Best Available Copy

DEVELOPMENT OF PERFORMANCE EVALUATIVE MEASURES

Personnel Psychophysics: Quantification of Malfunction Detection Probability

William Miehle
Arthur I. Siegel

prepared by

Applied Psychological Services
Science Center
Wayne, Pennsylvania

for the

Personnel and Training Branch
Office of Naval Research

under

-C-0107

Contract N00014-67-C0107
NR 153-177/7-5-66

December 1967

ABSTRACT

The logic of a technique for employing technician "confidence that a defect exists" for maximizing the probability of malfunction recognition is described. The technique is based on and drawn from parallel thinking in signal detection theory. Operator characteristic curves are derived for a variety of distributions of "confidence." Continuous and discrete distributions of "confidence" are considered as well as single and double criterion levels. The implications of the work for training and posttraining performance evaluation are pointed out.

ACKNOWLEDGMENTS

This work was completed under Contract N00014-67-C-0107 between the Personnel and Training Branch, Office of Naval Research and Applied Psychological Services. At the Office of Naval Research, a number of persons have contributed to our thinking both in terms of the work here reported and in terms of our prior studies from which the present report is partially drawn. These have included Dr. Glenn L. Bryan, Dr. Victor Fields, and Dr. James Regan.

We acknowledge our indebtedness to these persons who have sparked our curiosity and encouraged our efforts.

William Miehle
Arthur I. Siegel

APPLIED PSYCHOLOGICAL SERVICES
December 1967

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Previous research, completed by Applied Psychological Services, has attempted to set into focus methods for estimating personnel subsystem reliability. These studies, drawn from and based on a series of investigations into "personnel psychophysical" relationships, sought to establish techniques for quantitatively assessing the probability of successfully performing a given avionic maintenance act. In the first of these studies into personnel subsystem reliability (Siegel & Pfeiffer, 1966), the utility of the ratio of effective to effective plus ineffective performances was investigated as the basic ingredient for personnel subsystem reliability estimation. Measures of effective and ineffective performance were obtained for Fleet avionic maintenance personnel through magnitude estimation methods. The measures were anchored to a number of avionic job dimensions isolated through a multidimensional scaling analysis of the avionics maintenance job. When compounded through techniques which are analogous to those employed in traditional electronic reliability prediction, the evoked numerics were found to discriminate between squadrons known to be of different competency by other criteria and to discriminate within squadrons in accordance with reasonable expectation. Siegel and Pfeiffer concluded that:

> The obtained avionic personnel subsystem (reliability) indices seem to be useful for post-training performance appraisal, personnel placement, and squadron evaluative purposes.

Because of the reasonableness of the results, the ease of employment, and the general utility of the approach employed by Siegel and Pfeiffer, Siegel and Miehle (1967) extended the work to consider a number of circumstances not originally considered. Additionally, a measure of "effectiveness," as separate from "reliability," was derived.

The effectiveness calculation considered not only the reliability of the individual technicians performing the maintenance acts, but also elapsed time, amount of manpower employed, and job activity repetition. It was contended that the technique is useful for:

1. quantitative comparison of the effectiveness of different teams or individuals who perform the same task
2. prediction of the performance effectiveness of a team or of an individual on a task
3. derivation of training requirements
4. optimization of personnel assignments and maintenance procedures
5. evaluating the design of new systems or alternative designs of the same system

Additionally, Siegel and Miehle developed nomographs and tables to simplify the computational aspects of the procedures they described.

Purpose of Present Report

The previous studies consider only peripherally the trouble-shooting aspect of avionic maintenance. Within the technique described, the job activity, "electro cognition," is considered to encompass a host of mental acts involved in the maintenance and trouble-shooting of avionic equipment. Yet, it is known that malfunction recognition represents a particularly troublesome aspect of the maintenance procedure, whether periodic inspections or malfunction diagnosis and location are involved. Good components are often erroneously replaced and marginal or defective components overlooked.

As the logical extension of the previous work in "personnel psychophysics," it seemed reasonable to apply the psychophysical developments of signal recognition theory to the malfunction recognition issue. The central problem then becomes that of demonstrating a quantitative method for determining the probability that an avionic technician will say that: (1) a defect is present when a defect is in fact present, and (2) a defect is present when in fact no defect is present.

The methods to be described in subsequent chapters of this report are analogous to those used by Swets and his coworkers (cf., 1964). In the present report, the confidence that a defect exists replaces the perceived intensity of a signal in their work.

CHAPTER II

METHODS

In certain avionic maintenance activities, satisfactory performance may be said to occur when action is taken when it should be taken, and no action is taken when no action is warranted. Examples of this occur during routine equipment inspection, fault location and isolation, and malfunction correction. Let us consider all of these under the generic classification of defect recognition.

Defect Recognition

Let x be a measure of the maintenance technician's confidence that a defect exists, with $0 \le x \le 1$. For the situation in which the technician has complete confidence that a defect exists, $x = 1$. When the technician has complete confidence that no defect exists, $x = 0$. Intermediate confidence values may exist because of incomplete information, conflicting information, inadequate knowledge about the equipment system, etc.

Due to this uncertainty, a technician may have different values of confidence on different occasions, even though the situations presented to him are the same. This variation is representable by a confidence distribution.

When there is a defect, confidence that a defect exists is high (x is near 1) most of the time. When there is actually no defect, in most cases x is low.

Continuous Distribution of Confidence

For a continuous distribution of x (the measure of confidence that a defect exists), let f(x) be the frequency or probability density function. When a defect exists, this function is monotonically increasing, having a maximum value for $x = 1$ (Figure 1). When a defect does not exist, we have a different frequency function which is monotonically decreasing. It has a maximum for $x = 0$ (Figure 2).

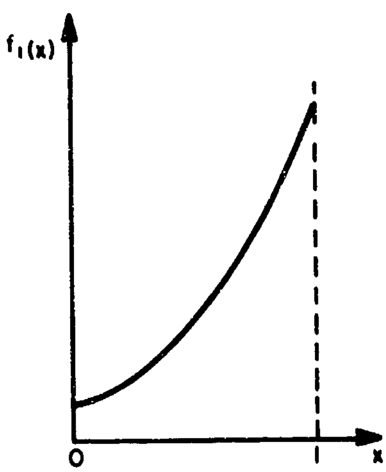

Figure 1 Hypothetical frequency function when defect exists

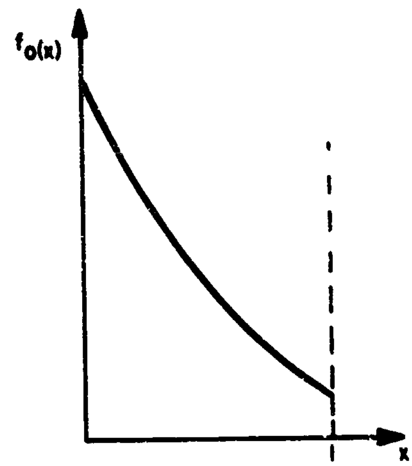

Figure 2 Hypothetical frequency function when there is no defect

Note that $\int_{-\infty}^{\infty} f(x)dx = 1$, i.e., the area under the curve is unity. Since $0 \le x \le 1$, $\int_0^1 f(x)\,dx = 1$ in our case.

The probability that the measure of confidence is less than or equal to a constant, k, is: $Pr[x \le k]$.

$$Pr[x \le k] = \int_{-\infty}^{k} f(x)dx$$

This probability is the area under the curve of f(x) to the left of $x = k$. The distribution function is $F(x) = \int_{-\infty}^{x} f(u)\,du$ and is the area under the curve of f(x) to the left of x. The probability that $a \le x \le b$ is $P_r[a \le x \le b] = \int_a^b f(x)\,dx = F(b) - F(a)$.

Note that $F'(x) = f(x)$. Suppose that a technician will report a defect (will say "yes") when his confidence that there is a defect is greater than or equal to some subjective criterion level, c, i.e., $x \ge c$. Then he will also report that a defect is not present (will say "no") when $x < c$.

For either a defect or no defect, the probability that a technician will say "yes" is $P_r[x \ge c] = \int_c^1 f(x)\,dx$. When there actually is a defect, then $P_r[x \ge c]$ is the probability of a correct response; call this P_r [yes|D]. When there is no defect, then $P_r[x \ge c]$ is the probability of an incorrect response (false alarm); call this P_r [yes|D'].

Probability of Defect Recognition

For either a defect or no defect, the probability that a technician will say "no" is $P_r[x < c] = \int_0^c f(x)\,dx$. When there actually is a defect, then $P_r[x < c]$ is the probability of an incorrect response (failure to detect). Call this P_r[no|D]. When there is no defect, then $P_r[x < c]$ is the probability of correct response. Call this P_r [no|D'].

Let the frequency function when defect exists be f_1 and for no defect be f_0. Then: $P_r[\text{yes}|D] = \int_c^1 f_1(x)\,dx$, $P_r[\text{yes}|D'] = \int_c^1 f_0(x)\,dx$, $P_r[\text{no}|D] = \int_0^c f_1(x)\,dx$, and $P_r[\text{no}|D'] = \int_0^c f_0(x)\,dx$.

The probability that the subject says "yes" and there is a defect is: $P_r[\text{yes} \wedge D] = P_r[\text{yes}|D]P_r[D]$. The probability that the answer is "no" and there is no defect is $P_r[\text{no} \wedge D'] = P_r[\text{no}|D']\,P_r[D']$.

The probability of a correct response (% of correct responses), R, is:

$$R = P_r[\text{yes} \wedge D] + P_r[\text{no} \wedge D']$$
$$= P_r[\text{yes} \mid D] P_r[D] + P_r[\text{no} \mid D'] P_r[D']$$
$$= P_r[D] \int_c^1 f_1(x)\,dx + P_r[D'] \int_0^c f_0(x)\,dx. \qquad (1)$$

The probability of an incorrect response is:

$$P_r[\text{yes} \wedge D'] + P_r[\text{no} \wedge D]$$
$$= P_r[\text{yes} \mid D'] P_r[D'] + P_r[\text{no} \mid D] P_r[D]$$
$$= P_r[D] \int_0^c f_1(x)\,dx + P_r[D'] \int_c^1 f_0(x)\,dx. \qquad (2)$$

The sum of these probabilities (correct and incorrect response) is unity. These relationships can be alternatively represented as indicated in Exhibit I.

	D	D'
Yes	$P_r[\text{yes} \wedge D]$	$P_r[\text{yes} \wedge D']$
No	$P_r[\text{no} \wedge D]$	$P_r[\text{no} \wedge D']$

Exhibit I

Exhibit I can be considered as a Venn diagram in which the large rectangle represents the universal set of all logical possibilities. The column labeled D represents the set of situations in which a defect exists. The row labeled "Yes" represents the set situations in which "yes" was the response. The cells represent intersections of the sets represented by rows and columns. Their measures are the joint probabilities, whereas the measures of the rows and columns are the marginal probabilities. Note that only if independence exists does $P_r[yes \wedge D] = P_r[yes] P_r[D]$. This situation is a special case of the formulas given earlier.

Reliability and Optimum Confidence Criterion Level

Reliability R (or the probability of correct response) is a function of $P_r[D]$ and the confidence distribution:

$$R = P_r[D] \int_c^1 f_1(x)\,dx + P_r[D'] \int_0^c f_0(x)\,dx$$

For maximum R, we set $\frac{dR}{dc} = 0$ to find the optimum value of c. Assume $P_r[D]$ is constant. The implicit equation:

$$\frac{dR}{dc} = -P_r[D]\, f_1(c) + P_r[D']\, f_0(c) = 0$$

gives optimum value(s) of c.

$$\frac{f_1(c)}{f_0(c)} = \frac{Pr[D']}{Pr[D]} = \frac{1 - Pr[D]}{Pr[D]}$$

Suppose that, in the long run, for an unchanging situation, the fraction of times a defect occurs is r. Then $P_r[D] \approx r$ and $P_r[D'] \approx 1 - r$, so that $\frac{f_1(c)}{f_0(c)} \approx \frac{1 - r}{r}$. For example, if a defect occurs 20% of the time $\frac{1 - r}{r} = \frac{1 - .2}{.2} = \frac{.8}{.2} = 4$. Then, for maximum reliability, the technician should (and probably will) make c such that $\frac{f_1(c)}{f_0(c)} = 4$. As example 1, assume the graphs of f_0 and f_1 as shown in Figure 3.

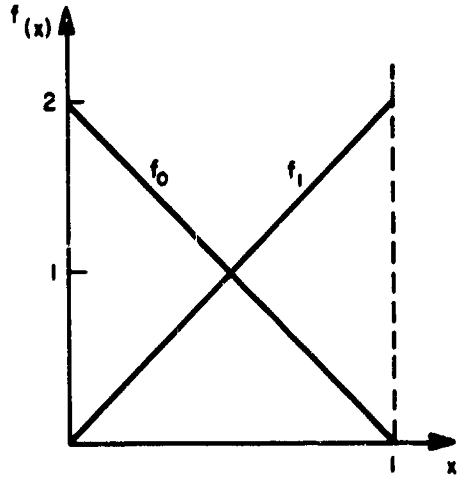

Figure 3 Linear frequency functions, with (f_1) and without (f_0) defect (first example)

Note that the vertical scale has been adjusted so that the area under each curve is equal to 1. Here, $f_0(x) = 2 - 2x$, and $f_1(x) = 2x$, so that $\frac{f_1(c)}{f_0(c)} = \frac{2c}{2-2c}$. For maximum reliability, for Pr[D] = .2, $\frac{2c}{2-2c} = 4$, or c = .8.

In this case, the technician should report a defect when his confidence value is .8 or above. For any Pr[D] = r,

$$\frac{2c}{2-2c} = \frac{1-r}{r}.$$

Solving for c gives: c = 1 - r. When only a few defects exist, r is low and consequently c is large, i.e., the subject will or should say "yes" only when he is very sure that a defect exists.

As another example, consider Figure 4. The area under each curve is still 1.

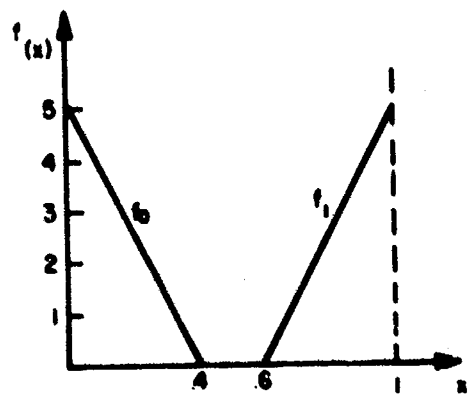

Figure 4 Linear frequency functions, with (f_1) and without (f_0) defect (second example)

$$f_0(x) = 5-12.5x \text{ and } f_1(x) = 12.5x - 7.5$$

On the basis of the distribution shown, when a defect exists, the technician always has a confidence level of .6 or greater. When no defect exists, the technician always has a confidence of .4 or less. Therefore, if his threshold or critical value c is between .4 and .6, he will always be correct (R = 1). This should be independent of Pr[D]. Let us compute R for $.4 \le c \le .6$.

$$R = Pr[D]\int_c^1 f_1(x)\,dx + P_r[D']\int_0^c f_0(x)\,dx$$

$$= Pr[D](1) + P_r[D'](1) = 1$$

If $c < .4$, the technician will sometimes report a defect when there is none. If $c > .6$, the subject will sometimes fail to notice a defect.

Consider the equation for optimum c:

$$\frac{f_1(c)}{f_0(c)} = \frac{1-r}{r} \quad \text{or} \quad \frac{12.5c - 7.5}{5 - 12.5c} = \frac{1-r}{r}.$$

When solved for c, this gives:

$$C = .4 + .2r.$$

As r varies from 0 to 1, c varies from .4 to .6. In this example, it turns out that all values of c in that interval are optimum.

Let us now calculate the reliability, R, for the example in which the curves intersected (first example):

$$R = r\int_c^1 2x\,dx + (1-r)\int_0^c (2-2x)\,dx$$

$$= r[x^2]_c^1 + (1-r)[2x - x^2]_0^c$$

$$= r(1-c^2) + (1-r)[2c - c^2 - 0]$$

$$= r + 2c - c^2 - 2rc.$$

For optimum defect recognition ($c = 1 - r$) this becomes:

$$R_m = r^2 - r + 1$$

R_m is the maximum value of reliability for the given confidence frequency functions and a given r. See Figure 5. The graph is a parabola.

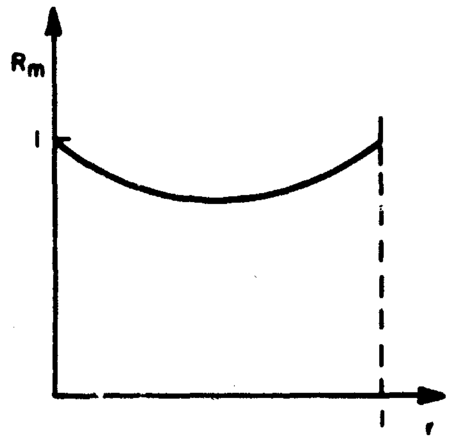

Figure 5 Maximum reliability vs. relative frequency of defect for function in Figure 3

For $r = .2$ (optimum $c = .8$),

$$R_m = .04 - .2 + 1 = .84$$

The minimum value of R_m (.75) occurs when $r = \frac{1}{2}$. When $r = 0$ or 1, $R_m = 1$. This relationship is, in a sense, in accord with information theory. Uncertainty is highest for equiprobable events as when $Pr[D] = Pr[D']$ (or $r = \frac{1}{2}$). Then, it is reasonable to expect least reliability R for $r = \frac{1}{2}$. For $R = .5$, the information is $H = \Sigma p_i \log_2 \frac{1}{p_i} = \frac{1}{2}\log 2 + \frac{1}{2}\log 2 = \frac{1}{2} + \frac{1}{2} = 1$ bit. For $r = .2$, the information is $.2 \log \frac{1}{.2} + .8 \log \frac{1}{.8} = .46 + .26 = .72$ bits. No attempt is made at the present to relate information to reliability in a quantitative manner.

Operating Characteristic Curve

In signal detection theory, receiver operating characteristic (ROC) curves are drawn to show the relationships between $Pr[yes|S]$ and $Pr[yes|S']$ where S stands for a signal being present. An analogous type of operating characteristic curve can be drawn for the present data. For the first example:

$$\Pr[\text{yes}\,|\,D'] = \int_c^1 f_0(x)\,dx = \int_c^1 (2 - 2x)dx = 2x - x^2 \Big|_c^1$$
$$= 2 - 1 - (2c - c^2) = 1 - 2c + c^2 = (1 - c)^2$$
$$\Pr[\text{no}\,|\,D] = \int_0^c f_1(x)dx = \int_0^c 2x\,dx = x^2 \Big|_0^c = c^2$$

We have already found that:

$$\Pr[\text{yes}\,|\,D] = \int_c^1 f_1(x)dx = 1 - c^2 \quad \text{and}$$
$$\Pr[\text{no}\,|\,D'] = \int_0^c f_0(x)dx = 2c - c^2 \,.$$

Note that $\Pr[\text{yes}\,|\,D] + \Pr[\text{no}\,|\,D] = (1 - c^2) + c^2 = 1$ and

$$\Pr[\text{yes}\,|\,D'] + \Pr[\text{no}\,|\,D'] = (1 - c)^2 + (2c - c^2) = 1.$$

Using corresponding values of Pr[yes | D] and Pr[yes | D'] for various values of c, the operating characteristic curve can be obtained. Table 1 gives such values.

Table 1

Conditional Probabilities for Various Values of Criterion Level

c	Pr[yes | D'] = $(1 - c)^2$	Pr[yes | D] = $1 - c^2$
0	1	1
.2	.64	.96
.4	.36	.84
.6	.16	.64
.8	.04	.36
1.0	0	0

A plot of the data of Table 1 is presented as Figure 6 and in Figure 11 as a = 1. (The similarity in the form of this curve and that of the receiver operating characteristic curves of signal detection theory is self evident. We have deliberately called our curves operating characteristic curves to accentuate the similarity.)

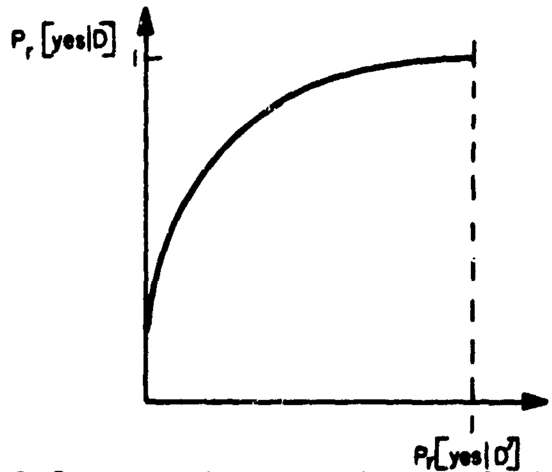

Figure 6 Operating characteristic curve for Example 1

Such a curve (Figure 6) is interpretable in several ways. First, if a technician's performance falls below the points of the curve, he is not performing malfunction diagnosis correctly. If his performance falls along the transverse line shown, he is performing at the chance level. Here, either training or equipment modification (automatic test equipment) might be indicated. Finally, for a fixed value of c, the ordinate of a point on the curve gives the probability of a malfunction being detected, given that a malfunction is present; the abcissa gives the corresponding probability that a malfunction will be reported when none exists, in fact.

Exponential Confidence Distribution

Let us assume that technician confidence possesses an exponential distribution. Exponential confidence frequency curves are shown in Figure 7.

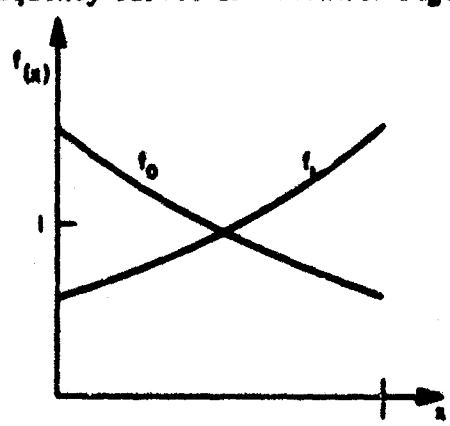

Figure 7 Frequency functions, with (f_1) and without (f_0) defect

The function for no defect is of the form $f_0(x) = ke^{-x}$. The area under the curve must be unity and accordingly:

$$\int_0^1 ke^{-x}dx = 1$$

$$-ke^{-x}\Big|_0^1 = -k(e^{-1} - 1) = 1$$

$$k = \frac{e}{e-1} = 1.582$$

$$f_0(x) = \frac{e}{e-1}e^{-x} = \frac{e^{1-x}}{e-1} = .582e^{1-x}.$$

The function when there is a defect is:

$$f_1(x) = \frac{e^x}{e-1} = .582e^x.$$

The various conditional probabilities are:

$$\Pr[\text{yes}|D] = \int_c^1 \frac{e^x}{e-1}dx = \frac{1}{e-1}e^x\Big|_c^1 = \frac{1}{e-1}(e-e^c) = \frac{e-e^c}{e-1} = \frac{1-e^{c-1}}{1-e^{-1}}$$

$$\Pr[\text{yes}|D'] = \int_c^1 \frac{e^{1-x}}{e-1}dx = -\frac{1}{e-1}e^{1-x}\Big|_c^1 = -\frac{e^0-e^{1-c}}{e-1} = \frac{e^{1-c}-1}{e-1}$$

$$\Pr[\text{no}|D] = 1 - \Pr[\text{yes}|T] = \int_0^c \frac{e^x}{e-1}dx = \frac{e^x}{e-1}\Big|_0^c = \frac{e^c-1}{e-1}$$

$$\Pr[\text{no}|D'] = 1 - \Pr[\text{yes}|T'] = \int_0^c \frac{e^{1-x}}{e-1}dx = -\frac{e^{1-x}}{e-1}\Big|_0^c = \frac{1-e^{-c}}{1-e^{-1}}$$

and

$$R = \Pr[D]\Pr[\text{yes}|D] + \Pr[D']\Pr[\text{no}|D']$$

$$= r\left(\frac{1-e^{c-1}}{1-e^{-1}}\right) + (1-r)\left(\frac{1-e^{-c}}{1-e^{-1}}\right).$$

Optimum Confidence Level for Exponential Cases

If one wishes to determine the confidence level that the technician should adopt in order to obtain a maximum value of R, the technician should adjust his c value so that:

$$\frac{f_1(c)}{f_0(c)} = \frac{1-r}{r}$$

$$\frac{e^c}{e-1} / \left(\frac{e^{1-c}}{e-1}\right) = \frac{e^c}{e^{1-c}} = e^{2c-1} = \frac{1-r}{r}$$

$$2c-1 = \ln\left(\frac{1-r}{r}\right)$$

$$2c = 1 + \ln\left(\frac{1-r}{r}\right)$$

$$c = \tfrac{1}{2} + \tfrac{1}{2}\ln\left(\frac{1-r}{r}\right).$$

Since $0 < c \leq 1$, the above formula for c holds only when:

$$\frac{1}{e+1} \leq r \leq \frac{e}{e+1} \quad \text{or} \quad .269 \leq r \leq .731.$$

When $r = \frac{1}{e+1} = .269$, $c = 1$, and when $r = \frac{e}{e+1} = .731$, $c = 0$ for maximum R. For these values of r and c, $R_m = \frac{e}{e+1} = .731$. When $r < \frac{1}{e+1}$, c remains 1 and R_m becomes $R_m = 1-r$. When $r > \frac{e}{e+1}$, c remains 0 and R becomes $R_m = r$.

The plot of R_m versus r is shown in Figure 8, where the curve from $x = \frac{1}{e+1}$ to $x = \frac{e}{e+1}$ is given by $R_m = \frac{e}{e+1}\left[1 - 2e^{-\frac{1}{2}}\sqrt{r(1-r)}\right]$. At $r = .5$, $R_m = .623$.

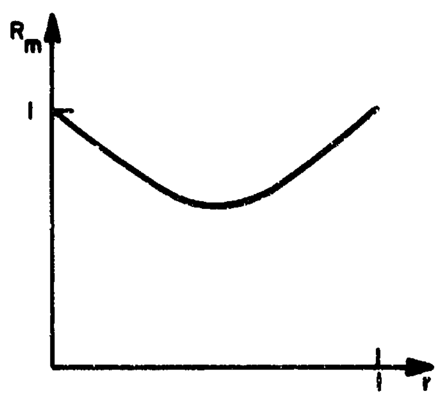

Figure 8 Maximum reliability curve for exponential frequency functions

Table 2 gives values of Pr[yes | D] and Pr[yes | D'] for drawing the operating characteristic curve.

Table 2

Conditional Probabilities for Various Values of Criterion Level for Exponential Frequency Function

c	Pr[yes | D']	Pr[yes | D]
0	1.000	1.000
.2	.713	.872
.4	.478	.713
.6	.286	.522
.8	.129	.286
1.0	.0	0

The plot of Table 2 is shown in Figure 11 as "exp."

Normal Confidence Distribution

The case in which confidence is normally distributed may be similarly developed. Normal confidence frequency curves ($\sigma = 1$) are shown in Figure 9.

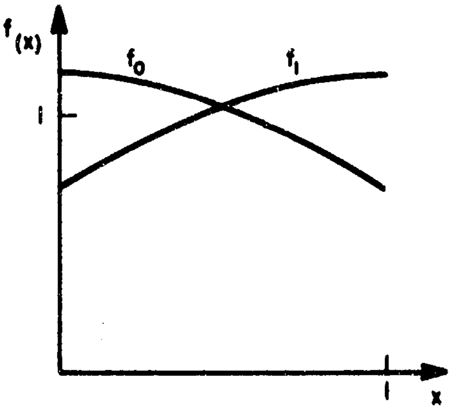

Figure 9 Normal frequency function, with (f_1) and without (f_0) defect

When no defect exists:

$$f_0(x) = \frac{k}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$\frac{k}{\sqrt{2\pi}} \int_0^1 e^{-\frac{x^2}{2}} dx = 1$$

$$k(.34134) = 1$$

$$k = \frac{1}{.34134} = 2.9296$$

$$f_0(x) = 2.9296 \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$$

$$f_1(x) = 2.9296 \frac{e^{-\frac{(x-1)^2}{2}}}{\sqrt{2\pi}}$$

$$Pr[yes|D] = 2.9296 \int_c^1 \frac{e^{-\frac{(x-1)^2}{2}}}{\sqrt{2\pi}} dx$$

let $u = x-1$
$du = dx$

when $x = c$, $u = c-1$
$x = 1$, $u = 0$

$$Pr[yes|D] = 2.9296 \int_{c-1}^{0} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du = 2.9296 \int_0^{1-c} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du$$

$$= 2.9296 \int_0^{1-c} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$$

$$Pr[no|D] = 2.9296 \int_{1-c}^{1} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx$$

$$Pr[yes|D'] = 2.9296 \int_c^1 \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx$$

$$Pr[no|D'] = 2.9296 \int_0^c \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx$$

The reliability is:

$$R = 2.9296 \left[r \int_0^{1-c} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx + (1-r) \int_0^{c} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx \right]$$

Optimum Confidence Level for Normal Distribution Case

For a relative maximum of R, the technician should set his c value so that:

$$\frac{f_1(c)}{f_0(c)} = \frac{1-r}{r}$$

$$e^{-\frac{(c-1)^2}{2}} / e^{-\frac{c^2}{2}} = \frac{1-r}{r}$$

$$e^{c-\frac{1}{2}} = \frac{1-r}{r}$$

$$c - \tfrac{1}{2} = \ln\left(\frac{1-r}{r}\right)$$

$$c = \tfrac{1}{2} + \ln\left(\frac{1-r}{r}\right) .$$

Since $0 \le c \le 1$, this formula holds only when:

$$\frac{1}{1+e^{+\frac{1}{2}}} \le r \le \frac{1}{1+e^{-\frac{1}{2}}} \quad \text{or} \quad .3775 \le r \le .6225.$$

When $r = .3775$, $c = 1$, and when $r = .6225$, $c = 0$ for maximum R. For these values of r and c, $R_m = .6225$. When $r < .3775$, c remains 1 and R_m becomes $R_m = 1-r$. When $r > .6225$, c remains 0 and R_m becomes $R_m = r$. Table 3 gives R_m for values of r between .3775 and .6225.

Table 3

Optimum Criterion Level and Maximum Reliability vs. Relative Frequency of Defects

r	c	R_m
.3775	1	.6225
.4	.9055	.602
.5	.5	.561
.6	.0945	.602
.6225	0	.6225

The plot of Table 3 is shown as Figure 10.

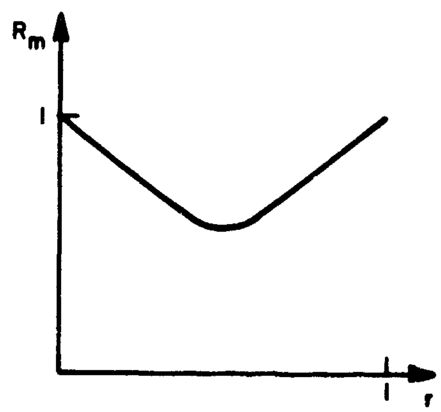

Figure 10 Maximum reliability curve for "normal" frequency curve ($\sigma = 1$)

Table 4 gives values of Pr(yes | D') and Pr(yes | D) for various values of c.

Table 4

Conditional Probabilities for Various Values of Criterion Level for Normal Distribution

c	Pr[yes | D']	Pr[yes | D]
0	1.000	1.000
.2	.767	.845
.4	.544	.662
.6	.338	.455
.8	.156	.232
1	0	0

The graph is shown in Figure 11 as $\sigma = 1$.

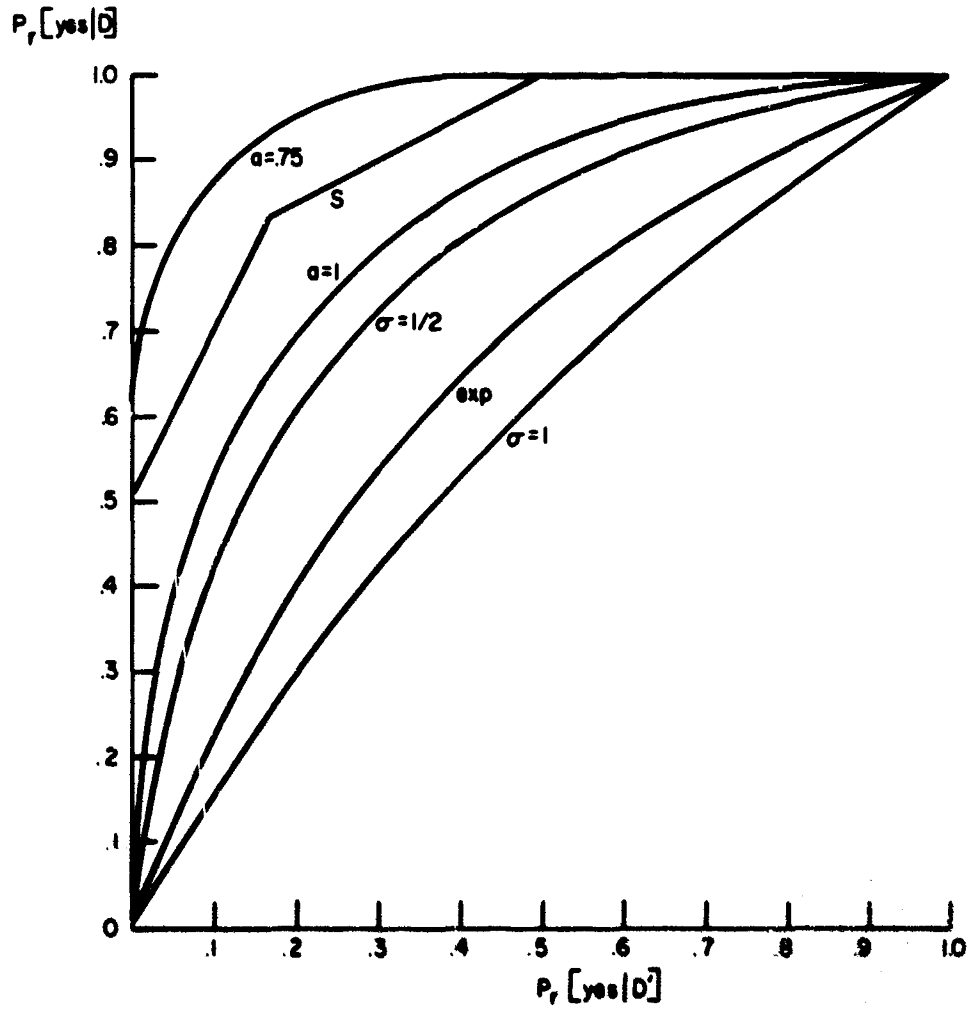

Figure 11 Operating characteristic curves for various confidence distributions

Normal Distribution with any Value of σ

The formula for the normal curve with a standard deviation of any value of σ is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} .$$

When no defect is present:

$$f_0(x) = \frac{k}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \qquad \text{and}$$

$$k\int_0^1 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx = 1$$

$$\text{let } u = \frac{x}{\sigma}, \quad du = \frac{dx}{\sigma}$$

$$\text{for } x = 0,\ u = 0; \text{ for } x = 1,\ u = \frac{1}{\sigma}, \text{ so}$$

$$k\int_0^{\frac{1}{\sigma}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{u^2}{2}} \sigma du = 1$$

$$k\int_0^{\frac{1}{\sigma}} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du = 1.$$

Table 5 gives k for various values of σ. As σ approaches 0, k approaches 2.

Table 5

Normalizing Constant (k) for Various Values of Standard Deviation

σ	k
2	5.222
1	2.930
1/2	2.095
1/3	2.005
1/4	2.000

When a defect is present,

$$f_1(x) = \frac{k}{\sigma\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2\sigma^2}} .$$

Normal Distribution with $\sigma = .5$

As an example of a value of σ different from unity, take $\sigma = .5$.

$$f_0(x) = 4.190\ e^{-2x^2}$$

$$f_1(x) = 4.190\ e^{-2(x-1)^2}$$

The confidence frequency curves are shown in Figure 12.

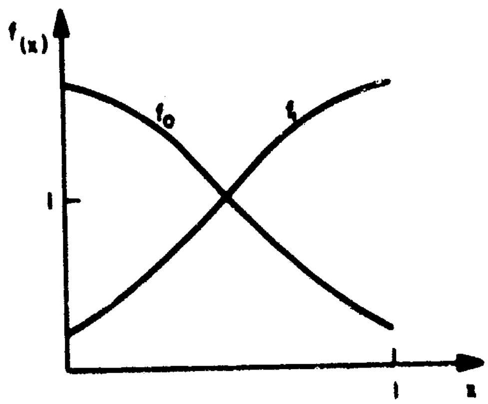

Figure 12 Normal frequency function ($\sigma = .5$), with (f_1) and without (f_0) defect

This represents a considerable improvement over $\sigma = 1$ shown in Figure 9.

$$Pr[yes|D] = k\int_0^{\frac{1-c}{\sigma}} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du$$

$$Pr[no|D] = k\int_{\frac{1-c}{\sigma}}^{\frac{1}{\sigma}} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du$$

$$\Pr[\text{yes}\,|\,D'] = k\int_{\frac{c}{\sigma}}^{\frac{1}{\sigma}} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\,du$$

$$\Pr[\text{no}\,|\,D'] = k\int_{0}^{\frac{c}{\sigma}} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\,du$$

$$R = k\left[r\int_{0}^{\frac{1-c}{\sigma}} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\,du + (1-r)\int_{0}^{\frac{c}{\sigma}} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\,du \right] .$$

Optimum Confidence Level for Normal Distribution with $\sigma = \frac{1}{2}$ Case

In the present case, for a relative maximum of R, the technician should set c such that:

$$\frac{f_1(c)}{f_0(c)} = \frac{1-r}{r}$$

$$\frac{e^{-\frac{(c-1)^2}{2\sigma^2}}}{e^{-\frac{c^2}{2\sigma^2}}} = \frac{1-r}{r} .$$

This reduces to:

$$e^{\frac{1}{\sigma^2}(c-\frac{1}{2})} = \frac{1-r}{r}$$

$$\frac{1}{\sigma^2}(c-\tfrac{1}{2}) = \ln\left(\frac{1-r}{r}\right)$$

$$c-\tfrac{1}{2} = \sigma^2 \ln\left(\frac{1-r}{r}\right)$$

$$c = \tfrac{1}{2} + \sigma^2 \ln\left(\frac{1-r}{r}\right) .$$

Since $0 \le c \le 1$, this formula holds only when:

$$\frac{1}{1+e^{\frac{1}{2\sigma^2}}} \le r \le \frac{1}{1+e^{-\frac{1}{2\sigma^2}}} .$$

When $\sigma = \frac{1}{2}$, this becomes $.119 < r < .881$. For $r \leq .119$, $c = 1$, and for $r \geq .881$, $c = 0$. For $r = .119$, and $r = .881$, $R_m = .881$. For $< .119$ ($c = 1$), $R_m = 1 - r$. For $r > .881$ ($c = 0$), $R_m = r$.

Table 6 gives R_m for values of r between .119 and .881.

Table 6

Optimum Criterion Level and Maximum Reliability vs. Relative Frequency of Defects (Normal Distribution with $\sigma = .5$)

r	c	R_m
.119	1	.881
.2	.846	.813
.3	.712	.757
.4	.601	.725
.5	.500	.715
.6	.399	.725
.7	.288	.757
.8	.154	.813
.881	0	.881

The graph is shown in Figure 13.

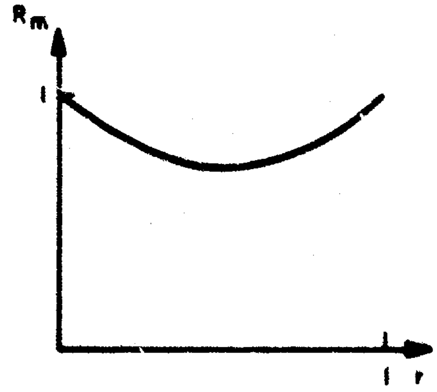

Figure 13 Maximum reliability curve for "normal" frequency function ($\sigma = .5$)

Table 7 gives values of Pr(yes|D') and Pr(yes|D) for various values of c.

Table 7

Conditional Probabilities for Various Values of Criterion Level
(Normal Frequency Function)

c	Pr[yes|D']	Pr[yes|D]
0	1.000	1.000
.2	.674	.933
.4	.396	.806
.6	.194	.604
.8	.067	.326
1.0	0	0

These are plotted in Figure 11 as $\sigma = .5$.

General Straight Line Confidence Distribution

Let us now consider the situation in which a more general straight line frequency function of confidence exists. This is shown in Figure 14.

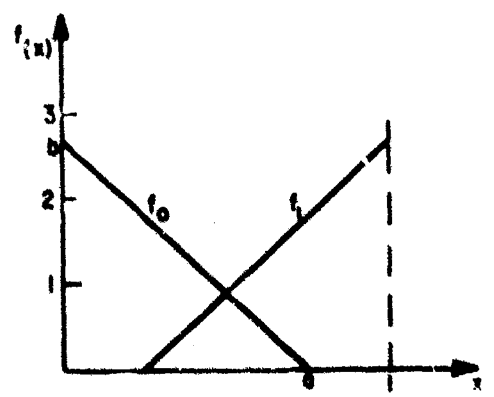

Figure 14 Linear frequency function, with (f_1) and without (f_0) defect

The x and y intercepts of $y = f_0(x)$ occur at a and b respectively, where a is arbitrary and $b = \frac{2}{a}$.

$$f_0(x) = \frac{2}{a}(1 - \frac{x}{a})$$

$$f_1(x) = \frac{2}{a^2}x + \frac{2}{a}(1 - \frac{1}{a})$$

$$Pr[yes\,|D] = 2(\frac{1-c}{a}) - (\frac{1-c}{a})^2$$

$$Pr[no\,|D] = 1 - 2(\frac{1-c}{a}) + (\frac{1-c}{a})^2$$

$$Pr[yes\,|D'] = (1 - \frac{c}{a})^2$$

$$Pr[no\,|D'] = \frac{2c}{a} - \frac{c^2}{a^2}$$

$$R = r\left[2(\frac{1-c}{a}) - (\frac{1-c}{a})^2\right] + (1-r)\left[\frac{2c}{a} - \frac{c^2}{a^2}\right] .$$

For a given value of r, the maximum R occurs when:

$$c = a + (1-2a)r \quad \text{and is given by:}$$

$$R_m = \frac{1}{a^2}\Big\{2ra[(1-a) - (1-2a)r] - r[(1-a) - (1-2a)r]^2$$
$$+ (1-r)[a + (1-2a)r][a - (1-2a)r]\Big\} .$$

As an example, let a = .75:

$$f_0(x) = \frac{8}{3}(1 - \frac{4}{3}x)$$

$$f_1(x) = \frac{32}{9}x - \frac{8}{9} .$$

The optimum c level exists when:

$$c = \frac{3}{4} - \frac{1}{2}r .$$

Values of R_m for various values of r are given in Table 8. The graph is shown in Figure 15.

Table 8

Maximum Reliability for Various Values of Relative Frequency of Defects

r	R_m
0	1
.1	.96
.3	.907
.5	.889
.7	.907
.9	.96
1	1

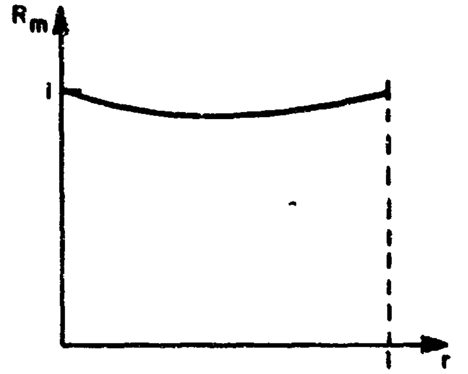

Figure 15 Maximum reliability curve for linear frequency function (a = .75)

Table 9 gives values of Pr[yes |D'] and Pr[yes |D] for various values of c.

Table 9

Conditional Probabilities for Various Values of Criterion Level (Linear Frequency Function)

c	Pr[yes |D']	Pr[yes |D]	c	Pr[yes |D']	Pr[yes |D]
0	1	1	.6	.04	.783
.1	.751	1	.7	.0045	.640
.2	.537	1	.75	0	.556
.25	.445	1	.8	0	.462
.3	.36	.996	.9	0	.250
.4	.218	.960	1.0	0	0
.5	.111	.889			

These are plotted in Figure 11 as a = .75.

Step Function Distribution

Step function confidence frequency curves as shown in Figure 16.

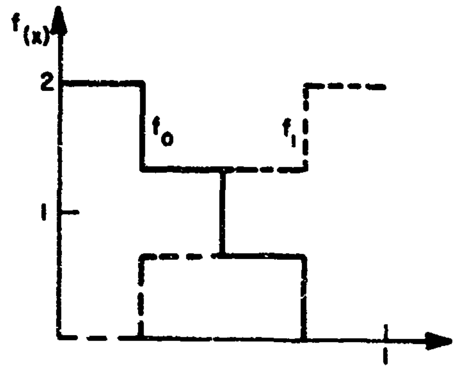

Figure 16 Step function frequency curves

Let us assume that the steps are of equal width and the vertical increments are equal. Let $S_\delta(x) = \begin{cases} 0 \text{ for } x < \delta \\ 1 \text{ for } x \geq \delta \end{cases}$. Its graph is shown in Figure 17.

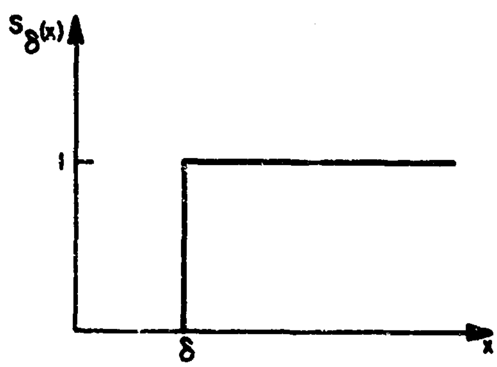

Figure 17 Step function curve

It is possible to represent $f_1(x)$ as a sum of step functions,

$$f_1(x) = k[S_{\frac{1}{4}}(x) + S_{\frac{1}{2}}(x) + S_{\frac{3}{4}}(x)]$$

The constant k must be such that the area under the curve is unity.

$$A = k\int_0^1 f_1(x)dx = k\int_0^1[(S_{\frac{1}{4}}(x) + S_{\frac{1}{2}}(x) + S_{\frac{3}{4}}(x)]dx$$

$$= k[.75 + .5 + .25] = 1.5k$$

$1.5k = 1$

$k = \frac{2}{3}$ so:

$$f_1(x) = \frac{2}{3}[S_{\frac{1}{4}}(x) + S_{\frac{1}{2}}(x) + S_{\frac{3}{4}}(x)]$$

$$f_0(x) = 2 - f_0(x) = 2 - \frac{2}{3}[S_{\frac{1}{4}}(x) + S_{\frac{1}{2}}(x) + S_{\frac{3}{4}}(x)] .$$

Let:

$$I_\delta(a, b) = \int_b^a S_\delta(x)dx = \begin{cases} b-a & \text{if } \delta \le a \le b \\ b-\delta & \text{if } a \le \delta \le b \\ 0 & \text{if } a \le b \le \delta \end{cases} .$$

Then :

$$\Pr[\text{yes}\,|\,D] = \int_c^1 f_1(x)dx = \frac{2}{3}\int_c^1[S_{\frac{1}{4}}(x) + S_{\frac{1}{2}}(x) + S_{\frac{3}{4}}(x)]dx$$

$$= \frac{2}{3}[I_{\frac{1}{4}}(c, 1) + I_{\frac{1}{2}}(c, 1) + I_{\frac{3}{4}}(c, 1)]$$

$$\Pr[\text{yes}\,|\,D'] = \int_c^1 f_0(x)dx = \int_c^1 \left\{2 - \frac{2}{3}[S_{\frac{1}{4}}(x) + S_{\frac{1}{2}}(x) + S_{\frac{3}{4}}(x)]\right\} dx$$

$$= 2 - 2c - \frac{2}{3}[I_{\frac{1}{4}}(c, 1) + I_{\frac{1}{2}}(c, 1) + I_{\frac{3}{4}}(c, 1)] .$$

Table 10 gives values of Pr[yes | D] and Pr[yes | D'] which are plotted in Figure 11 as curve S. The graph consists of four straight line segments (one being horizontal and one vertical).

Table 10

Conditional Probabilities for Various Values of Criterion Level (Step Function)

c	Pr[yes | D']	Pr[yes | D]
0	1	1
.1	.8	1
.25	.5	1
.4	.3	.9
.5	.167	.833
.6	.1	.7
.75	0	.5
.9	0	.2
1.0	0	0

$$Pr[no|D] = \int_0^c f_1(x)dx = \frac{2}{3}[I_{\frac{1}{4}}(0, c) + I_{\frac{1}{2}}(0, c) + I_{\frac{3}{4}}(0, c)]$$

$$Pr[no|D'] = \int_0^c f_0(x)dx = 2c - \frac{2}{3}[I_{\frac{1}{4}}(0, c) + I_{\frac{1}{2}}(0, c) + I_{\frac{3}{4}}(0, c)]$$

$$R = \frac{2}{3}r[I_{\frac{1}{4}}(c, 1) + I_{\frac{1}{2}}(c, 1) + I_{\frac{3}{4}}(c, 1)]$$

$$+ (1-r)\left\{2c - \frac{2}{3}[I_{\frac{1}{4}}(0, c) + I_{\frac{1}{2}}(0, c) + I_{\frac{3}{4}}(0, c)]\right\} .$$

Simplifying gives:

$$R = r + 2c(1-r) - \frac{2}{3}[I_{\frac{1}{4}}(0, c) + I_{\frac{1}{2}}(0, c) + I_{\frac{3}{4}}(0, c)] .$$

Due to the nature of the frequency functions, it is not possible to solve for optimum c for a given r value from the equation $\frac{f_1(c)}{f_0(c)} = \frac{1-r}{r}$. Table 11 gives values of $\frac{f_1(c)}{f_0(c)}$ and Table 12 gives values of $\frac{1-r}{r}$.

Table 11

Frequency Ratio vs. Optimum Criterion Level (Step Function)

c	$\frac{f_1(c)}{f_0(c)}$
0 to .25	0
.25 to .5	.5
.5 to .75	2
.75 to 1	∞

Table 12

Table of $\frac{1-r}{r}$ vs. r
(Step Function)

r	$\frac{1-r}{r}$
0	∞
.1	9
.2	4
.3	2.33
.4	1.5
.5	1
.6	.667
.7	.428
.8	.25
.9	.111
1.0	0

Suppose r = .3. The value of $\frac{1-r}{r}$ is 2.33. The closest value of $\frac{f_1(c)}{f_0(c)}$ is 2 and c is between .5 and .75. To find the maximum value of R for r = .3, the value of c would have to be varied between .5 and .75 and R calculated. When this is done, R is found to increase linearly from .833 to .85 as c increases from .5 to .75. Above .75, it drops linearly to .70 for c = 1. Therefore, for r = .3, R_m = .85.

General Case of Frequency Functions

The general case of arbitrary frequency functions will now be considered. As an illustration, two arbitrary curves were drawn with a French curve. Function values for f_0 and f_1 were read and used to compute the areas under the curves by Simpson's one-third rule. The ordinates were then scaled so that the area under each curve equaled unity, thus giving frequency curves. These are shown in Figure 18. Table 13 gives the function values.

Table 13

Normalized Frequency Function Values for the Hypothetical Example of Figure 18

x	$f_0(x)$	$f_1(x)$	x	$f_0(x)$	$f_1(x)$
0	1.86	0	.55	.95	1.18
.05	1.805	0	.60	.86	1.26
.10	1.75	0	.65	.74	1.33
.15	1.675	.22	.70	.61	1.40
.20	1.62	.41	.75	.50	1.48
.25	1.53	.57	.80	.37	1.55
.30	1.45	.70	.85	.24	1.625
.35	1.36	.81	.90	.130	1.70
.40	1.265	.92	.95	0	1.77
.45	1.17	1.02	1.00	0	1.85
.50	1.06	1.11			

Values of $A_0(x) = \int_0^x f_0(x)dx$ and $A_1(x) = \int_0^x f_1(x)dx$ are given in Table 14 and plotted in Figure 19.

Table 14

Areas under Frequency Curves from 0 to x

x	$A_0(x)$	$A_1(x)$
0	0	0
.1	.1805	0
.2	.348	.022
.3	.501	.078
.4	.637	.1595
.5	.754	.261
.6	.849	.379
.7	.923	.512
.8	.973	.660
.9	.998	.823
1.0	1.000	1.000

The condition for optimum value of c is still:

$$\frac{f_1(c)}{f_0(c)} = \frac{1-r}{r} .$$

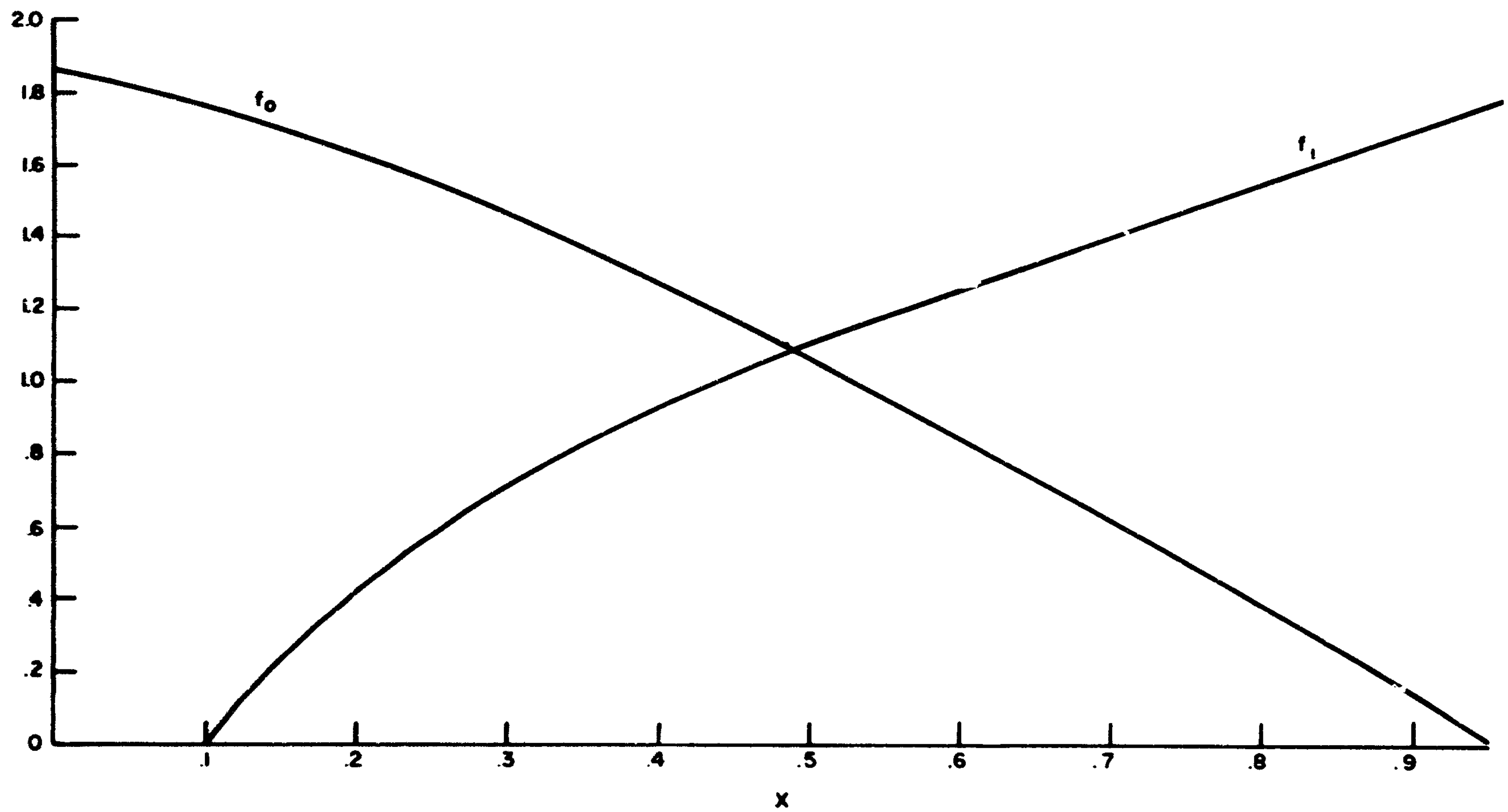

Figure 18 Arbitrary frequency curves

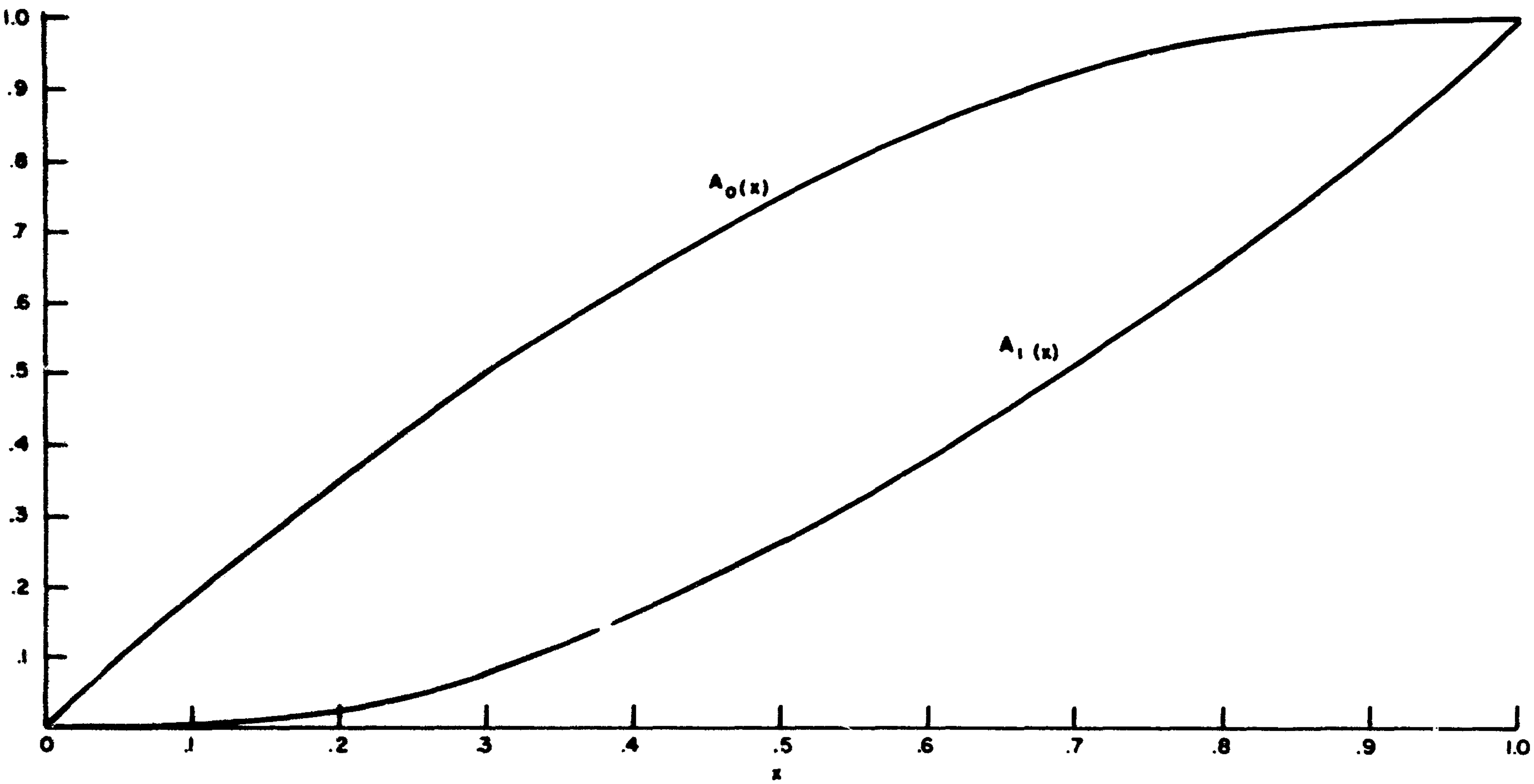

Figure 19 Distribution functions, with (A_1) and without (A_0) defect

For example, let $r = .3$. Then $\frac{1-r}{r} = \frac{7}{3}$. We must determine a value of x such that $\frac{f_1(x)}{f_0(x)} = \frac{7}{3}$. This can be done by drawing a template with two lines with slopes having ratios of $\frac{7}{3}$, as in Figure 20, on a transparent sheet.

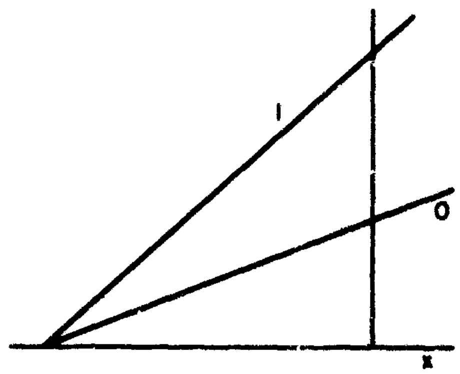

Figure 20 Template with lines having slopes 7:3 ratio

The height of every point on the upper line is $\frac{7}{3}$ the height of the corresponding point directly below it on the lower line. The x axis of this template is made coincident with the x axis of the graphs of f_0 and f_1 and displaced horizontally so that line 1 intersects f_1 and line 0 intersects f_0 for the same value of x. This occurs when $x = .707$ so that $c = .707$ for $r = .3$. The maximum value of R can now be computed for $r = 3$.

$$R_m = .3\int_{.707}^{1} f_1(x)dx + .7\int_{0}^{.707} f_0(x)dx$$

$$\int_{.707}^{1} f_1(x)dx = 1 - A_1(.707) = 1 - .520 = .480$$

$$\int_{0}^{.707} f_0(x)dx = .929$$

$$R_m = .3(.480) + .7(.929) = .794.$$

If other values of c are chosen (with r fixed at .3), lower values of R result. Table 15 illustrates this.

Table 15

Reliability vs. Criterion Level for Constant Frequency of Defect

c	R
.5	.75
.6	.78
.7	.793
.8	.783
.9	.75

$$Pr[yes|D'] = \int_c^1 f_0(x)dx = 1 - A_0(c)$$

$$Pr[yes|D] = \int_c^1 f_1(x)dx = 1 - A_1(c)$$

Table 16 gives values of Pr[yes|D'] and Pr[yes|D].

Table 16

Conditional Probabilities for Various Values of Criterion Level

c	Pr[yes|D']	Pr[yes|D]
0	1	1
.2	.652	.978
.4	.363	.841
.6	.151	.621
.8	.027	.340
1.0	0	0

The curve is so close to that labeled a = 1, that it was not plotted.

Discrete Distribution of Confidence

In the Naval operational situation, only a few levels of individual technician confidence are probably identifiable. When only a finite number of confidence values are involved, they can be represented as shown in Figure 21 and 22.

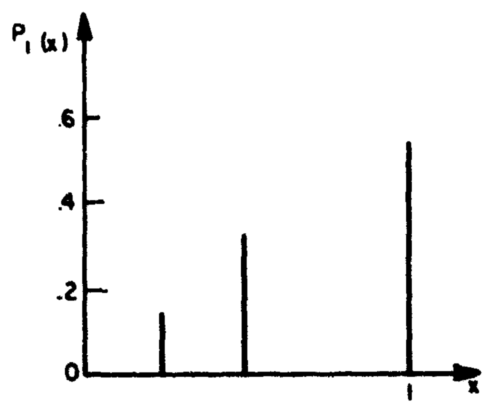

Figure 21 Probability distribution when a defect exists

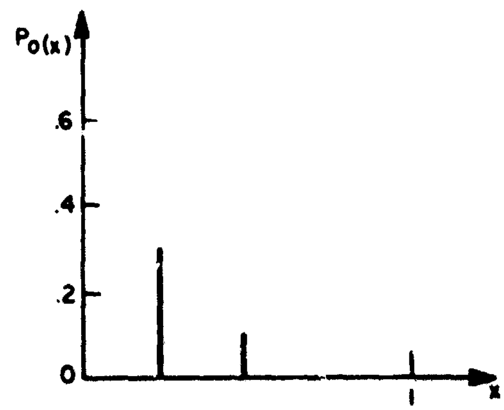

Figure 22 Probability distribution when there is no defect

The vertical lines have been drawn at various points to make the graphs more perspicuous. When a defect exists, the probability for $x = 1$ is highest (in this illustration, $p_1(1) = .54$). When there is no defect, the probability for $x = 0$ is highest ($p_0(0) = .6$). Values for this arbitrary example are shown in Table 17.

Table 17

Probability Distribution

x	$p_1(x)$	$p_0(x)$
0	0	.6
.23	.14	.3
.48	.32	.1
1	.54	0

In the discrete case, $p_0(x)$ and $p_1(x)$ correspond to $f_0(x)$ and $f_1(x)$ respectively. $\Sigma p_0(x)$ and $\Sigma p_1(x)$ correspond to $\int f_0(x)dx$ and $\int f_1(x)$ respectively. As before:

$$\Pr[\text{yes}] = \Pr[x \geq c]$$

$$\Pr[\text{yes}|D] = \sum_{x \geq c}^{1} p_1(x), \quad \Pr[\text{yes}|D'] = \sum_{x \geq c}^{1} p_0(x)$$

$$\Pr[\text{no}|D] = \sum_{x=0}^{x<c} p_1(x), \quad \Pr[\text{no}|D'] = \sum_{x=0}^{x<c} p_0(x) .$$

For the discrete case, when c is a value for which p_0 or p_1 is defined, then it makes a difference whether $\Pr[\text{yes}] = \Pr[x \geq c]$ or $\Pr[x > c]$. The two probabilities differ by $p(c)$.

$$R = r\sum_{x \geq c}^{1} p_1(x) + (1-r)\sum_{x=0}^{x<1} p_0(x)$$

This is also a maximum when:

$$\frac{p_1(c)}{p_0(c)} = \frac{(1-r)}{r} .$$

Consider an example which corresponds to the continuous example shown in Figure 3. Table 18 gives the values of p_0 and p_1. Table 18 also presents the (cumulative) distribution and ratios of p_1 to p_2. For $r = .8$, $\frac{1-r}{r} = .25$. From the table, we see that the nearest value of the ratio is $\frac{1}{3}$, so c should lie between 0 and .25.

Table 18

Probability Distribution and Ratios

x	$p_1(x)$	$p_0(x)$	$\sum_{x=0}^{x} p_1(x)$	$\sum_{x=0}^{x} p_0(x)$	$\frac{p_1(x)}{p_0(x)}$
0	0	.4	0	.4	0
.25	.1	.3	.1	.7	1/3
.50	.2	.2	.3	.9	1
.75	.3	.1	.6	1	3
1.0	.4	0	1.0	1	∞

Table 19 gives ranges of optimum c and values of R_m for various values of r.

Table 19

Range of Optimum Criterion Level and Maximum Reliability vs. Relative Frequency of Defects

r	c	R_m
0	.75 - 1	1
.2	.75 - 1	.88
.3	.5 - .75	.84
.4	.5 - .75	.82
.5	.5	.80
.6	.25 - .5	.82
.7	.25 - .5	.84
.8	0 - .25	.88
1.0	0 - .25	1

The values of R_m were obtained from Table 20, which gives values of R for various values of c and r.

Table 20

Reliability and Conditional Probabilities vs. Criterion Level

c	Pr[yes|D] $\sum_{x \geq c} p_1(x)$	$\sum_{x=0}^{x<c} p\ (x)$	R r = .2	r = .3	r = .4	r = .5	Pr[yes|D']
>0	1	.4	.52	.58	.64	.7	.6
.2	1	.4	.52	.58	.64	.7	.6
.4	.9	.7	.74	.76	.78	.8	.3
.5	.9	.7	.74	.76	.78	.8	.3
.6	.7	.9	.86	.84	.82	.8	.1
.8	.4	1	.88	.82	.76	.7	0
1.0	.4	1	.88	.82	.76	.7	0

The operating characteristic curve has four points: (.6, 1), (.3, .9), (.1, .7), and (0, .4). The R_m curve has a minimum of .8 for $r = .5$.

Double Criterion Level Detection Model

To this point, we have discussed a detection model with a single criterion level. When the confidence, x, of there being a defect is greater than or equal to the criterion level, c, the subject will say "yes" (there is a defect). When $x < c$, the subject will say "no."

Now consider two criterion levels, c_1 and c_2, where $c_1 < c_2$. If $x \leq c_1$, the subject will say "no." If $x \geq c_2$, he will say "yes." If $c_1 < x < c_2$, then the subject will make another attempt to determine the status of the equipment. This continues until he says either "yes" or "no." This type of procedure has also been called by Swets (1964) " sequential observations." A third possibility, not considered here, is that he may be allowed only a limited number of trials. If not able to decide, this might constitute a third type of unreliable operation on the part of the subject, the first two types occurring when subject says "no" with a defect, and "yes" with no defect.

Assume that the curves shown in Figure 23 represent frequency functions.

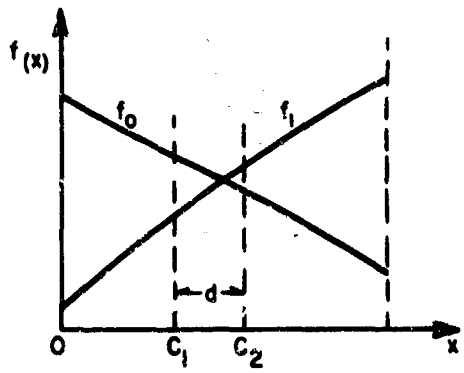

Figure 23 Hypothetical frequency function, with (f_1) and without (f_0) defect

The area under each would then be unity. The two criterion levels are represented by c_1 and c_2. The area under a curve f from c_2 to 1 [call it $A(c_2, 1)$] represents the probability (p_1) that the technician will say "yes" on one trial. Similarly, $A(0, c)$ is the probability (p_3) of saying "no" and is represented by the area under a curve f from 0 to c_1 on one trial. The probability (p_2) of being undecided on a single trial is the area under a curve f from c_1 to c_2. The trials are considered independent. The probability tree in Figure 24 shows only three stages. It holds for either a defect (D) or no defect (D').

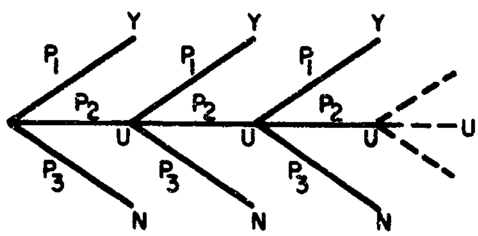

Figure 24 Probability tree

$$Pr[yes] = p_1 + p_2 p_1 + p_2^2 p_1 + p_2^3 p_1 \quad \ldots = \frac{p_1}{1 - p_2} = \frac{p_1}{p_1 + p_3}$$

$$Pr[no] = p_3 + p_2 p_3 + p_2^2 p_3 + p_2^3 p_3 \quad \ldots = \frac{p_3}{1 - p_2} = \frac{p_3}{p_1 + p_3}$$

Specifically:

$$Pr[yes \mid D] = \frac{A_1(c_2, 1)}{A_1(c_2, 1) + A_1(0, c_1)}$$

$$Pr[no \mid D'] = \frac{A_0(0, c_1)}{A_0(c_2, 1) + A_0(0, c_1)} .$$

Finally, reliability is:

$$R = r\,Pr[yes \mid D] + (1 - r) Pr[no \mid D']$$

$$= \frac{r A_1(c_2, 1)}{A_1(c_2, 1) + A_1(0, c_1)} + \frac{(1 - r) A_0(0, c_1)}{A_0(c_2, 1) + A_0(0, c_1)} .$$

The expected number of trials for a decision (yes or no), if up to n trials are allowed, was derived previously (Siegel & Miehle, 1967) as:

$$E_n = \frac{1}{1 - p_2} \left\{ 1 - [1 + n(1 - p_2)]p_2^n \right\} + np_2^n$$

$$\lim_{n \to \infty} En = \frac{1}{1 - p_2} = \frac{1}{A(c_2, 1) + A(0, c_1)} .$$

This holds for both the cases in which a defect or no defect exists. The case where an unlimited number of trials are allowed therefore is:

$$E_\infty = \frac{r}{A_1(c_2, 1) + A_1(0, c_1)} + \frac{1 - r}{A_0(c_2, 1) + A_0(0, c_1)} .$$

The increased reliability obtained by repeated trials is obtained at the expense of increased time and effort.

As an illustration, consider the linear frequency functions of Figure 3.

$$A_1(c_2, 1) = \int_{c_2}^{1} f_1(x)dx = \int_{c_2}^{1} 2xdx = x^2 \Big|_{c_2}^{1} = 1 - c_2^2$$

$$A_1(0, c_1) = x^2 \Big|_0^{c_1} = c_1^2$$

$$A_0(c_2, 1) = \int_{c_2}^{1} f_0(x)dx = \int_{c_2}^{1} (2 - 2x)dx = 2x - x^2 \Big|_{c_2}^{1} = 1 - 2c_2 + c_2^2$$

$$A_0(0, c_1) = 2x - x^2 \Big|_0^{c_1} = 2c_1 - c_1^2$$

Let $d = c_2 - c_1$:

$$R = \frac{r(1 - c_2^2)}{1 - c_2^2 + c_1^2} + \frac{(1 - r)(2c_1 - c_1^2)}{1 - 2c_2 + c_2^2 + 2c_1 - c_1^2}$$

$$= \frac{r(1 - c_2^2)}{1 - d(c_1 + c_2)} + \frac{(1 - r)c_1(2 - c_1)}{1 + d(c_1 + c_2) - 2d}$$

$$E_\infty = \frac{r}{1 - d(c_1 + c_2)} + \frac{1 - r}{1 - 2d + d(c_1 + c_2)} \; .$$

E_∞ is plotted as a function of $c = \frac{c_1 + c_2}{2}$ for r = .3, .5 for d = .2, .4, .6 (Figure 25). R is plotted as a function of c for d = 0, .2, .4, .6 for r = .3 in Figure 26, and for r = .5 in Figure 27.

It can be seen that the optimum value of c is .5 for r = .5 and that the maximum value of R increases as the separation (d) between the criterion levels increases. For r = .3, the optimum value of c decreases as d increases. The maximum values (R_m) of R increase with increasing d and are higher than those for r = .5 as expected.

The curves of E_∞ for d > .2 have relative minima whose location shifts to the left for increasing d, but they do not coincide (except for r = .5) with the maxima of R.

Table 21 gives values of optimum c, R_m, and E_∞ for various values of d, when r = .3 and .5.

Table 21

Optimum Average Criterion Level, Maximum Reliability, and Expected Number of Trials vs. Separation (d) of Criterion Levels

	r = .3			r = .5		
d	c	R_m	E_∞	c	R_m	E_∞
0	.7	.79	1	.5	.75	1
.2	.63	.831	1.22	.5	.80	1.25
.4	.59	.868	1.61	.5	.85	1.67
.6	.54	.910	2.41	.5	.90	2.5

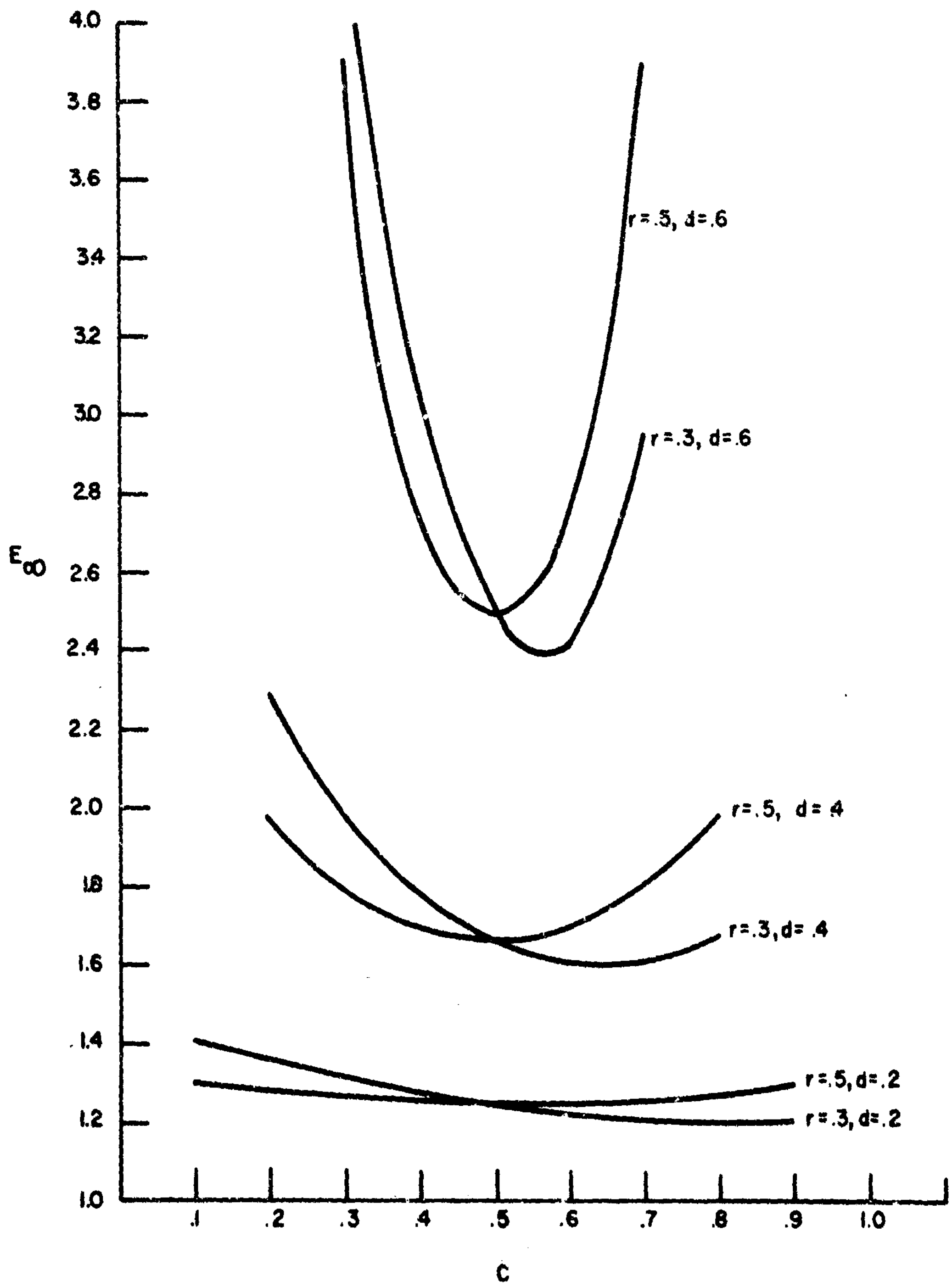

Figure 25 Expected number of trials vs. average criterion level

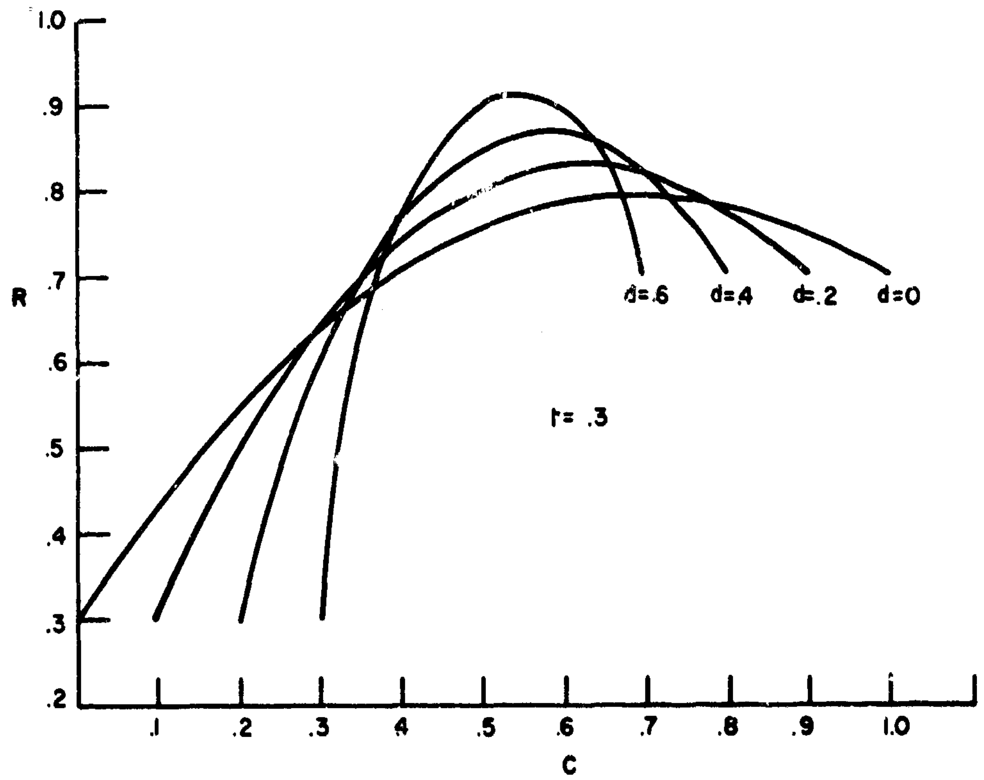

Figure 26 Reliability vs. average criterion level
(when r = .3)

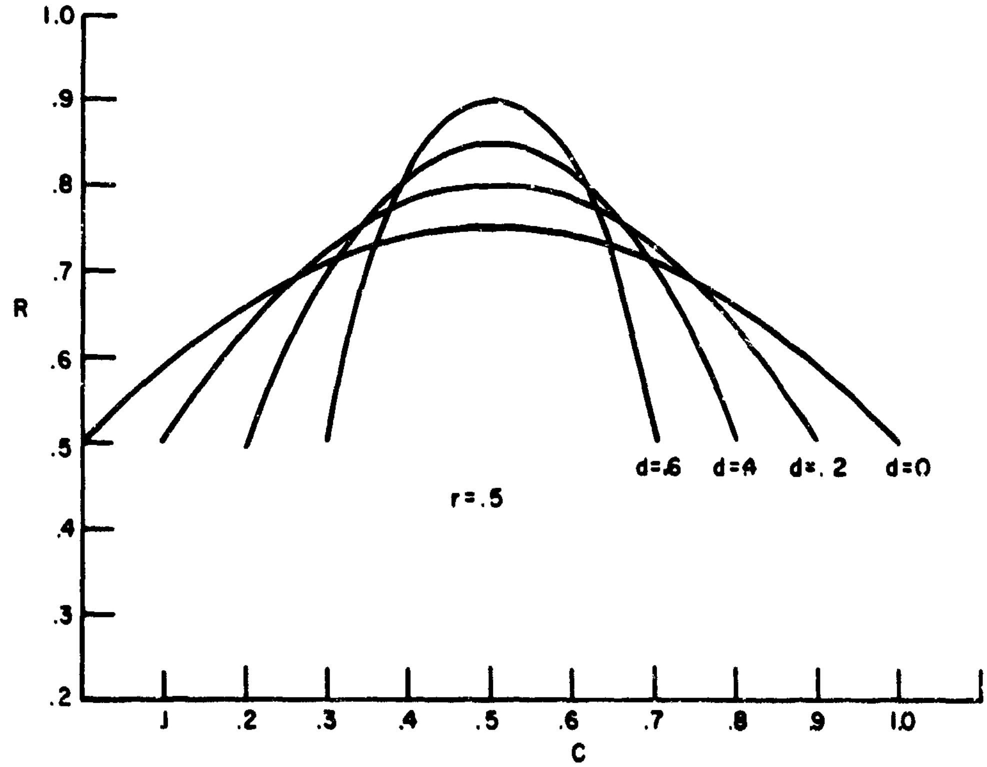

Figure 27 Reliability vs. average criterion level
(when r = .5)

CHAPTER III

DISCUSSION, SUMMARY, AND CONCLUSION

Discussion

In the present report, reliability and operating characteristic curves were considered for various types of "confidence of the existence of a defect" distributions. This is an adaptation of the method used by Tanner and Swets (1954), where stimulus is replaced by confidence. If the distributions for defect and no defect are symmetrical about $x = \frac{1}{2}$, then the maximum reliability (R_m) is lowest when a defect occurs with probability of $\frac{1}{2}$. This is the situation of maximum uncertainty. For all continuous distributions considered, the maximum reliability attainable equals 1 when a defect is never present ($r = 0$) and when a defect is always present ($r = 1$). The lowest value of R_m (at $r = .5$) depends on the distributions. The smaller the common area under both curves (D and D'), the larger the R_m values are (see Table 22). When the curves do not intersect, R_m is 1 for all values of r. When r is low, the optimum criterion level (c) is high, and when r is high, c is low. From the point at which c first becomes 0, the plot of R_m vs. r becomes a straight line ($R = r$). From the point where c first becomes 1, R_m vs. r becomes the straight line $R = 1 - r$.

Table 22

Comparison of Common Area under f_1 and f_0 Curves with Maximum Reliability for $r = .5$

Case	Area	R_m for $r = .5$
Non-intersecting	0	1
Straight line, $a = .75$	.222	.889
Straight line, $a = 1$	.5	.75
Normal, $\sigma = \frac{1}{2}$	.57	.715
Exponential	.755	.623
Normal, $\sigma = 1$	.876	.561

A similar progression holds for the operating characteristic curves. As the common area under the distribution curves decreases, the operating characteristic curves move closer to the left and upper boundaries.

The reliability value (R) may be improved in two ways. The first is to minimize or eliminate the intersection of the frequency curves, as indicated by Table 22. When the curves are non-intersecting, the maximum attainable reliability is 1. This reliability is obtained when the confidence criterion level is properly set. One method for achieving this level is to design the equipment so that the existence of a defect is more obvious. This will separate curves f_1 and f_0, making the setting of a criterion level less critical. With non-intersecting curves, the optimum criterion level is anywhere between the x intercepts of the curves.

With a given pair of frequency curves, the reliability may also be improved by training the technician so that he sets his criterion level towards the optimum value. This can be accomplished by providing the technician with feedback of information about his performance or possibly through programmed instructional techniques.

Summary and Conclusion

The present report has attempted to point out the relationship between the subjective criterion level that a maintenance technician sets or adapts for accepting or rejecting the hypothesis that a malfunction exists and the probability of his acceptance or rejection of the hypothesis being correct. The logic employed is similar to that used in signal detection theory, and operating characteristic curves were similarly derived. These may be interpreted in much the same way as the receiver operator curves derived in the psychophysical investigation of signal detection. The relationships were derived for a number of continuous distributions of confidence and for the discrete case. Both the single and double criterion instances were considered. It is believed that the discrete case with a single criterion level represents the most useful case, from the point of view of application. While we have not attempted to specify a method for measuring a technician's confidence criterion behavior, the production of such a change, and the consequent increase in correct hypothesis acceptance and false hypothesis rejection, does not seem unattainable.

The gains to be derived from the production of such a shift in criterion level are considerable. Consider example 1 (Figure 3). Assume the probability of a defect is $r = .1$. The optimum criterion level is $c = 1 - r = .9$. At that level, the maximum possible reliability is $R_m = r^2 - r + 1 = .01 - .1 + 1 = .91$. Suppose the technician actually assumes a level of $c = .5$. Then his reliability is

$$R = r + 2c - c^2 - 2rc$$

$$= .75$$

This is a considerable reduction from the maximum of .91.

It is possible that a programmed learning technique or program could be developed to allow individual technician training on acceptance criterion level setting. Thus, the acceptance criterion level training could be administered quite economically, with the usual advantages of programmed learning.

Alternatively, the employment of the acceptance level setting behavior of technicians as a fleet performance criterion seems tenable. Technicians who set their criterion of acceptance so as to maximize correct hypothesis and rejection are probably superior to those who do not. Thus, if for a pair of technicians and a given set of equipment conditions, technician A sets his criterion level less close to the optimum than technician B, then technician B may be said to be superior. Techniques for performing such measures should be readily derivable.

Finally, it was pointed out that the employment of the operating characteristic curve possesses implications for decisions regarding the development of automatic test equipment. Test equipment would be built for those situations in which malfunction detection deviates significantly from the optimum, or stated alternatively, where it approaches the chance level.

In conclusion, it seems that a method for establishing the optimum subjective criterion of acceptance level has been established for the malfunction recognition situation.

REFERENCES

Siegel, A. I., & Pfeiffer, M. G. Posttraining performance criterion development and application. Personnel psychophysics: Estimating personnel subsystem reliability through magnitude estimation methods. Wayne, Pa., Applied Psychological Services, 1966.

Siegel, A. I., & Miehle, W. Posttraining performance criterion development and application. Personnel psychophysics: Extension of a prior subsystem reliability determination technique. Wayne, Pa., Applied Psychological Services, 1967.

Swets, J. (Ed.) Signal detection and recognition by human observers. New York: Wiley, 1964.

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) Applied Psychological Services Science Center Wayne, Pennsylvania	2a REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b GROUP

3 REPORT TITLE

Personnel Psychophysics: Quantification of Malfunction Detection Probability

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

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

Mielhe, Wm.

Siegel, Arthur I.

6 REPORT DATE	7a TOTAL NO. OF PAGES	7b NO. OF REFS
December 1967	51	3

8a CONTRACT OR GRANT NO. N00014-67-C0107 b PROJECT NO. c d	9a ORIGINATOR'S REPORT NUMBER(S) 9b OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10 AVAILABILITY/LIMITATION NOTICES

Qualified requesters may obtain copies of this report through DDC

11 SUPPLEMENTARY NOTES	12 SPONSORING MILITARY ACTIVITY Personnel and Training Branch Psychological Sciences Division Office of Naval Research

13 ABSTRACT

The logic of a technique for employing technician "confidence that a defect exists" for maximizing the probability of malfunction recognition is described. The technique is based on and drawn from parallel thinking in signal detection theory. Operator characteristic curves are derived for a variety of distributions of "confidence." Continuous and discrete distributions of "confidence" are considered as well as single and double criterion levels. The implications of the work for training and posttraining performance evaluation are pointed out.

DD FORM 1 JAN 64 1473

UNCLASSIFIED

Security Classification

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Personnel and Training Psychophysics Quantitative Methods Reliability Performance Evaluation						

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.

2a. 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.

2b. 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.

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

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

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

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

9a. 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.

9b. 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.

UNCLASSIFIED
Security Classification