deviations of plus and minus one standard deviation occur with a given probability and are of no more importance than deviations of, say, plus and minus two or plus and minus one-half standard deviations. Therefore, in quoting results or performing calculations one must be careful to note that one allows a deviation of a given percentage with a specific probability, and that rms values are not maximum errors which occur.

2. SAMPLE SIZE CALCULATIONS FOR EQUIVALENCE OF MEANS AND VARIANCES

In any situation where one knows or hypothesizes a mean and variance of a Gaussian distribution, one may compute sample sizes necessary to properly test sample values against these theoretical values. Certain constraints must be imposed on the problem, such as specifying the level of significance and probability of Type II Error, which are explained below. The sample size may then be calculated which maintains these probabilities. In other cases one may be able to calculate required sample sizes based on a requirement to estimate a parameter with a specified percentage (rms) error.

A theoretical mean μ and variance σ^2 can sometimes be computed for a given distribution (or one can assume values). Using these theoretical values, one can then test obtained sample values to determine if the observed distribution can be considered to be the same as the theoretical distribution. In this case the statistical hypothesis is: "There is no evidence to conclude that the sample values are not the same as the theoretical values." These will be two-tailed tests, since deviations from the hypothesized values may occur in either direction.

Two types of errors can be made:

Type I Error - Rejecting the hypothesis when it is really true with probability α

Type II Error - Accepting the hypothesis when it is really false with probability β

To illustrate these two errors, one only needs to consider the sample mean values computed from two different random samples of observations drawn from the same underlying population. Clearly, with a certain small probability, say $\alpha = 10\%$, these sample mean values might differ enough to appear truly different. This is the Type I Error. On the other hand, if random samples are collected from two slightly different populations, clearly

by chance (say with probability $\beta = 10\%$), the mean values computed from these two samples might be so close together that they appear equivalent. This is the Type II Error.

As can be seen from this example, the farther apart the populations truly are, the smaller is the chance of the sample mean values appearing equivalent. Hence, in addition to specifying α and β, one must impose an additional restraint on the problem to allow the sample size to be calculated. That is, one must specify what particular deviation from the hypothesized parameter will allow the hypothesis to be accepted with probability β. In some specific situations one might have suspicions about the theory involved and anticipate some particular value other than the hypothesized value. In other cases one must use judgment in selecting values somewhat arbitrarily.

For the illustrative examples in this section, a ten percent difference in means and a fifty percent difference in standard deviations are selected as the values at which the probability of Type II Error will be held. Of course, other deviations from the theoretical values may be chosen, and have a specific associated probability of the hypothesis being accepted. Also, for simplicity, α and β will be chosen each equal to 10%.

2.1 SAMPLE SIZE FOR EQUIVALENCE OF MEANS

The calculation of the required sample size for the test of equivalent mean values is as follows: Let μ' be the mean value of the distribution which is to be detected with a probability $\beta = 10\%$, and $z_{1-\alpha/2}$ a normal (Gaussian) deviate such that

$$\text{Prob}\,(z \leq z_{1-\alpha/2}) = 1-\alpha/2$$

That is,

$$\text{Prob}\left[z \leq z_p\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_p} e^{-z^2/2}\, dz = p \qquad (1)$$

The reason for using $\alpha/2$ instead of α is to allow for two-sided deviations so that the Type I Error is α.

The following relations now hold where x_c is the critical point (see Figure 1). That is, x_c is a value such that if $\bar{x} > x_c$, the hypothesis is rejected, and if $\bar{x} < x_c$, the hypothesis is accepted where $\bar{x}$ is the calculated sample mean. The sample mean $\bar{x}$ is defined by the equation

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \qquad (2)$$

where x_i are the observations which make up the sample.

Figure 1 is not quite complete in the sense that deviations in either direction are considered. However, the symmetry is implied by using $\alpha/2$ and $\beta/2$ rather than α and β. In terms of $\alpha/2$,

$$z_{1-\alpha/2} = \frac{x_c - \mu}{\sigma/\sqrt{N}} \qquad (3)$$

while, in terms of $\beta/2$,

$$z_{\beta/2} = \frac{x_c - \mu'}{\sigma/\sqrt{N}} \tag{4}$$

where σ^2 is the theoretical variance of the distribution being sampled.

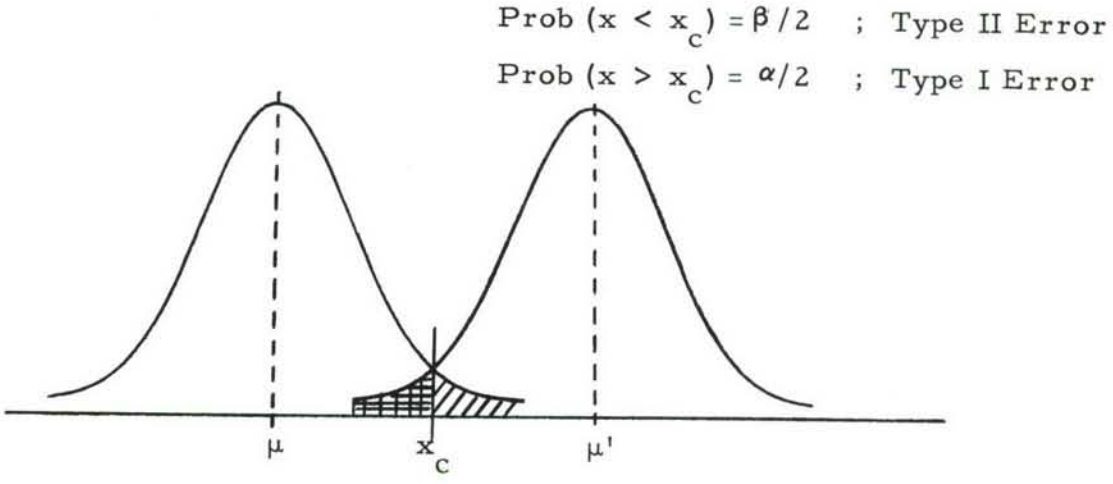


Figure 1. Illustration of Type I, Type II Errors

In the special situation where one sets $\alpha = \beta$, it follows that $z_{\alpha/2} = z_{\beta/2}$. Also, due to the symmetry of the normal distribution, $z_{1-\alpha/2} = -z_{\alpha/2}$. Hence, from Eqs. (3) and (4),

$$\frac{x_c - \mu}{\sigma/\sqrt{N}} = -\frac{x_c - \mu'}{\sigma/\sqrt{N}}$$

Solving for x_c one obtains

$$x_c = \frac{\mu + \mu'}{2} \tag{5}$$

Then substituting back in Eq. (4),

$$z_{\alpha/2} = \frac{\frac{\mu + \mu'}{2} - \mu'}{\sigma/\sqrt{N}} = \frac{\frac{\mu - \mu'}{2}}{\sigma/\sqrt{N}}$$

Letting $\Delta\mu = \mu - \mu'$, and solving for N:

$$N = \frac{\sigma^2 z^2_{\alpha/2}}{(\Delta\mu/2)^2} = 4\left(\frac{\sigma}{\Delta\mu} z_{\alpha/2}\right)^2 \qquad (6)$$

Note that the implicit assumption has been made that $\sigma^2 = (\sigma')^2$ where $(\sigma')^2$ is the alternative variance. Therefore, this test would be properly performed after the alternative variance $(\sigma')^2$ had been determined to be statistically equivalent to the theoretical value σ^2.

Computational Example

The calculation of N based on the test for equivalent means is illustrated as follows. Assume from independent considerations one obtains the theoretical values

$$\mu = 50$$

$$\sigma^2 = 25$$

The required sample size to detect a 10% difference in means (namely $\Delta\mu = 5$ here) with a Type II Error of $\beta = 10\%$ then is calculated by applying Eq. (6). For $\alpha = 10\%$, the term $z_{.95} = 1.645$. Thus

$$N = 4\left(\frac{25}{25}\right)(1.645)^2 \approx 11$$

This sample size of N = 11 will be compared later to the sample size required for equivalent variances.

2.2 SAMPLE SIZE FOR EQUIVALENCE OF VARIANCES

The reasoning for obtaining a formula to compute the sample size for the variance equality test proceeds in a similar manner. Let s_c^2 represent the critical point. Then, since $(N-1)\, s^2/\sigma^2$ has a χ^2 distribution with $(N-1)$ d.f., one has

$$s_c^2 = \frac{(\sigma')^2}{(N-1)} \chi^2_{\beta/2} \tag{7}$$

and

$$s_c^2 = \frac{\sigma^2}{(N-1)} \chi^2_{1-\alpha/2} \tag{8}$$

where $\chi^2_{\beta/2}$ and $\chi^2_{1-\alpha/2}$ are points of the χ^2 distribution with $(N-1)$ d.f. The sample (unbiased) variance s^2 is defined by the formula

$$s^2 = \frac{1}{(N-1)} \sum_{i=1}^{N} (x_i - \bar{x})^2 \tag{9}$$

Equating (7) and (8) and rearranging terms gives for the case $\alpha = \beta$.

$$\frac{(\sigma')^2}{\sigma^2} = \frac{\chi^2_{1-\alpha/2}}{\chi^2_{\alpha/2}} \tag{10}$$

Although $(N-1)$ cancels out, χ^2 is a function of $(N-1)$. Therefore, when σ^2 and $(\sigma')^2$ are specified, a trial and error inspection of a χ^2 table will give values of χ^2 for some number of d.f. such that Eq. (10) holds true.

Computational Example

For example, for $(N-1) = 29$ d.f., one finds in the χ^2 table

$$\frac{\chi^2_{.05}}{\chi^2_{.95}} = \frac{42.6}{17.7} = 2.41$$

For all practical purposes this corresponds to the desired ratio of standard deviations of 1.55 and 1.0. Therefore, a convenient sample size to test for variance equivalence is 30.

Note that the variance equivalence test has a larger required sample size than the mean equivalence test, namely $N = 30$ as compared to $N = 11$. Therefore, this would determine the over-all sample size for the experiment.

3. REDUCTION OF SAMPLE SIZE

Suppose that it is desired to reduce a large sample of size M to a smaller size n ($n < M$). The purpose of this section is to describe the method of reduction. It is not intended here to discuss how to obtain the original sample of M data points but the method also applies to the original sampling since one can consider that an infinite amount of data is reduced to M data points. Assume that all data are from the same population. If one suspects that they come from two different populations, one has to partition them into two disjoint groups before the analysis is performed. That problem belongs to the topic of classification analysis. It is not discussed here.

The reduction is conceived in two steps:

1. eliminate all bad observations (outliers)
2. reduce a sample consisting of a large number of data points to a smaller representative sample for detailed analysis

3.1 METHODS OF REDUCTION

Step 1. Elimination of Outliers (Bad Observations)

Often a sample data of size M contains some erroneous data which are called outliers. These errors result from such factors as instrumentation or human errors.

A statistic which is used to detect outliers is R/s, the range divided by the sample standard deviation. The sample standard deviation s is the independent external estimate of the standard deviation obtained from concurrent or past data, not from the sample on hand. A test of outliers can be performed if the percentile points of the R/s are available. These percentile points when the underlying data are from a normal distribution are shown in Table 1 (Reference 1).

d.f \ n	5	10	15	20
20	4.23	5.01	5.43	5.71
30	4.10	4.83	5.21	5.48
40	4.04	4.74	5.11	5.36
60	3.98	4.65	5.00	5.24
120	3.92	4.56	4.90	5.13
∞	3.86	4.47	4.80	5.01

Table 1. Table of 95 Percentiles, C(.95), of the Distribution of R/s

[The parameter n is the sample size and d.f. is the number of degrees-of-freedom in the independent standard deviation s.]

Let $x_1, x_2, \ldots, x_M$ be the sample of size M. Denote $x_1 = \text{Min}\{x_i\}$ and $x_M = \text{Max}\{x_i\}$. Then $R = x_M - x_1$ and the critical region for rejection is $R/s > c(\alpha)$, where $c(\alpha)$ is 100α percentile point from the available table. If $R/s > c(\alpha)$, the rejection rule is:

reject x_1 if $(\bar{x} - x_1) > (x_M - \bar{x})$

reject x_M if $(\bar{x} - x_1) < (x_M - \bar{x})$

reject both x_1 and x_M if $(\bar{x} - x_1) = (x_M - \bar{x})$

where $\bar{x}$ is the mean.

One application of the above technique in fatigue testing is as follows. Suppose that one has a sample of large size on hand. Now assume that a smaller second sample is obtained from the same system. It is suspected that a few data points of the second group are set apart from the others. One wonders whether or not they are far enough from the others so that one can reject them as being caused by some assignable but thus far unascertained cause. Now, one applies the above technique to decide whether the data should be kept or not. In this case the range is computed from the second sample and the standard deviation is computed from the first group to apply the rejection rule described in Step 1.

If rejection occurs, then the sample size is reduced to (M - 1) or (M - 2) from M. Now a new range R_2 and a new mean $\bar{x}_2$ are computed from the remaining data. Then the same procedure as described above is applied to detect the next possible outlier(s). The procedure is continued recursively until $R/s \leq c(\alpha)$. Let N denote the reduced sample size from which all outliers have been removed.

Example 1: Assume that the following 21 measurements are made from a normally distributed record. Further assume that the measurements are made sequentially at fixed intervals of time.

52	55	56	49	33	56	44
55	43	40	24	44	41	39
45	59	36	51	44	45	45

Suppose that it is desired to check for possible outliers. Assume that the above data are obtained from the same source as 1000 previous data in which sample standard deviation was found to be $s = 6.5$. Then proceed as follows.

$\bar{x} = 45.52$ $\qquad$ $c(.95) = 5.01$

$R = 59 - 24 = 35$ $\qquad$ $5.38 > 5.01$

$R/s = 35/6.5 = 5.38$

Therefore, outlier(s) are indicated. Since $(\bar{x} - 24) > (59 - \bar{x})$, one rejects 24 as being an outlier from the above data at 95% confidence level. Next, one computes that

$$\bar{x} = 46.6 \quad \text{(new mean)}$$

$$R_2 = 59 - 33 = 26$$

$$R_2/s = 26/6.5 = 4.00 < 5.01 = c(.95)$$

Thus, all the rest of the data are kept as good observations. The variance s^2 of 20 remaining data points is 52.46.

Step 2. Resampling

Now assume that it is desired to reduce the sample size from N to n, $(n < N)$. Suppose that the N data points are numbered 1 to N in an arbitrary manner. Three methods of reduction are discussed below.

a) Simple Random Reduction

This method is the simplest one. One simply randomly selects n points out of the sample of size N. A convenient method of selecting random samples is to apply commonly available "random number" tables. If one reads 23, 4, 13, ... from a table, then one selects the 23rd, 4th, 13th, ... data points from the N data until a total of n data points is obtained. The variance of the mean, $\mathrm{Var}(\bar{x})$, of the selected data in terms of original N data is

$$\mathrm{Var}(\bar{x}) = \frac{(N - n)}{Nn} s^2 \tag{11}$$

where s^2 is the sample variance of the N data points. Sometimes it is descriptive to talk about the precision of the estimate. Precision is defined as the reciprocal of the variance. Thus, in the case of Eq. (11), the precision of $\bar{x}$ obtained by a simple random reduction is

$$P(\bar{x}) = \frac{Nn}{(N-n)\,s^2}$$

It refers to the measure of precision of $\bar{x}$ obtained by repeated application of the same reduction procedure. It is obvious that the less the variance, the higher the precision of any estimate.

b) Systematic Reduction

Let $k = [N/n]$ be the greatest integer not larger than the quantity N/n. Partition the N data points into n disjoint subgroups of size k. In sampling theory these subgroups are called strata. Since N is not, in general, an integral multiple of n, different strata may vary by one data point in size. Select one sample data point from the first stratum at random and every kth data point thereafter. The variance of systematically reduced data is

$$\mathrm{Var}(\bar{x}) = \left(\frac{N-1}{N}\right) s^2 - \frac{1}{N} \sum_{i=1}^{n} \sum_{j=1}^{k} (x_{ij} - \bar{x}_i)^2 \qquad (12)$$

where s^2 is the variance of the N data points. x_{ij} denotes the jth sample point of ith stratum and $\bar{x}_i$ denotes the mean of ith sample. Equation (12) is proved in Reference 2. When Eq.(11) is compared with Eq. (12), one can state that the mean of a systematically reduced data is more precise than the mean of a simple random sample if and only if

$$\left(\frac{N-1}{N}\right) s^2 - \frac{1}{N} \sum_{i=1}^{n} \sum_{j=1}^{k} (x_{ij} - \bar{x}_i)^2 < \left(\frac{N-n}{Nn}\right) s^2$$

or

$$\left(\frac{n-1}{n}\right) s^2 < \frac{1}{N} \sum_{i=1}^{n} \sum_{j=1}^{k} (x_{ij} - \bar{x}_i)^2$$

Since $N/n \approx k$ one obtains the following condition.

$$s^2 < \frac{1}{k(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{k} (x_{ij} - \bar{x}_i)^2 \qquad (13)$$

This result implies that systematic reduction has more precision than simple random reduction if the variance within the systematic sample is larger than the original variance of N data. That is, systematic sampling is favorable when the reduced data are heterogeneous and unfavorable when they are homogeneous. If the population has a periodic trend, effectiveness of the method depends on the value of k. The least favorable case occurs if k is an integral multiple of the period. A favorable case occurs when k is an odd multiple of a half period. See Figure 2.

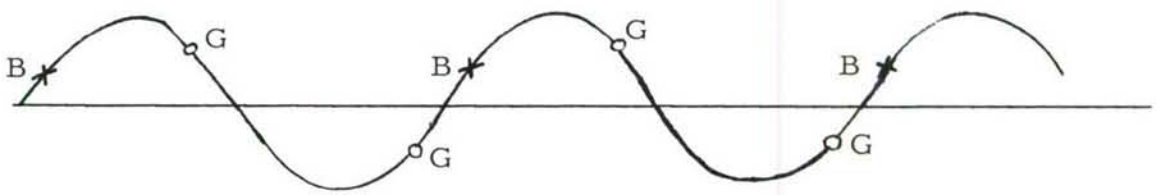


Figure 2. Periodic Variation. B denotes unfavorable case and G denotes favorable case.

In the case where the population values occur as a linear trend, the systematic method is the most efficient technique available.

Example 2: Suppose that it is desired to reduce the 20 data points of Example 1 to size 5 by the systematic reduction. Arrange the data in 5 strata as follows.

stratum No. i	data				$\sum_{j=1}^{4}(x_{ij}-\bar{x}_i)^2$	s_i^2
1	52	55	56	49	30	10.00
2	33	56	44	55	350	116.67
3	43	40	44	41	10	3.33
4	39	45	59	36	313	104.33
5	51	44	45	45	31	10.33

Now choose a data point at random from the first four measurements, say 55 (second data point). Then select every fourth data thereafter. Thus, the reduced sample data are

55 56 40 45 44

The mean and variance of the mean of a reduced data in this case is obtained by Eq. (12)

$$\bar{x} = 48.0$$

$$\text{Var}(\bar{x}) = (19/20)\,52.46 - (1/20)(30+350+10+313+31)$$

$$= 49.84 - 36.70 = 13.14$$

If a random reduction is used for the above data, one obtains by Eq. (11)

$$\text{Var}(\bar{x}) = \frac{(20-5)}{(20)(5)}\ 52.46 = 7.87$$

Thus, in the above case, the random reduction gives a more efficient result.

c) Stratified Reduction

In this method the N data are partitioned into n disjoint strata of approximately equal size according to some characteristic such as time or magnitude. Then one data point is selected from each stratum at random. (See Figure 2.) If one suspects any periodic trend and the period is unknown, then the stratified reduction is recommended over the method (b). The estimate of the mean by the stratified reduction is given by

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i}{n} \tag{14}$$

and its variance is

$$\text{Var}(\bar{x}) = \left(\frac{N - n}{N}\right) \frac{1}{n^2} \sum_{i=1}^{n} s_i^2 \tag{15}$$

where k is the stratum size, and s_i^2 denotes the variance of the ith stratum. Note that $k \approx \left[N/n\right]$.

The set of elements upon which the sample size reduction operation is performed is called a frame and in many practical situations a given population conceivably contains a number of different frames. In Figure 3, a system is stratified into three zones in two ways.

Frame I Frame II

Figure 3. Zone Stratifications of a System

Consider two frames of stratification as shown in Figure 3. Suppose it is known that the between-strata variance of three subregions A, B, and C of Frame I is greater than corresponding variance of Frame II. This implies that the sum of within-strata variance of the Frame I is less than that of the Frame II. Consequently, by Eq. (15), Frame I is preferred over Frame II. Thus, the choice of a frame is often an important aspect of sample design. In general, a good frame is one which is heterogeneous between subregions and homogeneous within each subzone. Total sample size is then allocated into three zones according to the importance of each zone, such as size and sensitivity.

Example 3: Consider again the data from Example 2. By stratified random reduction, one selects one data point from every stratum. That is, first point from (52, 55, 56, 49) and second point from (33, 56, 44, 59), etc. Thus, the reduced sample data might be

55 44 40 45 51

The mean and its variance of a reduced data in this case are obtained by Eqs. (14) and (15).

$$\bar{x} = 47.0$$

$$\text{Var}(\bar{x}) = \left[(20-5)/20\right] \frac{1}{5^2} (10.0+116.67+3.33+104.33+10.33)$$

$$= 7.34$$

Thus, stratified reduction method yields the most efficient estimate in the above example.

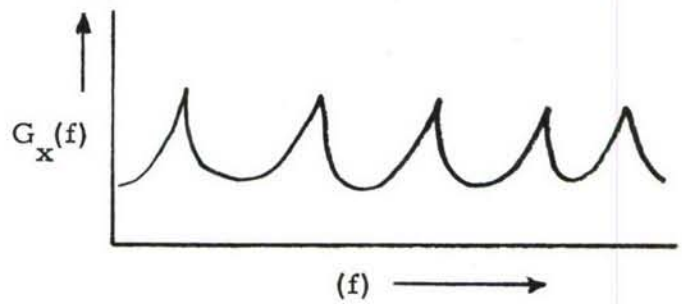
4. AUTOCORRELATION FUNCTION ESTIMATES

Certain arbitrary quantities must be decided upon for autocorrelation function estimates. First, an acceptable percentage normalized standard error ϵ referred to the value of $R(\tau)$ at $\tau = 0$ (the mean square value) must be established. Second, the bandwidth of the signal being analyzed must be known or estimated.

It can be shown, that under the assumption of a Gaussian process, the normalized standard error $\epsilon = \epsilon(0)$ is

$$\epsilon = \frac{1}{\sqrt{BT}} \tag{16}$$

where T is the record length used in the analysis, and B is the signal bandwidth appropriately defined. For Eq. (16) to theoretically hold true, the process x(t) should have a flat spectrum B cps wide with a perfectly sharp cutoff. In practice, B is much more difficult to define. For experiment planning purposes, one can only be careful to estimate the bandwidth B conservatively too small. When an experiment is completed or one has other reasons to know the shape of the spectrum, other problems arise. For example, suppose the spectrum of x(t) has the shape indicated in the sketch below.



In such a situation, probably one should choose as B the sum of the half-power point bandwidths of each peak. Similar judgments must be made in other complicated situations.

From Eq. (16), it is straightforward to compute a required record length T. For example, suppose it is desired to maintain $\epsilon = 10\%$ and B is known to be 2000 cps, then

$$T = \frac{1}{B\epsilon^2} = \frac{1}{2 \cdot 10^3 \times 10^{-2}} = .05 \text{ sec.}$$

If one collects N independent discrete observations, then the variance of the autocorrelation estimate is (see Reference 3, p. 358),

$$\text{Var}\left[\hat{R}_x(\tau)\right] = \frac{R_x^2(0) + R_x^2(\tau)}{N} \tag{17}$$

Note that this expression depends on the true (unknown in general) autocorrelation function $R_x(\tau)$ of the process being analyzed. The normalized standard error is

$$\epsilon(\tau) = \left\{ \frac{\left[R_x(0)/R_x(\tau)\right]^2 + 1}{N} \right\}^{\frac{1}{2}} \tag{18}$$

From this equation, if $R_x(\tau)$ is known, one can obtain the necessary sample size for any point on the correlation function.

Certain problems arise in deciding upon the necessary accuracy for a correlation estimate. For example, a typical correlation function has the form illustrated in Figure 4 below.

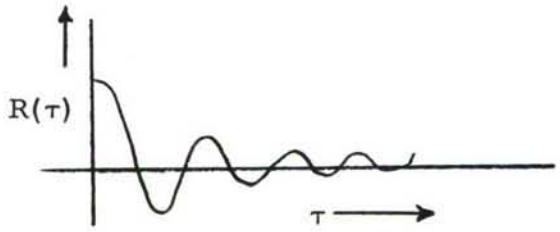


Figure 4. Typical Autocorrelation Function

A percentage-wise accurate estimate of $R(\tau)$ at one of the points where $R(\tau)$ is near zero would require an inordinately large sample size N. For this reason, it is not feasible to select sample sizes which will maintain small percentage of reading errors for all values of $R(\tau)$. A "percent of full scale" type error is a more reasonable quantity for this application. That is, the value of $R(\tau)$ at $\tau = 0$ (the maximum value of $R(\tau)$) should dictate sample size requirements for estimating the entire correlation function. Note that in Eq.(17) $R_x^2(\tau)$ is bounded above by $R_x^2(0)$ and below by zero so that the maximum variability takes place at $R_x(0)$. In this sense, basing sample size requirements for $R_x(\tau)$ entirely on the point $\tau = 0$ is conservative and proper. If one employs the often used relation $N = 2BT$ for relating continuous and discrete samples, then Eq. (18) will reduce to Eq. (17) at $\tau = 0$, namely,

$$\epsilon = \sqrt{\frac{2}{N}} = \frac{1}{\sqrt{BT}} \tag{19}$$

The relation $N = 2BT$ gives the degrees-of-freedom in a signal $x(t)$ with a flat spectrum of width B with a perfectly sharp cutoff. Degrees-of-freedom in this case means the number of independent points which uniquely determine $x(t)$. For this special situation, degrees-of-freedom is equivalent to the sample size N (i.e., N independent observations). Additional discussion of this point is given in Section 4 concerning power spectrum estimates.

When one is considering discrete observations which are statistically correlated, modifications have to be made to Eq. (19). In terms of the true autocorrelation function, the expression for ϵ becomes

$$\epsilon = \sqrt{\frac{2}{N} + \frac{4}{N^2} \sum_{r=1}^{N-1} (N - r) \frac{R_x^2(rh)}{R_x^2(0)}} \tag{20}$$

If one assumes the easily analyzed case of an exponential correlation function which occurs physically in the case of lowpass R-C filtered noise, then Eq. (20) becomes approximately

$$\epsilon = \sqrt{\frac{2}{N}} \sqrt{\frac{e^{2bh} + 1}{e^{2bh} - 1}} \tag{21}$$

where

$$R_x(rh) = R_x(0)\, e^{-b|rh|} \tag{22}$$

For Eq. (22) to apply, the requirements $N > 100$, $bh > 0.01$ should be met. This is only an approximation for other physically occurring situations, but should usually be conservative and quite useful. In Figure 5, several curves are drawn for various values of bh from which one can obtain N as a function of ϵ or vice versa. The parameter b in Eq. (21) and Eq. (22) is the noise bandwidth of the process $x(t)$. For example, assume $x(t)$ is sampled at an interval $h = 0.01$ sec. apart and that the noise bandwidth is $b = 60$ cps so that $bh = 0.60$. Further assume one wants to maintain $\epsilon = 5.0\%$. Then, by inspecting Figure 5, one notes that a sample size $N = 1490$ is necessary. Note that this implies a record length of $T = Nh = 14.9$ sec. This compares with $N = 800$, $T = 8.0$ sec. required for the case of independent samples which is given by the bottom curve for $bh = \infty$.

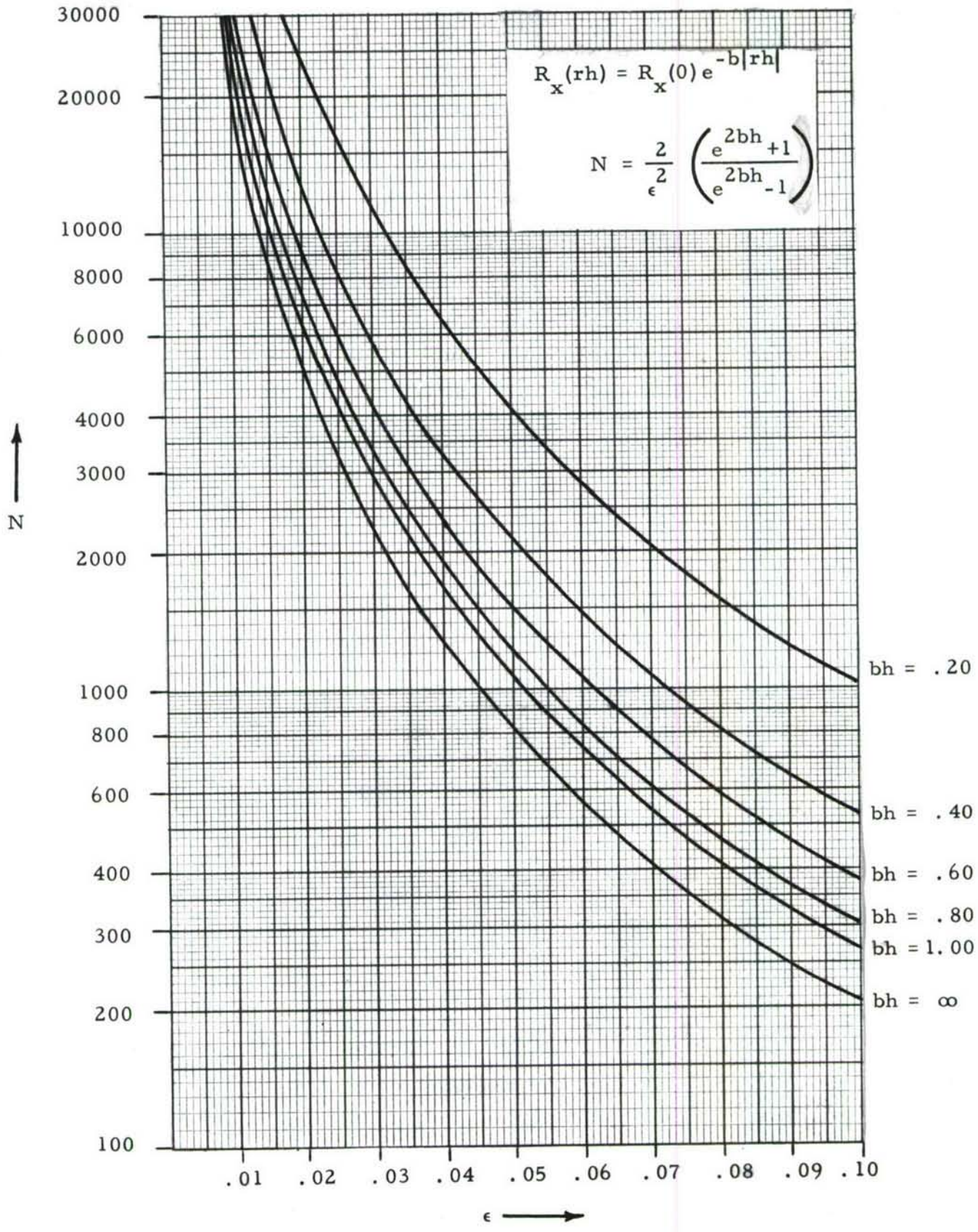


Figure 5. Standard Error for Correlated Products

5. POWER SPECTRUM ESTIMATES

As with the case of the correlation function, certain parameters must be specified in advance. These are the resolution bandwidth B of the analysis and the normalized standard error ϵ. The bandwidth B here should not be confused with the signal bandwidth. Depending on whether the analysis is to be performed with analog or digital methods, slightly different procedures are used.

Assume that x(t) is a Gaussian process with a true power spectrum G(f) which is approximately constant within the resolution bandwidth B. Then it may be shown that an estimate $\hat{G}(f)$ follows a χ^2 distribution with k degrees-of-freedom given by

$$\hat{G}(f) \sim \frac{G(f)\,\chi_k^2}{k} \tag{23}$$

Here it is assumed that m points of $\hat{G}(f)$ B cps apart are computed so that N = mk. The symbol "$\sim$" is to be read "distributed as." The variance of this quantity is then obtained directly as

$$\mathrm{Var}\left[\hat{G}(f)\right] = \frac{G^2(f)}{k^2}(2k) = G^2(f)\,\frac{2}{k} \tag{24}$$

The normalized standard error is

$$\epsilon = \frac{\sqrt{\mathrm{Var}\left[\hat{G}(f)\right]}}{G(f)} = \sqrt{\frac{2}{k}} = \frac{1}{\sqrt{BT}} \tag{25}$$

using k = 2BT with a proper interpretation here for B.

To illustrate Eq. (25), if it is desired to maintain ϵ at, say, 10%, and B = 10 cps, the required record length is

$$T = \frac{1}{B\epsilon^2} = \frac{1}{10(.01)} = 10 \text{ sec.}$$

and the required number of degrees-of-freedom in each individual point of $\hat{G}(f)$ is $k = 2BT = 200$.

A proper definition of the analysis bandwidth has not yet been given. It is actually a difficult problem to specify a proper correspondence between the degrees-of-freedom N and the BT product. In Reference 4 an "equivalent bandwidth" is given as

$$B_e = \frac{\left[\int_0^\infty G_{\Delta f}(f)\, df\right]^2}{\int_0^\infty \left[G_{\Delta f}(f)\right]^2 df} \tag{26}$$

where $G_{\Delta f}(f)$ is the power spectrum resulting from the process after it has been filtered by the analyzer filter. For practical purposes, either the noise bandwidth or half-power point bandwidth of the analyzer filter may be used instead of B_e due to the sharp cutoffs on modern spectrum analyzer filters. The details of these considerations are discussed in Reference 5.

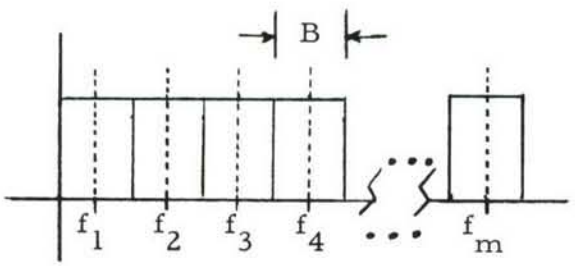
Therefore, in the case of analog power spectra computations, the half-power point bandwidth may be used in Eq. (25) for computing required record lengths. This, of course, assumes the filter bandwidth to be smaller than the signal bandwidth which must be the case for a proper analysis.

In the digital case, the analysis resolution bandwidth B is determined by the sampling interval Δt and the number of points, m, computed for the correlation function. In fact, for practical purposes, B is given by

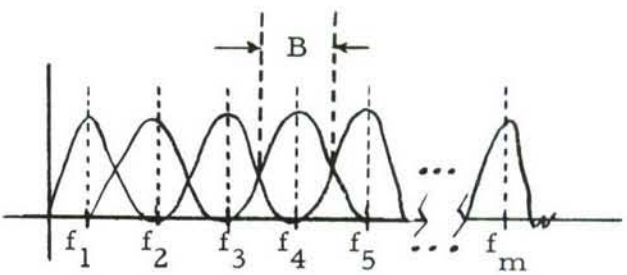
$$B = \frac{1}{m\Delta t} \tag{27}$$

As indicated in Reference 4, page 36, choosing B in this manner is not quite theoretically correct for actual filters but is satisfactory for almost all

practical purposes. This relation would be right if the filters effectively occurring in a usual digital computation had perfectly sharp cutoffs. In practice, they appear as indicated in the sketch below. The overlap of the filters creates a small amount of correlation in neighboring points of the spectrum estimate. This causes a slight inaccuracy in Eq.(17).



Ideal Filters for Digital Spectrum Estimates



Actual Filters for Digital Spectrum Estimates

A procedure for selecting the sample size and record length for a digital analysis is as follows. Assume frequencies up to a cutoff frequency, f_c = 2000 cps, are of interest and that a resolution bandwidth of B = 10 cps has been chosen. The maximum lag number for the correlation function is then

$$m = \frac{2f_c}{B} = 400 \tag{28}$$

To avoid aliasing below 2000 cps, the sampling frequency must be twice the frequency of interest, which accounts for the $2f_c$ in Eq. (28). This gives a sampling interval of

$$\Delta t = \frac{1}{2f_c} = \frac{1}{4000} = .00025 \text{ sec.}$$

Now if it is decided to maintain a normalized standard error $\epsilon = 10\%$, the required record length is

$$T = \frac{1}{B\epsilon^2} = \frac{1}{10(.01)} = 10 \text{ sec.}$$

The total number of observations required is

$$N = \frac{T}{\Delta t} = \frac{10}{.00025} = 40,000$$

The final calculations will then give 200 points of G(f) at 10 cps intervals each having 200 d.f.

There is a point to note when comparing the power spectra and correlation function estimates. When the assumption of a constant spectrum is approximately fulfilled, then one obtains independent estimates of G(f) while the point estimates for $R(\tau)$ are not independent but correlated. Therefore, the normalized standard error requirements only apply to any single given point of $R(\tau)$ at a time. However, the limits for G(f) apply to all computed points simultaneously. That is, one could draw the $\pm$ 10% confidence bounds about G(f) as a whole but not for $R(\tau)$ as a whole. However, the basic considerations for estimating sample sizes are not affected.

An additional point that one should realize is that specifying ϵ to be, say, 10% only means that about 68% of all the estimates obtained would be within $\pm$ 10% of the true value (assuming the estimates are normally distributed). Also, with this assumption, about 95% of the time, estimates will be within $\pm$ 20% of the true value. If one wants the estimates to be, say, within $\pm p\%$ of the true value 95% of the time, then one must choose $\epsilon = (p/2)\%$. Of course, if one draws a $\pm 1\epsilon = 10\%$ confidence band for 200 points in a power spectrum estimate, one would expect 68% of the estimates to be within 10% of the true value. This means 32% of 200 or 64 true points would be expected to lie outside the bands. One must adjust probabilities appropriately if it is desired for no true value to lie outside the confidence band with a given probability.

6. CROSS-CORRELATION ESTIMATES

Let $R_x(\tau)$ and $R_y(\tau)$ be the autocorrelation functions of x(t) and y(t) and let $R_{xy}(\tau)$ be the cross-correlation between x(t) and y(t). Then the variance of $R_{xy}(\tau)$ is

$$\mathrm{Var}\left[\hat{R}_{xy}(\tau)\right] = \frac{R_x(0)\,R_y(0) + R_{xy}^2(\tau)}{N} \tag{29}$$

Equation (29) is a direct generalization of Eq.(17) and is the formula for the variance when the processes are jointly Gaussian and the estimate of $R_{xy}(\tau)$ is based on N independent observations. The definition of the normalized standard error for $R_{xy}(\tau)$ is

$$\epsilon = \left\{\frac{\mathrm{Var}\left[\hat{R}_{xy}(\tau)\right]}{R_{xy}^2(\tau)}\right\}^{\frac{1}{2}} = \left\{\frac{\left[R_x(0)\,R_y(0)/\,R_{xy}^2(\tau)\right] + 1}{N}\right\}^{\frac{1}{2}} \tag{30}$$

In this case it is not convenient to talk about the normalized standard error at $\tau = 0$. The cross-correlation function $R_{xy}(\tau)$ does not necessarily have a maximum at $\tau = 0$ as do $R_x(\tau)$ and $R_y(\tau)$. However, one can show that the cross-correlation function is bounded by the product of the zero values of the two autocorrelation functions, namely,

$$R_{xy}^2(\tau) \leq R_x(0)\,R_y(0) \tag{31}$$

Therefore, if the cross-correlation function takes on a value close to the maximum possible value, then it makes sense to employ the error formula

$$\epsilon = \sqrt{\frac{2}{N}} = \sqrt{\frac{1}{BT}} \tag{32}$$

This is justified since if $R^2_{xy}(\tau)$ is close to $R_x(0)R_y(0)$, then

$$\frac{R_x(0)R_y(0)}{R^2_{xy}(\tau)} \approx 1$$

and Eq. (30) reduces to Eq. (32).

If one uses this relation for sample length requirements, a safety factor should be inserted since positive correlation in the observations will tend to reduce the effective sample size N. The relation N = 2BT then becomes less and less applicable. This is demonstrated in the graph of Figure 5 since increasing values of bh indicate larger and larger correlations of nearby sample points. No such convenient analytical guideline as Figure 5 is available for the cross-correlation case however since the forms of cross-correlation functions are not so conveniently classified.

A tacit assumption is made throughout this discussion that a common bandwidth B exists for the two signals. This, of course, is not necessarily true for practical applications. For experiment planning purposes a conservative choice should be made.

As an example, assume two signals x(t) and y(t) are to be cross correlated. If their bandwidths are estimated to be B_1 = 2000 cps and B_2 = 1000 cps, choose B = 1000 cps. Now, if a substantial peak is expected such as is the case when one signal is a time delayed version of another, then the relation

$$\epsilon = \frac{1}{\sqrt{BT}}$$

may be reasonably employed. This quantity represents a "percent of full scale" error now since the peak value of $R_{xy}(\tau)$ should be nearly as large as the product $R_x(0)R_y(0)$. If ϵ = 10% is the desired error, then for B = 1000 cps, the required record lengths are

$$T = 1/\epsilon^2 B = 0.1 \text{ sec.}$$

7. CROSS-SPECTRUM ESTIMATES

The cross-spectral density function is of major interest for the determination of frequency response functions of linear systems. Therefore, the discussion in the next Section 8 implicitly covers the most important aspects of cross-spectrum estimates. About the only time that the cross-spectrum would be of interest for its own sake is in determining a phase relation between two records. This, of course, would be equivalent to estimating a time delay between the two records in which case the cross-correlation function would be employed.

It is fortunate that the cross spectrum is not usually of direct interest itself. No convenient formulas exist for the variances and the sampling distributions are very complicated. However, the variance of the co-spectrum and quad-spectrum are bounded by (see Reference 6)

$$\begin{aligned} \mathrm{Var}\left[\hat{C}_{xy}(f)\right] &\leq \frac{G_x(f)\,G_y(f)}{BT} \\ \mathrm{Var}\left[\hat{Q}_{xy}(f)\right] &\leq \frac{G_x(f)\,G_y(f)}{BT} \end{aligned} \tag{33}$$

In the above equations

$$G_{xy}(f) = C_{xy}(f) - jQ_{xy}(f) \tag{34}$$

where $C_{xy}(f)$ is the real part (co-spectrum) of $G_{xy}(f)$ and $Q_{xy}(f)$ is the imaginary part (quad-spectrum) of $G_{xy}(f)$.

One encounters problems similar to that of cross correlation in trying to transform the quantity of Eq. (33) to a normalized standard error with respect to $\left|G_{xy}(f)\right|^2$ From basic theory it is known that

$$C_{xy}^2(f) + Q_{xy}^2(f) = \left|G_{xy}(f)\right|^2 \leq G_x(f)\,G_y(f) \tag{35}$$

but one does not know how much less $|G_{xy}(f)|^2$ is than $G_x(f)\,G_y(f)$. Therefore, in trying to divide by $|G_{xy}(f)|^2$, $C^2_{xy}(f)$, or $Q^2_{xy}(t)$ one cannot make any statement about the magnitude of the resulting normalized standard error. That is, if one defines

$$\epsilon^2(f) = \frac{G_x(f)\,G_y(f)}{|G_{xy}(f)|^2}\ \frac{1}{BT} \tag{36}$$

then an error formula of the usual form $(1/\sqrt{BT})$ can be employed only if the assumption that $G_x(f)\,G_y(f) \approx |G_{xy}(f)|^2$. This, at best, is most likely somewhat questionable. In fact, this is strictly true only in the case where one has a linear system relating $x(t)$ and $y(t)$ and there is no extraneous noise affecting the measurement of these quantities. Therefore, it is recommended that cross correlation or frequency response function error formulas be used for experiment planning purposes.

8. FREQUENCY RESPONSE FUNCTION ESTIMATES

Sampling variability for frequency response functions is more complicated than the previous functions. The frequency response is a complex-valued quantity which can be described in terms of a gain factor and a phase factor. The errors in gain factor estimates and phase factor estimates as a function of record length (sample size) are treated in Reference 7 and will be discussed here.

The frequency response function characterizes a linear system. If one knows the weighting function $h(t)$ for a constant parameter, time-invariant linear system, which is the response to a unit impulse input, then the frequency response function $H(f)$ is given as the Fourier transform of $h(t)$. In equation form,

$$H(f) = \int_{-\infty}^{\infty} h(t)\, e^{-j2\pi ft}\, dt \tag{37}$$

Also, $H(f)$ is a complex number in general and may be written in exponential form.

$$H(f) = \left| H(f) \right| e^{-j\phi(f)} \tag{38}$$

where $j = \sqrt{-1}$, $\left| H(f) \right|$ is the gain factor and $\phi(f)$ is the phase factor of the linear system.

When the frequency response function $H(f)$ is the end result of interest, a formula developed in Reference 7 gives error in the gain factor $\left| H(f) \right|$ and phase factor $\phi(f)$ of $H(f)$ as a function of the true coherence function $\gamma^2_{xy}(f)$ and degrees-of-freedom k.

The frequency response function $H(f)$ is related to the input power spectrum $G_x(f)$ and to the cross spectrum $G_{xy}(f)$ by the formula

$$H(f) = \frac{G_{xy}(f)}{G_x(f)} \tag{39}$$

The coherence function is also directly related to the input power spectrum $G_x(f)$, the output power spectrum $G_y(f)$, and to the cross spectrum $G_{xy}(f)$ by

$$\gamma_{xy}^2(f) = \frac{\left| G_{xy}(f) \right|^2}{G_x(f)\, G_y(f)} \tag{40}$$

The coherence function gives the degree of linear relationship (correlation) as a function of frequency, between the input $x(t)$ and the output $y(t)$.

The error formula for frequency response function measurements is

$$P = \text{Prob}\left[\left| \frac{\hat{H}(f) - H(f)}{H(f)} \right| < \sin\delta \quad \text{and} \left| \hat{\phi}(f) - \phi(f) \right| < \delta \right]$$

$$\approx 1 - \left[\frac{1 - \gamma_{xy}^2(f)}{1 - \gamma_{xy}^2(f)\cos^2\delta} \right]^{k/2} \tag{41}$$

where $k = 2BT$ degrees-of-freedom for the espectral estimates. For small values of δ, $\sin\delta \approx \delta$ so that both inequalities hold for the same numerical values. One applies the above formula by solving for k. Thus,

$$\left[\frac{1 - \gamma_{xy}^2(f)}{1 - \gamma_{xy}^2(f)\cos^2\delta} \right]^{k/2} = 1 - P$$

$$k \log \left[\frac{1 - \gamma_{xy}^2(f)}{1 - \gamma_{xy}^2(f)\cos^2\delta} \right] = 2 \log (1 - P)$$

$$k = \frac{2 \log (1 - P)}{\log \left[\frac{1 - \gamma_{xy}^2(f)}{1 - \gamma_{xy}^2(f)\cos^2\delta} \right]} \tag{42}$$

To apply Eq. (42), one chooses a value for δ, say 10% = .10, and a value for P, say P = .90. Assume for the moment $\gamma^2_{xy}(f)$ is known to be .90. Now, a value for k is calculated, in this case $k \approx 53$, which will maintain the sample gain factor $|\hat{H}(f)|$ within 10% of the true gain factor, and the sample phase $\hat{\phi}(f)$ within .10 radians of the true phase for approximately 90 out of 100 experiments. Different values for $\gamma^2_{xy}(f)$ lead to different k. This formula applies to one value of $|H(f)|$ and $\phi(f)$. One needs a total sample of N = mk for m points of the frequency response function.

The choice of a value in advance for $\gamma^2_{xy}(f)$ is strictly a matter of judgment if prior data is not available. From basic considerations, $0 \leq \gamma^2_{xy}(f) \leq 1$, analogous to the bounds on correlation coefficient. For purposes of planning an experiment, one must make a judgment based on the degree of linearity believed to exist and the amount of extraneous noise affecting the measurements. Both of these factors will reduce the coherence of the system from a theoretical maximum value of unity. Also note that coherence is a function of frequency so one must either restrict the range of frequency for which the computed k will apply or one must estimate a worst case in order to be conservative.

For convenience, several curves have been plotted giving k as a function of γ^2. Three sets of these curves are plotted corresponding to δ = .05, .10, and .15. In each set the curves correspond to P = .80, .85, and .90. These curves are displayed in Figure 6.

In converting k or N to a record length, the same considerations as for the ordinary power spectral density function apply. One does not have the problems associated with the cross-correlation function since in computing power spectra, the process is filtered by the analyzer (or effectively so in the case of digital methods) and the filter bandwidth is employed in the relation K = 2BT. In this situation one has the k required for a given accuracy of one estimate for a fairly narrow bandwidth B, where B is

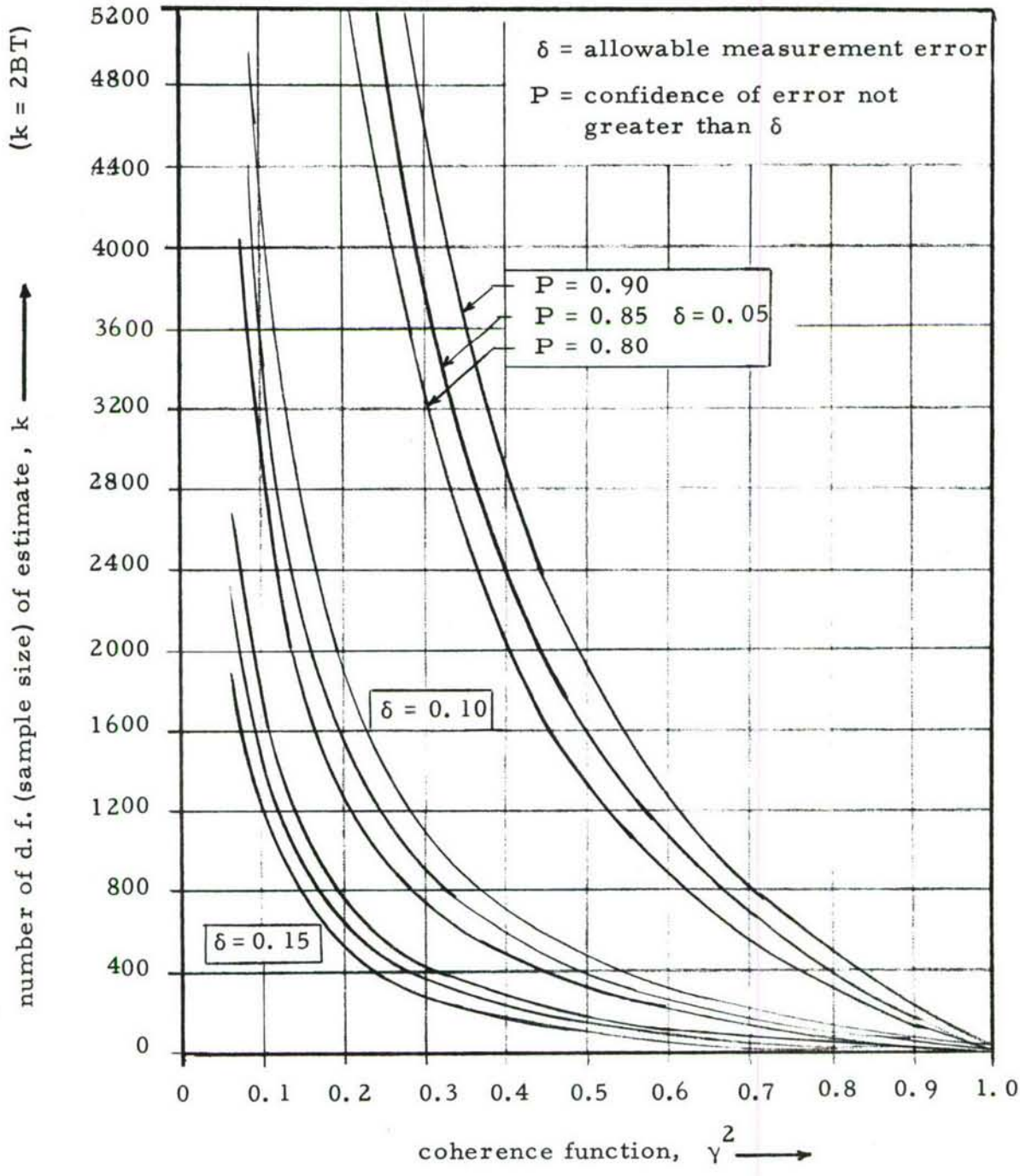


Figure 6. Data for Frequency Response Function Measurement Confidence

the analysis bandwidth. Therefore, since B is small, T must be relatively large. For the example where k = 53 and B = 10 cps,

$$T = \frac{k}{2B} = \frac{53}{20} = 2.65 \text{sec}$$

Another formula exists from which one may obtain confidence bands which are a function of the sample quantities obtained after the experiment has been performed. This is to be contrasted with the previous formula which is useful for planning where the true coherence must be estimated in advance. The $(1 - \alpha)$ confidence limits for gain and phase are given by

$$\left|\hat{H}(f)\right| - \hat{r}(f) \leq \left|H(f)\right| \leq \left|\hat{H}(f)\right| + \hat{r}(f) \tag{43}$$

$$\hat{\phi}(f) - \Delta\hat{\phi}(f) \leq \phi(f) \leq \hat{\phi}(f) + \Delta\hat{\phi}(f) \tag{44}$$

where

$$\hat{r}(f) = \left[\frac{1}{BT-1} F_{1-\alpha}(2, 2BT-2) \frac{(1 - \hat{\gamma}_{xy}^2)\hat{G}_y(f)}{\hat{G}_x(f)}\right]^{\frac{1}{2}} \tag{45}$$

and

$$\Delta\hat{\phi}(f) = \text{Arc sin} \frac{\hat{r}(f)}{\left|\hat{H}(f)\right|} \tag{46}$$

In Eq. (45), $F_{1-\alpha}(2, 2BT-2)$ is the $(1 - \alpha)$ percentile of the standard F distribution with degrees-of-freedom, $n_1 = 2$ and $n_2 = (2BT-2)$. Hence, Eqs. (43) and (44) give bounds that include the true gain $\left|H(f)\right|$ and true phase $\phi(f)$ with confidence $(1 - \alpha)$. Note that all quantities involved in the relations are sample values. These formulas are all special cases of the general equations found in Reference 8.

To illustrate these formulas, suppose the following values are obtained from a frequency response function estimation experiment.

$$\hat{G}_x(f_0) = 0.20\ g^2/\text{cps}$$

$$\hat{G}_y(f_0) = 0.10\ g^2/\text{cps}$$

$$\hat{\gamma}^2_{xy}(f_0) = 0.80$$

$$\left|\hat{H}(f_0)\right|^2 = 0.40$$

$$\hat{\phi}(f_0) = \pi/4 \text{ radians} = 45^\circ$$

$$B = 10 \text{ cps}$$

$$T = 1 \text{ sec}$$

$$\alpha = .05$$

$$F_{.95}(2, 18) = 7.21 \quad \text{(see Reference 1 tables for example)}$$

Note that in these hypothetical values, the square of the gain factor does not equal the ratio of the output to the input spectra. This might happen in practice as a result of the effects of extraneous noise or nonlinearities. From Eqs. (45) and (46), the following values are obtained.

$$\hat{r}(f_0) = \left[\frac{1}{10-1}\,(7.21)\,\frac{(1 - 0.80)(0.10)}{0.20}\right]^{\frac{1}{2}} = .283$$

$$\Delta\hat{\phi}(f_0) = \text{Arc sin}\,\frac{.283}{.632} = 26^\circ 37'$$

There, 95% confidence intervals corresponding to Eqs. (43) and (44) are:

$$.63 - .28 \leq \left|H(f_0)\right| \leq .63 + .28$$

$$45^\circ - 26^\circ 37' \leq \phi(f_0) \leq 45^\circ + 26^\circ 37'$$

9. PROBABILITY DENSITY ESTIMATES

Considerable experimental and theoretical work is still in the process of being performed to develop proper error formulas for probability density estimates. Experiments have been conducted in the past and are described in Reference 9. The use of the error formula developed experimentally in that report is recommended at the present time. This formula will now be presented along with its limitations.

If one neglects certain bias terms which are unimportant in usual applications, then theory predicts a variance for probability density estimates of

$$\epsilon^2\left[\hat{p}(x)\right] = \frac{\mathrm{Var}\left[\hat{p}(x)\right]}{E^2\left[\hat{p}(x)\right]} = \frac{1}{N\,p(x)\,\Delta x} = \frac{1}{2BT\,p(x)\,\Delta x} \qquad (47)$$

In Eq. (47), B is the bandwidth of the process where a perfectly sharp cutoff in the spectrum is assumed. Also, T is record length, p(x) is the true value of the probability density, Δx is the amplitude "window" or resolution of the measurement, and N is the number of independent observations (sample size) used in the estimate.

The requirement of independent samples is not fulfilled in existing analog measurements nor is it necessarily in digital procedures. Experiments were performed (Reference 9) which indicate that this is a significant factor. The results of those experiments indicate that a usable error formula is

$$\epsilon\left[\hat{p}(x)\right] = \frac{0.20}{\sqrt{BT\,p(x)\,\Delta x}} \qquad (48)$$

It must be emphasized that the above equation was developed only for one specific instrument and is possibly valid for only that instrument. For example, if a process was digitally sampled, and the observations were uncorrelated, then, if the density function was calculated on a digital computer,

Eq. (47) rather than Eq. (48) applies. Experiments are presently being designed so that error formulas similar to Eq. (48) may be developed for other specific instruments.

In Eq. (48), $\hat{p}(x)$ would be used rather than the true value $p(x)$ if one was establishing limits about a measurement of $p(x)$. Also, in practice, noise bandwidth or half-power bandwidth can be used for B. As mentioned above and described in Reference 9, the above formula was obtained with only one specific instrument and only approximately Gaussian signals were analyzed. Therefore, as with most other error formulas in existence, one must be prudent in its application when the underlying assumptions (such as non-Gaussian noise and different instruments) are not satisfied. However, Eq. (48) represents the best available result and does provide one with reasonable guidelines for experiment planning purposes.

As an example of the application of Eq. (48), assume one has a signal $x(t)$ with a flat spectrum out to $B = 2000$ cps. Further assume an error $\epsilon = 1\%$ is desired for a point one standard deviation (1.0σ) away from the mean and that the resolution is to be $\Delta x = 0.1\sigma$. For planning purposes, suppose one expects a near normal density function. Then one obtains $p(1.0\sigma) = .242$ from tables of the normal density function. The required record length for the experiment then is

$$T = \frac{.04}{B\,p(x)\,\Delta x\,\epsilon^2} = \frac{.04}{2000(.242)(.1)(.0001)} = 8.26 \text{ sec} \qquad (49)$$

For expected density functions other than Gaussian, one substitutes the appropriate value for $p(x)$. Also, it will be most convenient to work in terms of standardized units, e.g., x/σ, rather than any absolute terms.

The final fact to be emphasized is that Eq. (47) or Eq. (48) apply only to a given single point selected in advance on the probability density function. The correlation from one point estimate to the next is not known and one cannot draw confidence bands about the entire curve simultaneously. One can only do this for one given individual point.

10. THE WEIBULL DISTRIBUTION FOR FATIGUE TESTS

When constructing a statistical model for life length or fatigue failure rate of a structure one often finds that an assumption of normality is not satisfied. For example, many life length distributions are markedly skewed. The instantaneous failure rate or so-called hazard rate (see Section 10.1 for the definition) of the normal distribution is a strictly increasing linear function of time which is not desirable in many structural fatigue models.

In order to describe the random behavior of fatigue life, a number of probability distributions have been proposed. Among these, the exponential distribution is best known and most widely used in electronic, chemical, and other application areas.

The exponential distribution has a number of desirable statistical properties, but its usefulness is limited because of the following property: If the life length T of a structure has an exponential distribution, then previous use does not affect its future life length. This fact is easily seen in the following relation.

Let T be the random variable distributed with an exponential probability density.

$$f(t) = \frac{1}{\alpha} e^{-t/\alpha} \qquad \text{if} \quad t \geq 0$$

$$= 0 \qquad \text{if} \quad t < 0$$

Then

$$P(T > a+b \mid T > b) = \frac{P(T > a+b,\ T > b)}{P(T > b)}$$

Now, if $T > a+b$, then it is simultaneously larger than b, hence

$$\frac{P(T > a+b,\ T > b)}{P(T > b)} = \frac{P(T > a+b)}{P(T > b)}$$

$$= \frac{e^{-(a+b)/\alpha}}{e^{-a/\alpha}} = e^{-b/\alpha} = P(T > b)$$

In words, given that a structure has lasted for b or more units of time, then probability of lasting "a" or more additional units of time is the same as the probability of a new unit lasting a + b or more units of time. In short, a characteristic of the exponential distribution is a constant failure rate when it is used as a failure rate distribution. That is, the average number of failures which occur in a unit time period remains constant with time. Thus, if a structure has a constant failure rate the exponential distribution is a tailormade model.

In general, the distribution of life length for some object has a unique shape, scale, and location. For example, any particular structure has a unique failure rate. The Weibull distribution has three parameters which determine shape, scale, and location. Thus, the life length of many structures can be suitably modeled by determining each of three parameters. Among others, one of the advantages of using the Weibull distribution is to be able to model so many different fatigue life distributions; it can be exponential, Rayleigh or approximately normal. The following sections present some of the statistical properties and applications of the Weibull distribution.

10.1 DEFINITION OF THE WEIBULL DISTRIBUTION

The general form of the Weibull distribution is:

$$W(t) = \begin{cases} 1 - e^{-\left[(t-\gamma)/\alpha\right]^{\beta}} & \text{if } t > \gamma \\ 0 & \text{if } t \leq \gamma \end{cases} \tag{50}$$

and the density function is

$$w(t) = \begin{cases} \beta \dfrac{(t-\gamma)^{\beta-1}}{\alpha^{\beta}} e^{-\left[(t-\gamma)/\alpha\right]^{\beta}} & \text{if } t > \gamma \\ 0 & \text{if } t < \gamma \end{cases} \tag{51}$$

In the above equations, α, β, and γ are parameters of the distribution generally named as follows:

α = scale parameter

β = shape parameter

γ = location parameter

(α and β are not to be confused with level of significance and probability of Type II Error which they often denote.) The parameter α is analogous to the variance of the normal distribution in that it's value affects the scaling of the distribution. The parameter γ is a location parameter as is the mean of a normal distribution in that it translates the distribution. This parameter γ may be interpreted in life length testing as the minimum length of time that passes before any failure can occur. The parameter β is termed the shape parameter since it affects the basic shape of the distribution.

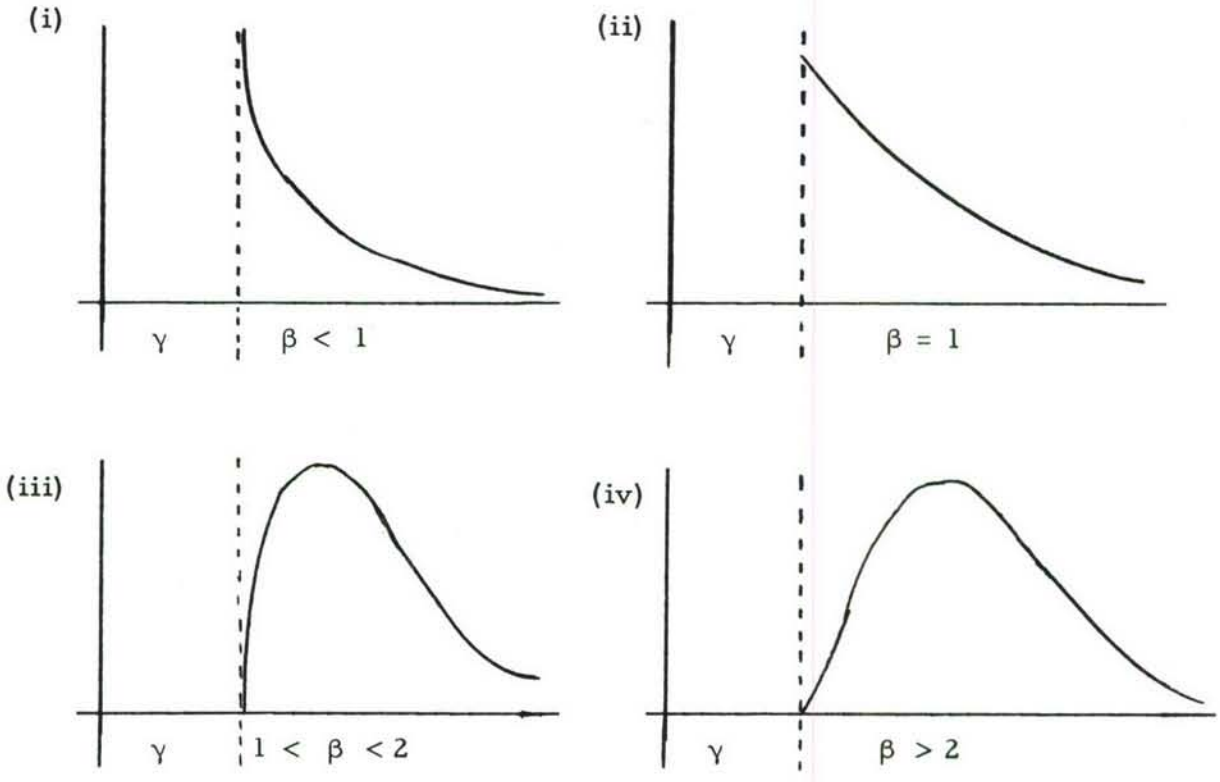


Figure 7. Four Shapes of the Weibull Density Function

Four different shapes depending on β are illustrated in Figure 7. The shape parameter β describes the mode of failure. Thus, for $\beta = 1$ the failure rate is constant over time. $\beta < 1$ indicates that the failure rate is a decreasing function of time, while for $\beta > 1$ the rate is increasing with time. A more descriptive way of looking at a statistical fatigue model is by considering the (instantaneous) failure rate or so-called hazard rate. It is the instantaneous rate of change in failure probability. That is, the hazard rate $H(t)$ is defined by

$$H(t) = \lim_{\Delta t \to 0} \frac{P(T > t) - P(T > t+\Delta t)}{\Delta t\, P(T > t)} = \frac{f(t)}{P(T > t)}$$

where $f(t)$ is the value of density function at $T = t$. Thus, the hazard rate is the density function of time to failure given that the system has not failed prior to time t.

In the case of the Weibull distribution, the hazard rate is

$$H(t) = \beta \frac{(t - \gamma)^{\beta - 1}}{\alpha^{\beta}} \tag{52}$$

Equation (52) shows that the Weibull distribution reduces to the exponential distribution when $\beta = 1$. From Eq. (52), it is noted that the exponential distribution is associated with a constant hazard rate $1/\alpha$. This fact was shown using basic probability notations in the introduction to Section 10.

The hazard rates of the Weibull distribution for several values of β and with the other parameters fixed are illustrated in Figure 8.

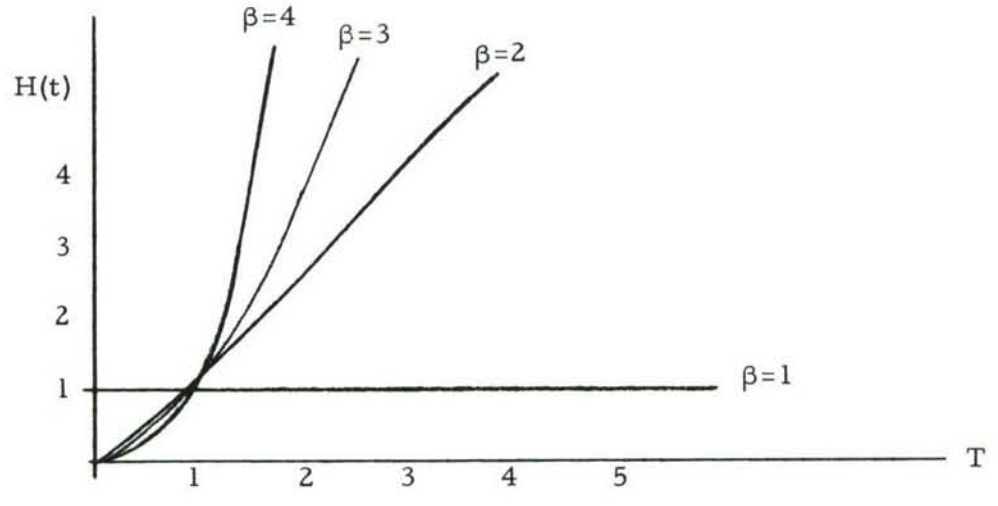


Figure 8. H(t) with $\gamma = 0$, $\alpha = 1$, $\beta = 1, 2, 3, 4$

The shape parameter β may be interpreted in terms of the hazard rate as follows:

$\beta > 1$ an increasing hazard rate

$\beta = 1$ a constant hazard rate

$\beta < 1$ a decreasing hazard rate

Let $b = 1/\beta$. Then one finds the following values of the first two moments and the median for the Weibull distribution (Reference 10):

$$\text{mean} = \mu = \gamma + \alpha(b!) \tag{53}$$

$$\text{variance} = \sigma^2 = \alpha^2\left[\, 2(b!) - (b!)^2\right] \tag{54}$$

$$\text{median} = m_{.5} = \gamma + \alpha(\log 2)^b \tag{55}$$

When $\beta = 3.57$, then the Weibull distribution becomes a good approximation of the normal distribution. This is because

$$b! = \Gamma\left(\frac{4.57}{3.57}\right) \approx (.693)^{1/3.57} = (\log 2)^b \tag{56}$$

and the mean and median are approximately equal when $\beta = 3.57$.

Note that the distribution has positive skewness if $\beta < 3.57$. The Weibull distribution becomes the Rayleigh distribution (χ^2 with 2 d.f.) when $\beta = 2$.

10.2 WEIBULL PARAMETER ESTIMATES

When the Weibull distribution is assumed as a statistical model, the problem of estimating three parameters arises. Two methods, maximum likelihood and minimum chi-square methods are generally applied to estimation problems. It is not intended to give an exhaustive discussion here on statistical inference concerning the Weibull distribution. Only a few applicable results are presented.

(a) Maximum likelihood estimate (M.L.E) of α^{β} when γ and β are known. The likelihood function to be maximized is

$$L\left[w(t)\right] = \frac{\beta^n}{\alpha^{n\beta}} \prod_{i=1}^{n} (T_i - \gamma)^{\beta-1} \exp\left[- \sum_{i=1}^{n} \left(\frac{T_i - \gamma}{\alpha} \right)^{\beta} \right] \qquad (57)$$

where $T_1, T_2, \ldots, T_n$ are the sample data (times to failure). Logarithms are now taken which simplify the solution of Eq. (57):

$$\log L\left[w(t)\right] = n \log\left(\frac{\beta}{\alpha^{\beta}}\right) + \sum_{i=1}^{n} \log (T_i - \gamma)^{\beta-1} - \sum_{i=1}^{n} \left(\frac{T_i - \gamma}{\alpha} \right)^{\beta} \qquad (58)$$

When Eq. (58) is maximized with respect to α^{β} assuming γ and β are known constants, one finds

$$\hat{\alpha^{\beta}} = \frac{\sum_{i=1}^{n} (T_i - \gamma)^{\beta}}{n} \qquad (59)$$

(b) M.L.E's of β and γ. When Eq. (58) is maximized simultaneously with respect to β and γ, one obtains the two equations

$$\frac{n}{\beta} - n \log \alpha + \sum_{i=1}^{n} \log (T_i - \gamma) - \sum_{i=1}^{n} \left(\frac{T_i - \gamma}{\alpha} \right)^{\beta} \log \left(\frac{T_i - \gamma}{\alpha} \right) = 0 \qquad (60)$$

$$-(\beta - 1) \sum_{i=1}^{n} \frac{1}{T_i - \gamma} + \frac{\beta}{\alpha} \sum_{i=1}^{n} \left(\frac{T_i - \gamma}{\alpha} \right)^{\beta - 1} = 0 \qquad (61)$$

These two equations along with Eq. (59) have to be solved to yield M.L.E's for α, β, and γ. This cannot be done explicitly (Reference 10), but an iterative solution by computer is possible.

(c) A very useful result is derived by Menon (Reference 11). Let $x_i = T_i - \gamma$. Let y_i be formed of those negative values obtained from $\log x_i - \log \alpha$. while the z_i are formed of those positive values obtained from $\log x_i - \log \alpha$. That is, define y_i and z_i as follows.

$$y_i = \begin{cases} \log x_i - \log \alpha & \text{if } \log x_i - \log \alpha < 0 \\ 0 & \text{otherwise} \end{cases}$$

$$z_i = \begin{cases} \log x_i - \log \alpha & \text{if } \log x_i - \log \alpha > 0 \\ 0 & \text{otherwise} \end{cases}$$

Then, when α and γ are known,

$$\hat{\beta} = \frac{n}{-.74 \sum_{i=1}^{n} y_i + 1.85 \sum_{i=1}^{n} z_i} \qquad (62)$$

$$\text{Var}(\hat{b}) = \frac{.66\, b^2}{n} \qquad (63)$$

where $b = 1/\beta$. Note that the summations on y_i and z_i run to n since n values for each of these variables is defined even though some of the values are zero. This is convenient for this theoretical work but for computational purposes, one summation would run from 1 to p and the other from 1 to q where $N = p + q$.

When α is unknown but γ is known,

$$\hat{\beta} = \left\{ \frac{1}{6} \left[\frac{\pi^2 (n-1)}{\sum_{i=1}^{n} (\log x_i)^2 - \left(\sum_{i=1}^{n} \log x_i \right)^2 \Big/ n} \right] \right\}^{\frac{1}{2}} \tag{64}$$

$$\text{Var}(\hat{b}) \approx \frac{1.1\, b^2}{n} \tag{65}$$

where $b = 1/\beta$.

Example: Suppose

$$\sum_{i=1}^{15} \log y_i = -.5 \quad , \quad \sum_{i=1}^{15} \log z_i = 5.4$$

Then, using Eq. (62), one obtains

$$\hat{\beta} = \frac{1}{(-.74)(-.5) + (1.85)(5.4)} = .0965$$

and using Eq. (63), the variance of $\hat{b}$ is computed to be

$$\text{Var}(\hat{b}) = \frac{(.66)\left(\frac{1}{.0965}\right)^2}{15} = 4.72$$

(d) Graphical methods may be used for estimation. Through the use of the Weibull probability graph paper, a simple method is available for obtaining estimates of the parameters α and β. This technique is discussed in Reference 12.

10.3 WEIBULL PARAMETER CONFIDENCE LIMITS

(a) Confidence limits for α^β when β and γ are known are now given. Let $y_i = T_i - \gamma$ for $i = 1, 2, \ldots, n$ and $y = T - \gamma$. Then

$$P(y^\beta < y_0) = P\left(y < y_0^{1/\beta}\right)$$

$$= 1 - e^{-y_0/\alpha^\beta}$$

It is clear that y^β has an exponential distribution with a single parameter α^β. Reference 13 shows that $2n\hat{\alpha}^\beta/\alpha^\beta$ is distributed as chi-square with $2n$ degrees-of-freedom (χ^2_{2n}). Thus, the desired $(1 - \epsilon)$ confidence interval is defined by the equation

$$P\left[\chi^2_{2n(\epsilon/2)} < \frac{2n\hat{\alpha}^\beta}{\alpha^\beta} < \chi^2_{2n(1-\epsilon/2)}\right] = P\left[\frac{2n\hat{\alpha}^\beta}{\chi^2_{2n(1-\epsilon/2)}} < \alpha^\beta < \frac{2n\hat{\alpha}^\beta}{\chi^2_{2n(\epsilon/2)}}\right] = 1 - \epsilon \qquad (66)$$

where $\chi^2_{2n(\epsilon/2)}$ and $\chi^2_{2n(1-\epsilon/2)}$ are lower and upper tail $\epsilon/2$ percentiles of the χ^2_{2n} distribution.

Example: Suppose $\alpha^\beta = 30$ (hours), $n = 9$, and $\epsilon = .1$. Then

$$P\left(\frac{540}{28.87} < \alpha^\beta < \frac{540}{9.39}\right) = P(18.70 < \alpha^\beta < 57.51) = .9$$

Thus, 90% confidence limits for α^β is (18.70, 57.51).

(b) Rather conservative confidence limits for β can be constructed using Chebyshev's inequality and Eqs. (62) and (63) or (64) and (65). Chebyshev's inequality states that

$$P\left[\left|\hat{b} - b\right| < \epsilon\sqrt{\text{Var } \hat{b}}\right] \geq 1 - \frac{1}{\epsilon^2}$$

or

$$P\left[\hat{b} - \epsilon\sqrt{\text{Var } \hat{b}} < b < \hat{b} + \epsilon\sqrt{\text{Var } \hat{b}}\right] \geq 1 - \frac{1}{\epsilon^2}$$

Thus, if α is known

$$P\left[\hat{\beta}(1 - \delta) < \beta < \hat{\beta}(1 + \delta)\right] \geq 1 - \frac{.66}{n\delta^2} \tag{67}$$

and if α is unknown

$$P\left[\hat{\beta}(1 - \delta) < \beta < \hat{\beta}(1 + \delta)\right] \geq 1 - \frac{1.1}{n\delta^2}$$

where $\delta > 0$ is a constant.

Example: Suppose one wishes to construct 90% confidence limits on β given the data of the example in Section 10.2(c). Let $1 - \frac{.66}{n\delta^2} = .9$, then $\delta = .663$. When γ and α are known, one obtains from Eqs. (61), (62), and (66)

$$P\left[.0965(1-.663) < \beta < .0965(1+.663)\right] = P(.033 < \beta < .160) \geq .9$$

Thus, the 90% confidence limits for β are (.033, .160).

10.4 HYPOTHESIS TESTS

(a) Suppose one wishes to test the hypothesis $H_0: \theta = \theta_0$, (θ may be α, β, or γ).

The likelihood ratio test is used here when sample size n is large. Let

$$\lambda = \frac{\prod_{i=1}^{n} w(t_i | \theta_0)}{\prod_{i=1}^{n} w(t_i | \hat{\theta})} \tag{68}$$

where $w(t_i | \theta_0)$ is the density function given in Eq. (51) when $\theta = \theta_0$. Similarly, $w(t_i | \hat{\theta})$ denotes the value of the density function when $\theta = \hat{\theta}$ where $\hat{\theta}$ is the M. L. E. of θ. When n is fairly large $-2\log_e \lambda$ is approximated by the χ_1^2 distribution. Hence, a value of λ is obtained from the sample data and $-2\log \lambda$ may be compared with an appropriate value of χ_1^2 obtained from a table.

Example: Suppose one wishes to test the hypothesis $H_0: \beta = 1.5$ with sample size $n = 30$. Assume that the computation by Eq. (68) yields the value $\lambda = .11$. Then $-2 \log_e \lambda = 4.41$. Since the 5% critical value of χ_1^2 is 3.84, the hypothesis is rejected at 95% confidence level.

(b) The χ^2 distribution may be used for a Weibull goodness of fit test. A criterion for testing the goodness of fit of a Weibull distribution from n samples is

$$X_{k-2}^2 = \sum_{i=1}^{k} \frac{f_i^2}{np_i} - n \tag{69}$$

where p_i is the probability that a failure occurs in the interval t_{i-1} to t_i. f_i is the actual number of failures in the interval t_{i-1} to t_i. The quantity k is the number of cells such that approximately $np_i \geq 5$ for each i. If this condition is satisfied then X_{k-2}^2 is approximated by χ_{k-2}^2 distribution.

10.5 RELIABILITY PROBLEMS

Suppose one has several systems each of which could be used for the same purpose. Very often one has to choose the one which is best suited for the purpose. Define reliability $R(t)$ as the probability that a system will perform satisfactorily for at least a given period of time t. In the case of the Weibull distribution

$$R(t) = P(T > t) = e^{\left[(t-\gamma)/\alpha\right]^{\beta}}$$

A commonly used decision procedure is to choose a system with the maximum reliability $R(t_0)$ if the system has to last to time t_0 in case of non-replacement policy. This policy applies when the system is not replaceable or one does not wish to replace the system if it fails. Another situation is the replacement policy; this is a policy of immediate replacement when the system fails. A simple decision procedure in this case is to choose a system with the maximum mean life.

Example: Suppose a structure is characterized by $\gamma = 10$ hours, $\alpha = 80$ hours, and $\beta = 2$. Then

$$R(50) = e^{-\left[(50-10)/80\right]^2} = .779$$

$$\text{mean life} = 10 + 80\left(\frac{1}{2}!\right) = 80.9 \text{ hours}$$

That is, the probability that the structure will last at least 50 hours is .779. The graph of $R(t)$ for the Weibull distributions with some fixed parameters is given in Figure 9.

The above mentioned decision procedure when the system is non-replaceable, does not account for the average reliability or cost which in general are functions of time. Thus, it seems that a more sophisticated technique is necessary depending on the situation. The following two procedures are proposed examples.

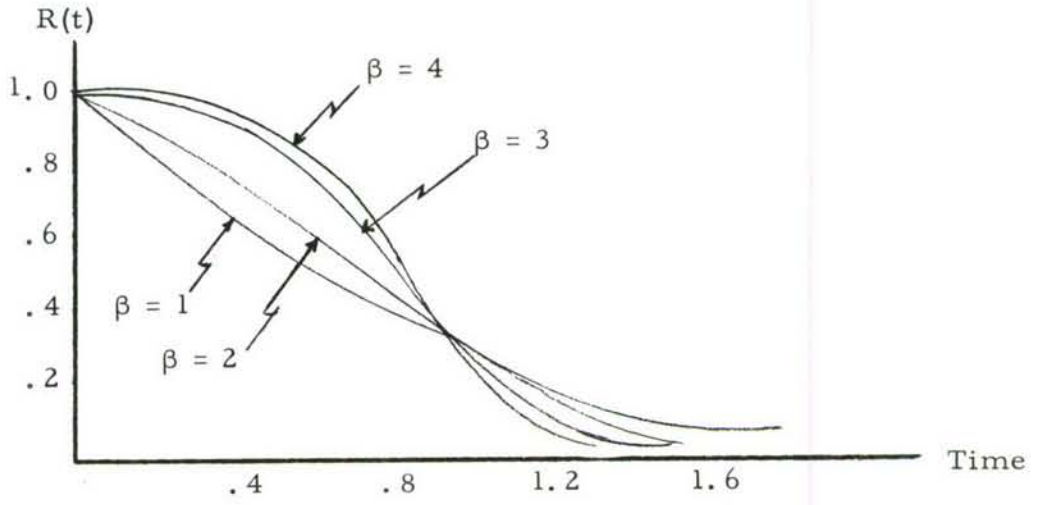


Figure 9, $R(t) = e^{-t^{\beta}}$ ($\gamma = 0$ and $\alpha = 1$)

(a) To guard against the worst case select the system j associated with θ_j which will maximize

$$\min_t \phi\left\{c(t)\right\} R(t|\theta_j) \tag{70}$$

where $\phi\left\{c(t)\right\}$ is a function of the cost at time t. $R(t|\theta_j)$ is a reliability of a system with the parameter θ_j. This decision procedure is an extremely conservative one. One simply selects the system which is the best in the worst case. (See Figure 10)

(b) Suppose one is interested in over-all performance during the time of operation t_0. In this case a better decision is to select the system j associated with θ_j which will maximize

$$\int_0^t \phi\left\{c(t)\right\} R(t|\theta_j)\, dt \tag{71}$$

That is, a system which is the best in average during the time of operation should be selected. In Figure 10, one clearly prefers system 3 over systems 1 and 2 by the above method (a) which only considers the point t_0 in this case, but if one is interested in over-all operation time, system 1 is most preferable.

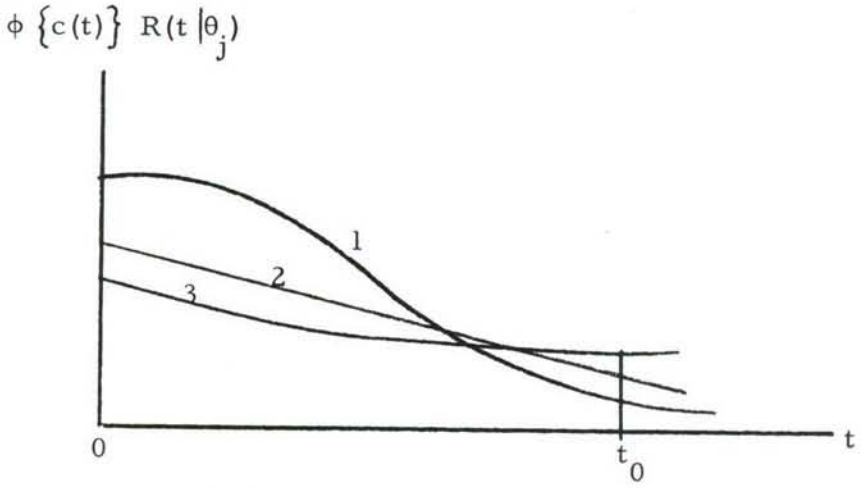


Figure 10. $\phi\{c(t)\}\ R(t|\theta_j)$

Example.

Assume $\phi\{c(t)\} = 1$ for $j = 1$ or 2. Suppose that $t_0 = 2$, $\gamma = 0$, $\alpha = 1$. If there is a choice of two systems with $\beta = 1$ and $\beta = 2$, then since $e^{-2^2} < e^{-1}$ the decision by Eq. (70) will choose a system with $\beta = 1$ to guard against the worst case. While the relation

$$\int_0^2 e^{-t^2}\,dt > \int_0^2 e^{-t}\,dt$$

indicates that a system with $\beta = 2$ is better if one is interested in average reliability during the time period of operation.

10.6 SAMPLE SIZE

Methods of determining sample sizes needed for estimating various quantities are proposed here.

(a) Suppose a confidence interval about α^{β} is desired whose length is L. The parameters β and γ are assumed to be known, (see Section 10.6 (c) below). The sample size n is determined such that n satisfies the following two conditions:

$$\frac{2n\hat{\alpha}^{\hat{\beta}}}{\chi^2_{2n(\epsilon/2)}} - \frac{2n\hat{\alpha}^{\hat{\beta}}}{\chi^2_{2n(1-\epsilon/2)}} \leq L$$

and

$$P\left[\frac{2n\hat{\alpha}^{\hat{\beta}}}{\chi^2_{2n(1-\epsilon/2}} < \alpha^{\hat{\beta}} < \frac{2n\hat{\alpha}^{\hat{\beta}}}{\chi^2_{2n(\epsilon/2)}}\right] = 1 - \epsilon$$

(b) Now consider a confidence interval about β whose desired length is $\leq$ L. Assume that α and γ are known. Equation (67) states:

$$P\left[\hat{\beta}(1-\delta) < \beta < \hat{\beta}(1+\delta)\right] \geq 1 - \frac{.66}{n\delta^2} \qquad (72)$$

where $\delta > 0$ is a constant. If the confidence coefficient is p_0, then

$$1 - \frac{.66}{n\delta^2} = p_0 \quad \text{or} \quad \delta = \sqrt{\frac{.66}{(1-p_0)n}}$$

Thus,

$$2\delta\hat{\beta} = 2\sqrt{\frac{.66}{(1-p_0)n}}\,\hat{\beta} = 1.62\hat{\beta}\sqrt{\frac{1}{(1-p_0)n}} \leq L$$

and

$$\frac{2.62\hat{\beta}^2}{(1-p_0)L^2} \leq n$$

Since the inequality (72) has two unknowns n and $\hat{\beta}^2$, n has to be determined iteratively. That is, n is increased until the inequality (72) is satisfied. This procedure is illustrated in the following example.

Example: Suppose that one wishes to estimate β by Eqs. (62) and (63) within an accuracy of $\pm .5$ with confidence of 90%. One has to determine n from inequality (62) iteratively. Suppose one guesses $n = 16$, and obtains an estimate $\hat{\beta} = 2$. Then inequality (62) is not satisfied since

$$\frac{(2.62)(4)}{(.1)(1)} = 104.8 > 16$$

Now suppose a sample of size $n = 100$ is used, and an estimate $\hat{\beta} = 1.95$ is obtained. The inequality (62) is satisfied, thus 100 is a sufficiently large sample size for the purpose.

$$\frac{(2.62)(1.95)^2}{(.1)(1)} = 99.6 < 100$$

(c) A confidence interval about β whose expected length is L when γ is known but α is unknown will now be derived. A similar approach to (b) yields

$$1 - \frac{1.1}{n\delta^2} = p_0 \quad \text{or} \quad \delta = \sqrt{\frac{1.1}{(1 - p_0)n}}$$

where p_0 is the confidence coefficient. The desired sample size n is given by

$$\frac{4.39\hat{\beta}^2}{(1 - p_0)L^2} \leq n$$

Again the above inequality has to be solved for n iteratively.

(d) As a further example for sample size computations, it will be assumed that $\beta = 2$. This would seem to correspond to a reasonable idea of the shape of a density function for time to fatigue failure. Choosing a specific reasonable value for the parameter β and further assuming that γ is known simplifies many of the parameter estimation problems. It has been noted that for $\beta = 2$, the Weibull distribution becomes a χ^2 distribution with two degrees-of-freedom if the scale parameter is chosen as $\alpha = \sqrt{2}$ and the location parameter $\gamma = 0$. This is also the Rayleigh distribution. The equation for the density function in this case is

$$W(x) = x e^{-x^2/2} \tag{73}$$

It can be shown (Reference 10) that the mean value of a random variable x having the distribution $W(x)$ is (when $\beta = 2$)

$$E(x) = \gamma + \alpha \frac{\sqrt{\pi}}{2} \tag{74}$$

and the variance is

$$\text{Var}(x) = \alpha^2 \left[\Gamma(2) - \Gamma^2\left(\frac{3}{2}\right) \right] = \alpha^2 \left[1 - \frac{\pi}{4} \right] \tag{75}$$

where $\Gamma(n)$ is the Gamma function.

To obtain an estimate of the sample size needed to estimate the mean of the distribution (the mean time to fatigue failure) one examines estimates of $E(x)$ of Eq. (74). If γ is assumed to be known, then the maximum likelihood estimate of α^2 is

$$\hat{\alpha}^2 = \frac{\sum_{i=1}^{N} x_i^2}{N} \tag{76}$$

where $x_i = T_i - \gamma$ and T_i is the time to failure of the ith test item. The total number of test items is denoted by N. From Section 10.3 the quantity $2n\hat{\alpha}^2/\alpha^2$ has a χ^2 distribution with 2N d. f. Hence, $\hat{\alpha}^2$ has the distribution

$$\alpha^2 \sim \frac{\alpha^2 \chi^2(2N)}{2N} \tag{77}$$

Therefore, the expected value and variance of $\hat{\alpha}^2$ are

$$E(\hat{\alpha}^2) = \frac{\hat{\alpha}^2}{2N} \; E\left[\chi^2(2N)\right] = \alpha^2 \tag{78}$$

and

$$\text{Var}\,(\hat{\alpha}^2) = \frac{\alpha^4}{4N^2} \; \text{Var}\left[\chi^2(2N)\right] = \frac{\alpha^4}{4N^2}\, 4N = \frac{\alpha^2}{N} \tag{79}$$

The normalized standard error ϵ is then determined by

$$\epsilon^2 = \frac{\text{Var}(\hat{\alpha}^2)}{E^2(\hat{\alpha}^2)} = \frac{\alpha^4}{N}\,\frac{1}{\alpha^4} = \frac{1}{N} \tag{80}$$

or

$$\epsilon = \frac{1}{\sqrt{N}} \tag{81}$$

These formulas could be employed to estimate α^2 directly. However, since the unsquared quantity α appears in Eq. (74), this is the quantity of interest.

For large N, if a variable x has a $\chi^2(2N)$ distribution, then $\sqrt{2x}$ is approximately normal with mean $\sqrt{2N-1}$ and variance of unity. See page 371, Reference 14. The unit variance in this transformation is an approximation but only terms of the form 1/4N are neglected. Hence,

for even moderate values of N the approximation is good. Using these facts one calculates for the mean value:

$$E\left(\hat{\alpha}\,\frac{\sqrt{\pi}}{2}\right) = \frac{\sqrt{\pi}}{2}\,\frac{\alpha}{\sqrt{4N}}\;E\left(\sqrt{2\chi^2(2N)}\right) \approx \frac{\sqrt{\pi}}{2}\,\frac{\alpha}{\sqrt{4N}}\,\sqrt{2N-1} \qquad (82)$$

and the variance:

$$\mathrm{Var}\left(\hat{\alpha}\frac{\sqrt{\pi}}{2}\right) = \frac{\pi}{4}\,\frac{\alpha^2}{4N}\;\mathrm{Var}\left(\sqrt{2\chi^2(2N)}\right) \approx \frac{\pi}{4}\,\frac{\alpha^2}{4N} \qquad (83)$$

The normalized standard error ϵ is then obtained from

$$\epsilon^2 = \frac{\mathrm{Var}\left(\alpha\frac{\sqrt{\pi}}{2}\right)}{E^2\left(\alpha\frac{\sqrt{\pi}}{2}\right)} = \frac{\frac{\pi}{4}\,\frac{\alpha^2}{4N}}{\frac{\pi}{4}\,\frac{\alpha^2}{4N}\,(2N-1)} = \frac{1}{2N-1} \qquad (84)$$

Finally, the normalized standard error of the unknown portion of the mean time to fatigue failure is given by

$$\epsilon = \frac{1}{\sqrt{2N-1}} \qquad (85)$$

One may now use Eq. (85) for experiment planning purposes. Large sample sizes should be expected to properly employ Eq. (85), but for the lack of better available techniques, one may use it as a reasonable guideline for relatively small samples also.

For example, assume it is desired to estimate the unknown portion $\left(\alpha\,\frac{\sqrt{\pi}}{2}\right)$ of mean time to fatigue failure (with γ assumed known) where a normalized standard error of 20% is allowed. Then

$$\sqrt{2N - 1} = \frac{1}{.2}$$

and

$$N = \frac{1}{2}\left[\left(\frac{1}{.2}\right)^2 + 1\right] = 13$$

As one can see from this example, obtaining fairly good estimates of fatigue life will require extensive panel and other structural testing. To work the formula the other way, assume N = 20 panels are tested to failure. Then the normalized standard error is

$$\epsilon_1 = \frac{1}{\sqrt{2N - 1}} = \frac{1}{\sqrt{39}} = .16$$

In these estimates, the minimum time to failure, γ, has been assumed known. Inclusion of this figure will reduce percentage errors. This reduction will be significant if γ is large compared to α.

The main reason for the assumptions of known β and known γ are that considerably more complication enters the estimation procedures when all three parameters are to be estimated from the data. The above discussed procedures provide suitable guidelines for fatigue experiment planning. After a fair amount of data is collected, revised estimates can be obtained for the parameters and more accurate procedures can be used for subsequent experimental program design.

10.7 EXAMPLE OF FATIGUE TEST

The following data are obtained from an actual sonic fatigue test result of a certain panel.

	Time to Failure in 100 sec (T_i)	$\log (T_i - \gamma)$	$\left\{\log (T_i - \gamma)\right\}^2$
	35	2.71	7.34
	37	2.83	8.01
	38	2.89	8.35
	33	2.56	6.55
	39	2.94	8.64
	37	2.83	8.01
	30	2.30	5.29
	29	2.20	4.84
	30	2.30	5.29
	27	1.95	3.80
	39	2.94	8.64
	49	3.37	11.36
	39	2.94	8.64
	23	1.10	1.21
	24	1.39	1.93
	34	2.64	6.97
	26	1.79	3.20
	28	2.08	4.33
	29	2.20	4.84
	30	2.30	5.29
Total	656	48.26	122.53

Table 2. Summary of Test Data

Assume that it is guaranteed or known from long experience that the panel never fails before 2000 sec. Thus, let $\gamma = 20$ and let $x_i = T_i - \gamma$ for each i. Then

$$(\Sigma \log x_i)^2 = 2329.03 \quad \Sigma \log x_i^2 = 122.53$$

Using Eq. (64),

$$\hat{\beta} = \left\{\frac{1}{6}\left[\frac{\pi^2 (19)}{122.53 - 116.45}\right]\right\}^{\frac{1}{2}} = \sqrt{5.14} = 2.27$$

The estimated β value of 2.27 indicates that although the fatigue failure distribution is almost symmetrical (see Figure 11), the distribution has slight positive skew. A plot of the test data also confirms the positive skew and a normal distribution may be a poor model in this example as in many fatigue test analyses. Also, it indicates that the failure rate is a sharply increasing function of time (see Figure 8). Now assuming that $\hat{\beta}$ is the true value α can be estimated by Eq. (59). A simple approximation can be obtained by Eq. (53) as follows.

$$\hat{\mu} = \overline{T} = \gamma + \left(\frac{1}{\hat{\beta}}\right) ! \hat{\alpha}$$

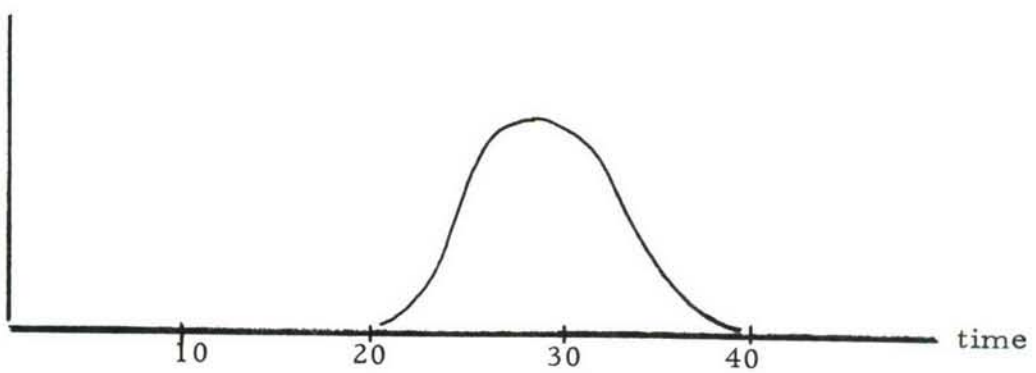


Figure 11. Distribution of Failure Time

In the above example

$$\hat{\mu} = \overline{T} = \frac{\Sigma T_i}{20} = 32.8 = 20 + (.394) ! \hat{\alpha}$$

$$\hat{\alpha} \approx \frac{12.8}{.888} = 14.41$$

Thus, desired density function as a mathematical model is estimated by

$$\hat{w}(t) = 2.27 \frac{(t - 20)^{1.27}}{14.41^{2.27}} e^{-\left[\frac{t-20}{14.41}\right]^{2.27}} \qquad \text{for } t \geq 20$$

$$= 0 \qquad \text{for } t < 20$$

The reliability function is estimated by

$$\hat{R}(t) = e^{-\left[\frac{t-20}{14.41}\right]^{2.27}} \qquad \text{for } t \geq 20$$

$$= 0 \qquad \text{for } t < 20$$

In the following reliability functions based on the Weibull and normal distributions are denoted $R(t)$ and $R_N(t)$ respectively, and the actual percent of sample panels which have not failed at time t is denoted $R_s(t)$. By the definition of reliability,

$$R_N(t) = \int_t^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

$$= \int_{\frac{t-\mu}{\sigma}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \qquad (86)$$

where μ is the mean and σ is the standard deviation of the normal distribution.

In the above example, $\overline{T} = 32.8$ and $s = 6.38$. Thus, $R_N(t)$ is estimated by

$$\hat{R}_N(t) = \int_{(t-32.8)/6.38}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx$$

The actual percent of samples that did not fail at time t is

$$R_s(t) = 1 - \frac{\text{number of failure before } t}{\text{sample size}}$$

The following table compares three estimates.

t	$\hat{R}(t)$	$\hat{R}_N(t)$	$R_s(t)$
20	1.00	.98	1.00
25	.91	.89	.90
27	.82	.82	.85
30	.65	.67	.65
33	.45	.49	.50
35	.34	.36	.40
38	..19	.21	.25
40	.12	.13	.05
45	.03	.03	.05

Table 3. Estimates of Reliabilities

For example, the probability that the panel survives 3000 seconds is .65. In actual sample of 20 tests, 65% of the panels survived 3000 seconds.

In order to describe the failure probability of the panel, the survival curve which is a graph of R(t) is given in Figure 12.

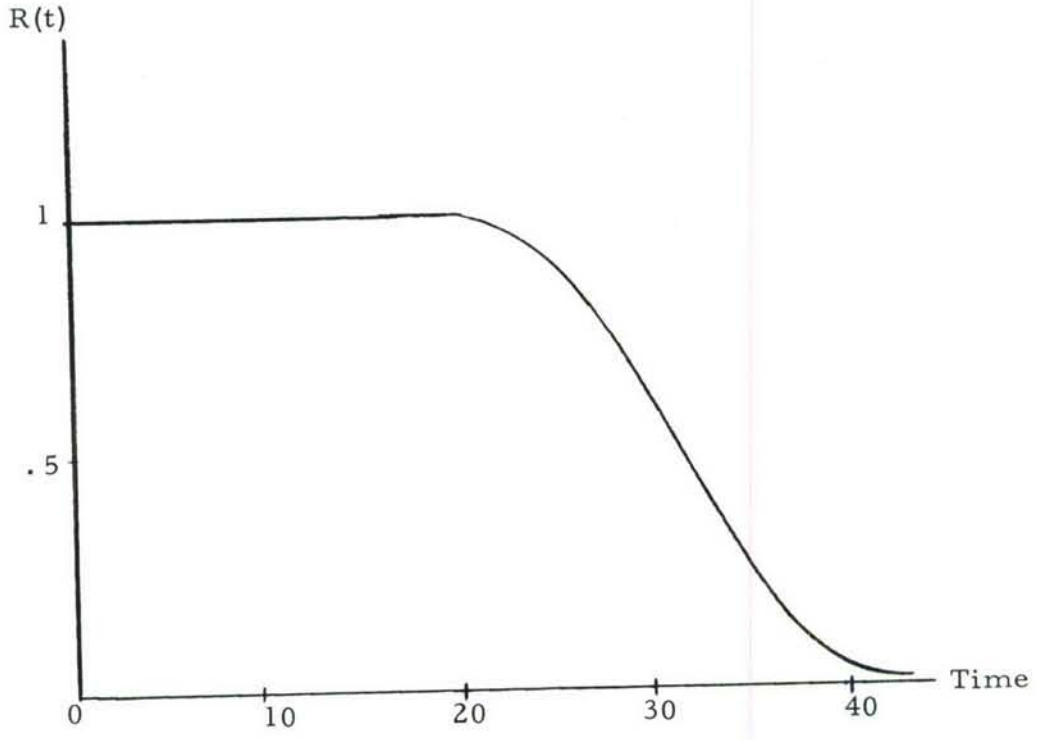


Figure 12. Survival Curve

11. RECOMMENDATIONS

The properties of the Weibull distribution have been discussed. Indications are that the Weibull distribution is suitable as a mathematical model of life length when the three parameters are properly estimated. The best procedure may be to estimate the parameters from the simultaneous equations given in Section 10.4. However, the analytical approach does not yield explicit solutions. Thus, development of a computer program is recommended to provide iterative solutions.

More research is recommended to improve the method of estimation of the Weibull parameters. The properties of estimators such as consistency, efficiency, and asymptotic distribution should be investigated.

Some decision procedures for distinguishing between two sets of life data in terms of reliability have been presented. This is another area which needs further study for more useful applications. The applications of the minimax, Bayes, and other decision procedures may be explored.

Additional studies in some areas of sample size reduction would be fruitful. Clearly, it is laborious to reduce very large sample size by the technique described in Section 3. One has to compute a sequence of means and ranges in the process of eliminating outliers. Then one needs a random number table. The technique is recursive in nature and as a result it is extremely suitable to adopt a computer solution. The above method can be easily and compactly programmed for a computer solution. It is recommended that basic methods of sampling be studied further. Such criteria as relative efficiency, net relative efficiency, costs, bias, and practicality should be evaluated for useful applications.

REFERENCES

1. Dixon, W.J., and F.J. Massey, Introduction to Statistical Analysis, McGraw-Hill Book Company, New York. 1957. p. 442.

2. Cochran, W.G., Sampling Techniques, John Wiley and Sons, Inc. New York. 1963. p. 209.

3. Cramer, Harald, Mathematical Methods of Statistics, Princeton University Press, Princeton, New Jersey. 1946.

4. Blackman, R.B., and J.W. Tukey, The Measurement of Power Spectra, Dover Publications, Inc. New York. 1958.

5. Piersol, A.G., "The Measurement and Interpretation of Ordinary Power Spectra for Vibration Problems," National Aeronautics and Space Administration, Washington, D.C. NASA CR-90. September 1964.

6. Parzen, E., "On Consistent Estimates of the Spectrum of a Stationary Time Series," Annals of Mathematical Statistics, Vol. 28. June 1957. pp. 329-348.

7. Goodman, N.R., "On the Joint Estimation of the Spectra, Co-Spectrum and Quadrature Spectrum of a Two-Dimensional Stationary Gaussian Process," Scientific Paper No. 10, Engineering Statistics Laboratory, New York University. 1957.

8. Goodman, N.R., "Measurement of Matrix Frequency Response Functions and Multiple Coherence Functions," AFFDL-TR-65-56, Research and Technology Division, USAF, Wright-Patterson AFB, Ohio. May 1965.

9. Bendat, J.S., Enochson, L.D., Klein, G.H., and A.G. Piersol. "Advanced Concepts of Stochastic Processes and Statistics for Flight Vehicle Vibration Estimation and Measurement," ASD TDR 62-973, Aeronautical Systems Division, AFSC, USAF, Wright-Patterson AFB, Ohio. December 1962. (AD 297 031).

10. Lehman, E.H., Jr., "Shapes, Moments and Estimators of the Weibull Distribution," IEEE Trans. on Reliability, Vol. R-12, No. 3. September 1963. pp. 32-38.

11. Menon, M.V., "Estimation of the Shape and Scale Parameters of the Weibull Distribution," Technometrics, Vol. 5, 1963. pp. 175-182.

12. ARINC Research Corp., "Reliability Engineering," 1964. pp. 5.27-5.30.

13. Epstein, B., and M. Sobel, "Life Testing," Journal of American Statistical Association, Vol. 48. 1953. p. 486..

14. Kendall, M.G., and A. Stuart, The Advanced Theory of Statistics, Vol. I, Hofner Publishing Co., New York. 1963. pp. 371, 372.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*

Measurement Analysis Corporation
10962 Santa Monica Blvd
Los Angeles, California 90025

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

3. REPORT TITLE

Random Fatigue Test Sampling Requirements

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Final

5. AUTHOR(S) *(Last name, first name, initial)*

Choi, S. C. and Enochson, L. D.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
August 1965	67	14

8a. CONTRACT OR GRANT NO.

AF 33(615)-1314

b. PROJECT NO.

4437

c.

Task 443706

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

AFFDL-TR-65-95

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

MAC 402-02A

10. AVAILABILITY/LIMITATION NOTICES

Qualified requesters may obtain copies of this report from DDC. This report has been furnished to CFSTI for sale to the public.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

Air Force Flight Dynamics Laboratory
Wright-Patterson AFB, Ohio 45433

13. ABSTRACT

Basic considerations are discussed for determining sample sizes and record lengths for various statistical tests and estimates which are important to random fatigue testing. Methods for determining minimum sample sizes when comparing means and variances of normally (Gaussian) distributed random variables are described. Procedures for reducing a relatively large sample to a smaller sample are presented. Elimination of outliers and systematic resampling are two methods given.

An explanation is presented of the requirements and problems involved in the determination of record lengths necessary for an estimate of a given accuracy for autocorrelation functions, ordinary power spectral density functions, cross-correlation functions, cross-spectral density functions, frequency response functions, and probability density functions.

Due to its importance in random fatigue testing applications, the basic properties of the Weibull distribution in terms of its parameters and the failure rate are summarized. A presentation is given of estimation and statistical testing problems related to the Weibull distribution. The best available methods of estimating the parameters are described. Methods of determining sample sizes needed for various analyses are developed. Some problems of reliability analysis applicable in fatigue testing are discussed. New methods of decision techniques for comparing two or more systems are proposed in terms of reliability. The report concludes with an example of the application of the Weibull distribution to actual fatigue test data.

DD FORM 1 JAN 64 1473

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Random Sampling Statistics Data Processing						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2*a*. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS), (S), (C), or (U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.