SOME MEASURES FOR DISCRIMINATING BETWEEN NORMAL
MULTIVARIATE DISTRIBUTIONS WITH EQUAL
COVARIANCE MATRICES

BY
HERMAN CHERNOFF

TECHNICAL REPORT NO. 73
AUGUST 28, 1972

PREPARED UNDER CONTRACT
NO0014-67-A-0112-0051 (NR-042-993)
FOR THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
SOME MEASURES FOR DISCRIMINATING BETWEEN NORMAL MULTIVARIATE DISTRIBUTIONS WITH EQUAL COVARIANCE MATRICES

BY

HERMAN CHERNOFF

TECHNICAL REPORT NO. 73
AUGUST 28, 1972

PREPARED UNDER CONTRACT N00014-67-A-0112-0051 (NR-042-993) FOR THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
1. **Summary and Introduction**

In a previous paper [5], a measure $S$ was described which indicates how well one may discriminate between two normal multivariate distributions using a linear discriminant function. This measure, when applied to an example in design of experiments, led to a somewhat unexpected conclusion.

Briefly, suppose that under $H_1$, $X$ is a $N(1,1)$ random variable and $Y$ is independently $N(1,9)$ where $N(\mu,\sigma^2)$ represents the normal distribution with mean $\mu$ and variance $\sigma^2$. Under $H_2$, $X$ is $N(-1,9)$ and $Y$ is independently $N(-1,1)$. The statistician who wishes to discriminate between $H_1$ and $H_2$ is permitted $n$ observations on $X$. In addition he is given a choice between $n$ more observations on $X$ or $n$ observations on $Y$. Applying the measure $S$ he is led to prefer the unbalanced choice of $2n$ observations on $X$ to the balanced one of $n$ observations on each $X$ and $Y$. Thus it appears that if his first observations are more precise under $H_1$ than $H_2$, there is a premium on taking additional observations which are more precise under $H_1$ than $H_2$. 
Is this result a consequence of restricting attention to linear
discriminant functions? In this paper we extend the measure to \( T \)
which is appropriate for using the likelihood-ratio test. It is shown
that using \( T \) there still is a premium for the unbalanced choice.

Section 4 contains a brief discussion of the Kullback-Leibler
information numbers and how they relate to \( S \) and \( T \).

2. The Measure \( S \)

Becker [2] suggested that

\[
S = \frac{|\mu_1 - \mu_2|}{\sigma_1^2 + \sigma_2^2}
\]

is a useful measure of separation between two distributions \( F_i \) with
mean \( \mu_i \) and variance \( \sigma_i^2 \), \( i = 1, 2 \). This measure appeared in [3],
and the multivariate extension

\[
S = \sup_{b \neq 0} \frac{|b'(\mu_1 - \mu_2)|}{(b'\Sigma_1 b)^{1/2} + (b'\Sigma_2 b)^{1/2}}
\]

(2.1)

was shown to be relevant in discriminating between two multivariate
normal distributions \( F_i = N(\mu_i, \Sigma_i) \), \( i = 1, 2 \) when linear discriminant
functions are used. In that case the linear discriminant function which
minimizes the maximum error probability consists of selecting \( F_1 \) if

\[
b'_o X > \frac{(b'_o\Sigma_1 b'_o)^{-1/2}\mu_1 + (b'_o\Sigma_2 b'_o)^{-1/2}\mu_2}{(b'_o\Sigma_1 b'_o)^{-1/2} + (b'_o\Sigma_2 b'_o)^{-1/2}}
\]
where \( b_0 \) is the vector which minimizes the expression in (2.1). The corresponding error probabilities are

\[
\varepsilon_1 = \varepsilon_2 = \Phi(-S).
\]

If \( X \) is replaced by the sample mean of \( n \) independent observations, then the error probabilities approach zero exponentially fast in \( n \). In fact, the error probabilities are

\[
(2.2) \quad \varepsilon_{1n} = \varepsilon_{2n} = \Phi(-\sqrt{n}S) \approx \frac{1}{\sqrt{2\pi nS^2}} e^{-nS^2/2}.
\]

The following theorem, presented in [5], is essentially a restatement of results in [1] and [6]. It may be derived by applying the method of Lagrange multipliers to the relatively simple calculation of \( S \) for the multivariate distributions \( N(\mu_i, \Sigma_i), i = 1, 2, \delta = \mu_1 - \mu_2 \).

We assume that the \( \Sigma_i \) are positive definite.

**Theorem 1:**

\[
(2.3) \quad S^2 = t(1-t)\delta \Sigma^{-1} \delta
\]

where

\[
(2.4) \quad \Sigma = t \Sigma_1 + (1-t) \Sigma_2,
\]

and \( t \) is the unique solution between \( 0 \) and \( 1 \) of

\[
(2.5) \quad R(t) = \delta \Sigma^{-1} [t^2 \Sigma_1 - (1-t)^2 \Sigma_2] \Sigma^{-1} \delta = 0.
\]
The optimal value of \( b \) is given by

\[(2.6) \quad b_0 = \Sigma^{-1}\delta \]

and is unique up to a multiplicative constant. Furthermore,

\[(2.7) \quad |b_0'\delta| = 8'\Sigma^{-1}\delta = 8^2/t(1-t) \]

and

\[(2.8) \quad t^2b_0'\Sigma_1b_0 = (1-t)^2b_0'\Sigma_2b_0 = 8^2 \]

The fact that

\[(2.9) \quad \frac{dR}{dt} = 2\delta'\Sigma_1^{-1}\Sigma_1\Sigma_2^{-1}\Sigma_2\Sigma_1^{-1}\delta > 0 \quad \text{for} \quad \delta \neq 0 \]

permits us to apply the Newton iterative technique where an approximation \( t^* \) to \( t \) is improved to

\[
t^{**} = t^* - \left[ \frac{dR(t^*)}{dt} \right]^{-1} R(t^*) \]

Theorem 1 presents \( S^2 \) as a multiple of \( \delta'\Sigma^{-1}\delta \) which may be regarded as a Mahalanobis distance with respect to the weighted average \( \Sigma \) of the two covariance matrices \( \Sigma_1 \) and \( \Sigma_2 \). Note that if \( \Sigma_1 = \Sigma_2 \), the minimizing value of \( t \) is 0.5, \( \Sigma = \Sigma_1 = \Sigma_2 \) and \( S^2 = (\delta'\Sigma^{-1}\delta)/4 \).

Suppose \((X,Y)\)' is \( N(\mu_i, \Sigma_i) \) under \( H_i, i = 1,2 \), where \( \mu_1 = (1,1) \), \( \mu_2 = (-1,-1) \), \( \Sigma_1 = (1 0 \ 0 9) \), and \( \Sigma_2 = (9 0 \ 0 1) \). Then \( S^2 \) corresponding to \( X \) is 0.25, while \( S^2 \) corresponding to \((X,Y)\) is 0.40. Applying (2.2) the exponential rate at which the error probabilities approach zero as the sample size approaches \( \omega \) is determined by \( S^2 = 0.25 \) per..
observation for $X$ and $S^2/2 = 0.20$ per observation for $(X,Y)$ if $(X,Y)$ counts for two observations.

This illustration can be interpreted by the remark in the introduction to the effect that when using linear discriminant functions, the fact that some of the data are more precise under $H_1$ than under $H_2$ implies that there is a premium on additional data of the same sort rather than on data which are more precise under $H_2$ than under $H_1$. This remark is relevant in the nonsequential case where the choice of data to be observed is to be made before any data are gathered.

Is this phenomenon due to the fact that the linear discriminant function neglects relevant information? Would it disappear if one applied the likelihood-ratio test? To answer this question we proceed to Section 3.

3. The Measure $T$

In [3] it was shown that if the likelihood-ratio is selected to minimize $L_1 \log \epsilon_{1n} + \lambda \epsilon_{2n}$ for fixed $\lambda > 0$ or to minimize $\max(\epsilon_{1n}, \epsilon_{2n})$ on the basis of $n$ independent observations on $X$ with density $f_i(x)$ under $H_i$, $i = 1, 2$, then

$$\lim [n^{-1} \log \epsilon_{1n}] = -I$$

where

$$I = -\log \inf_{t \leq 1} \int_{1-t}^{t} f_1^{-t}(x)f_2^{t}(x)dx.$$
Furthermore, the test which decides according to the sign of the logarithm of the likelihood-ratio attains error probabilities satisfying (3.1). Thus comparison with (2.3) shows that $S^2/2$ is comparable to $I$ or that $T$, defined by

\[(3.3) \quad I = T^2/2,\]

is comparable to $S$ as a measure of separation between two distributions.

The following theorem characterizes $T$ for two distinct multivariate normal distributions with positive definite matrices.

**Theorem 2:**

\[(3.4) \quad T^2 = \sup_{0 \leq t \leq 1} \left\{ t(1-t)\delta^t \Sigma^{-1} \delta + \log \left| \frac{\Sigma}{\Sigma_1^{1-t} \Sigma_2^{t-1}} \right| \right\} \]

where

$$\Sigma = t\Sigma_1 + (1-t)\Sigma_2$$

and if the expression in braces in (3.4) is $H(t)$,

\[(3.5) \quad H'(t) = \delta^t \Sigma^{-1} [(1-t)\Sigma_2^t + t\Sigma_1^t] \Sigma^{-1} \delta + \log |\Sigma_1^{-1} \Sigma_2^{-1}| + \text{tr}[(\Sigma_1^{-1} - \Sigma_2^{-1}) \Sigma^{-1}] \]

and

\[(3.6) \quad H''(t) = -2\delta^t \Sigma^{-1} \Sigma_1^{-1} \Sigma_2^{-1} \delta - \text{tr}[(\Sigma_1^{-1} - \Sigma_2^{-1}) \Sigma^{-1} (\Sigma_1^{-1} - \Sigma_2^{-1})] < 0 . \]

Before deriving this result we remark that the concavity of log determinant implies that $T^2 > S^2$ which is anticipated from the fact
that the likelihood-ratio test is at least as good as the best linear
discriminant function. This theorem leads to a computational procedure
for $T^2$ which is very little different from that indicated for $S^2$
and only slightly more involved.

Proof:

$$f_{1}^{1-t} f_{2}^{t} = (2\pi)^{-k/2}|\Sigma_{1}|^{-1/2} |\Sigma_{2}|^{-t/2} \exp\left[\frac{-1}{2} \left((1-t)(x-\mu_{1})'\Sigma_{1}^{-1}(x-\mu_{1})
+ t(x-\mu_{2})'\Sigma_{2}^{-1}(x-\mu_{2})\right)\right].$$

Completing the square, the expression in the square brackets may be
written as

$$(x-\mu)'A(x-\mu) + c$$

where by matching coefficients we have

(3.7a) \hspace{1cm} A = (1-t)\Sigma_{1}^{-1} + t\Sigma_{2}^{-1}

(3.7b) \hspace{1cm} A\mu = (1-t)\Sigma_{1}^{-1}\mu_{1} + t\Sigma_{2}^{-1}\mu_{2}

and

(3.7c) \hspace{1cm} \mu'A\mu + c = (1-t)\mu_{1}'\Sigma_{1}^{-1}\mu_{1} + t\mu_{2}'\Sigma_{2}^{-1}\mu_{2} .

Then

(3.8) \hspace{1cm} \int f_{1}^{1-t}(x)f_{2}^{t}(x)dx = |\Sigma_{1}|^{-(1-t)/2} |\Sigma_{2}|^{-t/2} |A|^{-1/2} e^{-c/2} .
We have

\begin{equation}
\Sigma_1 A \Sigma_2 = \Sigma_2 A \Sigma_1 = \Sigma = t \Sigma_1 + (1-t) \Sigma_2
\end{equation}

\[ \mu' A \mu = [(1-t) \mu_1^{(1)} \Sigma_1^{(-1)} + t \mu_2^{(1)} \Sigma_2^{(-1)}] A^{-1} [(1-t) \Sigma_1^{(-1)} \mu_1 + t \Sigma_2^{(-1)} \mu_2] \]

and

\[ c = (1-t) \mu_1^{(1)} \Sigma_1^{(-1)} \mu_1 - (1-t)^2 \mu_1^{(1)} \Sigma_1^{(-1)} A^{-1} \Sigma_1^{(-1)} \mu_1 + t \mu_1^{(1)} \Sigma_1^{(-1)} A^{-1} \Sigma_2^{(-1)} \mu_2 - t^2 \mu_2^{(1)} \Sigma_2^{(-1)} A^{-1} \Sigma_2^{(-1)} \mu_2 - t(1-t)[\mu_1^{(1)} \Sigma_1^{(-1)} A^{-1} \Sigma_2^{(-1)} \mu_2 + \mu_2^{(1)} \Sigma_2^{(-1)} A^{-1} \Sigma_1^{(-1)} \mu_1] . \]

Now

\[ A^{-1} = \Sigma_1 \Sigma_2^{(-1)} = \Sigma_2 \Sigma_1^{(-1)} \]

\[ \mu_1^{(1)} \Sigma_1^{(-1)} \mu_1 = \mu_1^{(1)} \Sigma_1^{(-1)} (t \Sigma_1 + (1-t) \Sigma_2) \Sigma_1^{(-1)} \mu_1 = t \mu_1^{(1)} \Sigma_1^{(-1)} \mu_1 + (1-t) \mu_1^{(1)} \Sigma_2 \Sigma_1^{(-1)} \mu_1 \]

\[ \mu_1^{(1)} \Sigma_1^{(-1)} A^{-1} \Sigma_1^{(-1)} \mu_1 = \mu_1^{(1)} \Sigma_1^{(-1)} \Sigma_2 \Sigma_1^{(-1)} \mu_1 \]

\[ \Sigma_1^{(-1)} A^{-1} \Sigma_1^{(-1)} = \Sigma_2 \Sigma_1^{(-1)} \Sigma_2 \Sigma_1^{(-1)} = \Sigma_1^{(-1)} . \]

By symmetry with respect to the interchange of 1 with 2 and t with (1-t), it follows that

\[ c = t(1-t)[\mu_1^{(1)} \Sigma_1^{(-1)} \mu_1 + \mu_2^{(1)} \Sigma_1^{(-1)} \mu_2 - \mu_1^{(1)} \Sigma_1^{(-1)} \mu_2 - \mu_2^{(1)} \Sigma_1^{(-1)} \mu_1] \]

\begin{equation}
(3.10) \quad c = t(1-t) \delta' \Sigma^{(-1)} \delta
\end{equation}

Applying (3.9) and (3.10) to (3.8) we have

\[ \log[\int_1 f_1^{1-t}(s) f_2^t(x) dx] = \frac{1}{2} H(t) \]
where
\[ H(t) = t(1-t)\delta\Sigma^{-1}\delta + \log \frac{|\Sigma|}{|\Sigma_1|^t|\Sigma_2|^{1-t}}. \]

We recall the expansions
\[ (A+h\Delta)^{-1} = A^{-1} - hA^{-1}\Delta A^{-1} + h^2A^{-1}\Delta A^{-1}\Delta A^{-1} + \cdots \]
\[ \log|A+h\Delta| = \log|A| + h\text{tr}(A^{-1}\Delta) - \frac{h^2}{2}\text{tr}(A^{-1}\Delta A^{-1}\Delta) + \cdots \]
from which it follows that
\[ H'(t) = (1-2t)\delta\Sigma^{-1}\delta - t(1-t)\delta\Sigma^{-1}(\Sigma_1-\Sigma_2)\Sigma^{-1}\delta \]
\[ + \text{tr}[\Sigma^{-1}(\Sigma_1-\Sigma_2)] + \log|\Sigma_1^{-1}\Sigma_2| \]
and
\[ H''(t) = -2\delta\Sigma^{-1}\delta - 2(1-2t)\delta\Sigma^{-1}(\Sigma_1-\Sigma_2)\Sigma^{-1}\delta + 2t(1-t)\delta\Sigma^{-1}(\Sigma_1-\Sigma_2)\Sigma^{-1}(\Sigma_1-\Sigma_2)\Sigma^{-1}\delta \]
\[ - \text{tr}\Sigma^{-1}(\Sigma_1-\Sigma_2)\Sigma^{-1}(\Sigma_1-\Sigma_2). \]

Using
\[ \Sigma^{-1} = \Sigma^{-1}(t\Sigma_1+(1-t)\Sigma_2)\Sigma^{-1} \]
in the first term for \( H'(t) \) yields (3.5). To obtain (3.6), use the above relation for the second \( \Sigma^{-1} \) as well as
\[ \Sigma^{-1} = \Sigma^{-1}(t\Sigma_1+(1-t)\Sigma_2)\Sigma^{-1}(t\Sigma_1+(1-t)\Sigma_2)\Sigma^{-1} \]
in the first term in the expression for \( H'' \), yielding
$$H''(t) = -2\delta \Sigma^{-1}(t \Sigma^{-1} \Sigma_2 + (1-t) \Sigma_2 \Sigma^{-1} \Sigma_1) \Sigma^{-1} \delta - \text{tr}[\Sigma^{-1}(\Sigma_1 - \Sigma_2) \Sigma^{-1}(\Sigma_1 - \Sigma_2)].$$

Equation (3.6) follows when we recall that $A^{-1} = \Sigma_1 \Sigma^{-1} \Sigma_2 = \Sigma_2 \Sigma^{-1} \Sigma_1$.

From (3.7) we note that $A$ and hence $A^{-1}$ is positive definite. Thus the first term of $H''$ is negative unless $\delta = 0$. The fact that the second term of $H''$ is negative unless $\Sigma_1 = \Sigma_2$ can be derived from the concavity of log determinant or more directly by applying a non-singular linear transformation which simultaneously orthogonalizes $\Sigma_1$ and $\Sigma_2$. Thus if $\Sigma_1 = R\Lambda_1 R'$ and $\Sigma_2 = R\Lambda_2 R'$ where $\Lambda_1$ and $\Lambda_2$ are diagonal positive definite matrices, the second term becomes

$$-\text{tr}[\Lambda^{-1}(\Lambda_1 - \Lambda_2)\Lambda^{-1}(\Lambda_1 - \Lambda_2)]$$

where $\Lambda = t\Lambda_1 + (1-t)\Lambda_2$. Hence $H'' < 0$ as long as the two multivariate distributions are distinct.

The algebra of this derivation could have been reduced considerably by relating $H(t)$ of Theorem 1 to the derivative of the first term of $H(t)$ and applying Theorem 1.

The expression (3.4) represents $T^2$ as the sum of two terms. One may be regarded mainly as a Mahalanobis distance corresponding to a weighted average of $\Sigma_1$ and $\Sigma_2$ (the weights may be close to those of Theorem 1 but will typically be different). This term essentially measures how "far" apart the means are. The second term is essentially a measure of the information contributed by the differences between the covariance matrices $\Sigma_1$ and $\Sigma_2$.

Applying the measure $T^2$ to independent observations on the variables $X$ and $Y$ of the example of Section 2 yields the following table of "separation per unit observation".

10
<table>
<thead>
<tr>
<th>$s^2/2$</th>
<th>$(X_1, X_2)$</th>
<th>$(X_1, Y_1)$</th>
<th>$(Y_1, Y_2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.25</td>
<td>0.20</td>
<td>0.25</td>
</tr>
<tr>
<td>$t^2/2$</td>
<td>0.81</td>
<td>0.71</td>
<td>0.81</td>
</tr>
</tbody>
</table>

This indicates that even when the likelihood-ratio test is used, the phenomenon described in Section 2 remains.

4. The Kullback-Leibler Information Numbers

Another measure of separation is the Kullback-Leibler information number [7], $I(F_1, F_2)$ defined by

\[(4.1) \quad I(F_1, F_2) = \int \log \frac{f_1(x)}{f_2(x)} f_1(x) dx .\]

For $k$ variate multivariate normal distributions we have

\[(4.2) \quad I(F_1, F_2) = \frac{1}{2} \left[ 8' \Sigma_2^{-1} \delta + \log |\Sigma_1^{-1} \Sigma_2| - k + \text{tr}(\Sigma_2^{-1} \Sigma_1) \right] .\]

Here $I$ measures the rate at which $\epsilon_{2n}$ approaches zero when the likelihood-ratio test is used and $\epsilon_{1n}$ is kept bounded away from zero and one. That is,

\[(4.3) \quad \lim_{n \to \infty} n^{-1} \log \epsilon_{2n} = -I(F_1, F_2) .\]

Thus $I(F_1, F_2)$ is comparable to $I = t^2/2$ of Section 3 and to $s^2/2$. Since $\epsilon_{2n}$ approaches zero more rapidly when $\epsilon_{1n}$ is bounded away
from zero than when $\varepsilon_{1n}$ and $\varepsilon_{2n}$ approach zero at the same rate, it follows that

$$(\varepsilon, \delta) \quad I(F_1, F_2) > T^2/2 > S^2/2 .$$

The Kullback-Leibler numbers are additive in the sense that the information for two independent experiments is the sum of the two informations. For the illustration of Section 2, $I_x(F_1, F_2)$, the information corresponding to $X$ is 0.877 while $I_y(F_1, F_2)$ is 4.901. From the point of view of having $\varepsilon_{2n}$ approach zero most rapidly when $\varepsilon_{1n}$ is bounded away from zero, $Y$ is more informative than $X$ and is preferred to $X$ whenever possible and not simply to attain a balance.

This information number also has an interpretation in terms of sequential experimentation. It is a measure of how well one can do using large scale sequential experiments [4]. More precisely, suppose that the cost per independent observation on $X$ is $c \to 0$, and that the choice between $F_1$ and $F_2$ is made using a Bayes sequential procedure. The risk associated with this procedure when $F_1$ is the true distribution is asymptotically equivalent to $(-c \log c)/I_x(F_1, F_2)$. Thus $I$ determines how good $X$ is for discriminating between $F_1$ and $F_2$ sequentially when $F_1$ is the true distribution. In view of this interpretation, it is not very surprising that one experiment is preferred to another and that there is no premium on mixing experiments when one is moderately sure of which is the correct hypothesis and experimentation is carried out sequentially.
References


Suppose that a statistician is permitted access to data which are more precise under $H_1$ than under $H_2$ where each hypothesis specifies a multivariate normal distribution. He is also allowed a choice between additional data more precise under $H_1$ than under $H_2$ or data in which the reverse is true. In a previous paper it was shown that if a linear discriminant function is used there is a premium on selecting the additional data to be more precise under $H_1$. In this paper this result is extended to the case where the likelihood-ratio test is used. The results involve several alternate measures for discriminating between normal multivariate distributions with unequal covariance matrices.
multivariate normal
discriminant function
Mahalanobis Distance
Kullback-Leibler Information
covariance matrix

**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. **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.

3. **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.

4. **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 classification in all capitals in parenthesis immediately following the title.

5. **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.

6. **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.

7. **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.

8. **TOTAL NUMBER OF PAGES:** Enter the total number of pages containing information.

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

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

11. **PROJECT NUMBER:** Enter the appropriate military project number, such as project number, subproject number, system numbers, task number, etc.

12. **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.

13. **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).

14. **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 DDC."
   4. "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through DDC."
   5. "All distribution of this report is controlled. Qualified DDC users shall request through DDC."

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.

15. **SUPPLEMENTARY NOTES:** Use for additional explanatory notes.

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

17. **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.

18. **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, roles, and weights is optional.