INFORMATION AND CONTROL 5, 324–346 (1962)

The Hypersphere in Pattern Recognition*

PAUL W. COOPER

Applied Research Laboratory, Sylvania Electronic Systems, Waltham, Massachusetts

Statistical classification (pattern recognition) in n-dimensional space consists in partitioning the space into category regions with decision boundaries and assigning an unknown to the category in whose region it falls. This paper demonstrates the wide utility of a particular form of decision boundary—the hypersphere—which, while especially easy to implement, is fully optimum for large classes of distributions which may arise in real problems. Of the broad spectrum of distributions described for which the hypersphere is optimum, particular interest centers on the normal, and the Pearson Type II and Type VII distributions; and methods for obtaining the boundary parameters are prescribed. Ordering of the coordinate directions according to their relative significance in contributing to the decision is examined, thereby indicating the most efficient reduction of dimensionality where this may be desired in order to allow further computational simplicity. A partial listing of error probabilities is also included.

I. INTRODUCTION

Statistical classification in n-dimensional space consists in partitioning the space into category regions with decision boundaries and assigning an unknown to the category in whose region it falls. Any real pattern recognition problem can be formulated in terms of this statistical model if the real data can be described with a set of numerically represented attributes (Cooper, 1961a). Members of the categories, both the known and the unknown, are represented as points in an n-dimensional hyperspace. The coordinates in the n-space can correspond to components of the vector describing a continuous signal in function space, or they can represent distinct measured attributes of the category members. The methods of statistical inference lead to decision procedures optimally satisfying a prescribed criterion and making best use of the known sam-

* This paper was prepared while the author was a Staff Consultant at the Melpar Applied Science Division, Watertown, Massachusetts.

ples ("learning" by statistical estimation) and any prior knowledge of the category distributions (Anderson, 1958; Cooper, 1961a; Kendall, 1960; Middleton and VanMeter, 1955).

Whereas the usual decision theoretic approach consists in attempting to estimate the best decision rule (boundary) for a particular set of distributions, this paper begins with the fundamental observation that any decision boundary is fully optimum for many distributions. With this in mind we approach the classification problem in a new way in which we select certain boundary forms which are simple to implement and demonstrate their wide applicability by describing the classes of distributions for which they are the optimum partitioning boundary. Two boundary forms—the hyperplane and the hypersphere—stand out because of their simplicity of implementation despite their wide generality. The importance of this should not be underestimated. Although optimum decision procedures generally can be formulated, at least in principle, the specific solution may not necessarily be readily determined, and even if it is it may be grossly unmanageable computationally. Practically, recourse must then be had to simpler solutions which may be good approximations to the true one. This paper is devoted to the hypersphere; a previous one dealt with the hyperplane (Cooper, 1961b).

We treat the case where there are two categories. The hypersphere decision boundary, which is implemented simply by comparing a threshold with the Euclidean distance between the unknown and a fixed point, can be an excellent approximation boundary particularly in situations where the category distributions display a spherical symmetry and differ in variance. But even more, it is the fully optimum decision boundary for a number of important classes of distributions of varied shapes and representing a broad spectrum of possible real situations. This paper introduces classes of distributions for which this is true and prescribes the optimum boundary parameters associated with them. Throughout this paper the criterion for *optimality* is that the total probability of misclassification be minimal. This is equivalent to the Bayes criterion when the category prior probabilities are equal and the misclassification costs are equal as are the costs of correct classification, and the resultant decision boundary is that contour on which the density functions for both categories have the same value. Near the end of Section III it is shown how the boundary parameters can be modified to define a hypersphere boundary which is optimum for more general criteria. Returning to our primary criterion, solutions are then presented for the multiple-category case.

In order to render a classification problem computationally more manageable one usually wishes to reduce dimensionality by eliminating those coordinate directions which are relatively insignificant in contributing toward the decision. Accordingly, Section IV indicates the relative ordering of the coordinate directions according to their importance for various spherical boundaries. Appendix II lists some error probabilities for various distributions.

II. HYPERSPHERE IMPLEMENTATION

The hypersphere in n-space having center $\mathbf{c} = (c_1, c_2, \cdots, c_n)$ and radius R_n is

$$\sum_{i=1}^{n} (x_i - c_i)^2 = R_n^2. \tag{1}$$

As a matter of convention, solely for expository convenience in this paper, we shall label categories so that the category indexed by the higher subscript is associated with the interior of the hypersphere. That is, in 2-category classification, the center of the sphere is associated with category 2, and the exterior region with category 1.

An unknown $\mathbf{x}$ is then classified as follows:

$$\text{If} \quad \sum_{i=1}^{n} (x_i - c_i)^2 \leqq R_n^2, \qquad \begin{array}{l}\text{decide on category 2,}\\ \text{otherwise, category 1.}\end{array} \tag{2}$$

In terms of a coordinate set centered at the sphere center, an unknown $\mathbf{y} = \mathbf{x} - \mathbf{c}$ is classified as follows:

$$\text{If} \quad \sum_{i=1}^{n} y_i^2 \leqq R_n^2, \qquad \begin{array}{l}\text{decide on category 2,}\\ \text{otherwise, category 1.}\end{array} \tag{3}$$

Although this paper is concerned primarily with the mathematical aspects of classification with hyperspheres, which can be implemented on a digital computer directly from the equations given, it may be of interest to note that simple analog implementation can also be feasible. By any of a number of sampling-synthesis techniques (Shannon, 1949), a function which is essentially band-limited at W and time-limited at T can be represented as a vector in a 2TW-dimensional space. Accord-

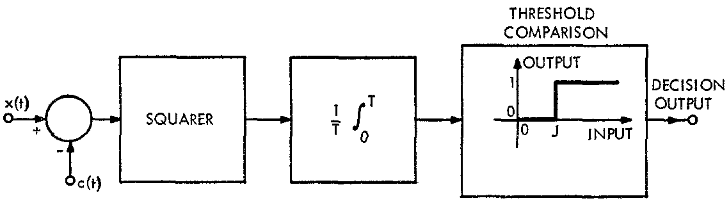

FIG. 1. Analog implementation for hypersphere decision boundary

ingly, utilizing the respective correspondence of vector **x** with time waveform $x(t)$, and **c** with $c(t)$, we can view the decision procedure for classifying the unknown $x(t)$ as follows:

$$\text{If} \quad \frac{1}{T}\int_0^T [x(t) - c(t)]^2 \, dt \leqq \frac{R_n^2}{2TW} = \frac{R_n^2}{n} = J, \tag{4}$$

decide on category 2,
otherwise, category 1.

A simple block diagram is shown in Fig. 1. If the output is 0 decide on category 2, if it is 1 decide on category 1.

III. OPTIMUM CLASSES

The hypersphere is the optimum dividing boundary for a large class of probability distribution pairs. Included in this class are many distribution pairs having varied and irregular forms for which a hypersphere constitutes the optimum boundary for a particular combination of values of the mean and the dispersion parameters only, but for which a drastically different boundary form occurs for other values of these parameters. These specialized distributions are acknowledged, but the remainder of this paper is confined to consideration of the more interesting cases where the sphere is fully optimum for all distributions of a prescribed form, regardless of specific parameter values.

A. NOTATION AND DEFINITIONS

The notation used to describe the probability density functions is herewith defined. A vector in n-space is represented with a bold-face symbol, and its components with subscripts i; e.g., $\mathbf{x} = (x_1, x_2, \cdots, x_n)$. A vector specifically associated with the kth category has subscript k, and its components have a double subscript; e.g., the mean of the kth category is $\mathbf{u}_k = (u_{k1}, u_{k2}, \cdots u_{kn})$. A scaling of the random variable is represented with ω, ($\omega > 0$), and a single subscript here references the category. A double subscript on ω is interpreted as with the vectors; the first subscript references the category, and the second the coordinate direction in n-space. Density functions (with zero mean) are represented in terms of a function $f(x)$ and constant A_n as follows:

$$\text{Univariate } p(x) = A_1\omega f(\omega x). \tag{5}$$

A radial measure in n-space is represented by r_n, where

$$r_n^2 = \sum_{i=1}^{n} x_i^2.$$

The subscript will be dropped from the r_n except where it is needed for clarity.

A spherically symmetric density function in n-space having a univariate section of the same form as (5) is

$$p(\mathbf{x}) = p(x_1, x_2, \cdots, x_n) = A_n\omega^n f(\omega r). \tag{6}$$

The surface area (Sommerville, 1958) of a hypersphere in n-space is

$$S(r_n) = S_n r^{n-1} = \frac{2\pi^{n/2}}{\Gamma(n/2)} r^{n-1}. \tag{7}$$

The radial density function corresponding to (6) is

$$p(r_n) = S(r_n)p(\mathbf{x}) = S_nA_n\omega^n r^{n-1} f(\omega r_n) = G_n\omega^n r^{n-1} f(\omega r_n). \tag{8}$$

This method of developing a multivariate distribution as an extension of a univariate one is predicated upon the resultant function in n-space being integrable. Perhaps the best way to obtain A_n is to first obtain G_n from integration of (8), and to divide the result by S_n.

Where appropriate, a second subscript on A will reference a parameter m associated with the particular functional form of $f(\omega x)$.

There will be occasion to consider certain distributions with ellipsoidal constant probability contours having principal axes parallel to the coordinate axes. For the kth category,

$$p_k(\mathbf{x}) = p_k(x_1, x_2, \cdots, x_n) = A_n\left[\prod_{i=1}^{n} \omega_{ki}\right] \cdot \left\{f\left(\left[\sum_{i=1}^{n} (\omega_{ki}\, x_i)^2\right]^{1/2}\right)\right\}. \tag{9}$$

In accordance with our convention that the interior of the spherical decision boundary be associated with the category indexed with the higher subscript, the following relationships prevail:

$$\omega_k > \omega_j, \quad \text{and} \quad \omega_{ki} > \omega_{ji}, \quad \text{for} \quad k > j.$$

B. Spherically Symmetric Normal Distributions

Of all the distributions for which the hypersphere is the fully optimum decision boundary, the one which is perhaps of greatest interest is the spherically symmetric normal distribution,

$$p(\mathbf{x}) = A_n \sigma^{-n} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - u_i)^2\right], \tag{10}$$

where

$$A_n = (2\pi)^{-n/2}, \quad \text{and} \quad \omega = 1/\sigma.$$

Two distributions differ then only in $\mathbf{u}$ and ω. Following our convention, $\omega_2 > \omega_1$, or $\sigma_2 < \sigma_1$. The center and radius of the hypersphere are, respectively,

$$\mathbf{c} = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 - \sigma_2^2}\left[\frac{\mathbf{u}_2}{\sigma_2^2} - \frac{\mathbf{u}_1}{\sigma_1^2}\right], \tag{11}$$

$$R_n^2 = 2n\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 - \sigma_2^2}\log\left(\frac{\sigma_1}{\sigma_2}\right) + \frac{\sigma_1^2\sigma_2^2}{(\sigma_1^2 - \sigma_2^2)^2}\sum_{i=1}^{n}(u_{2i} - u_{1i})^2. \tag{12}$$

When the means are the same, these equations simplify to

$$\mathbf{c} = \mathbf{u}, \quad \text{and} \quad R_n^2 = 2n\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 - \sigma_2^2}\log\left(\frac{\sigma_1}{\sigma_2}\right). \tag{13}$$

Given M_k known vectors $\{\mathbf{z}_k^{(j)}\}$, $j = 1, 2, \cdots, M_k$, from category k, the maximum-likelihood estimators for its distribution parameters are:

$$\hat{\mathbf{u}}_k = \frac{1}{M_k}\sum_{j=1}^{M_k}\mathbf{z}_k^{(j)} \tag{14}$$

$$\hat{\sigma}_k^2 = \frac{1}{n(M_k - 1)}\sum_{j=1}^{M_k}\sum_{i=1}^{n}(z_{ki}^{(j)} - \hat{u}_{ki})^2 \tag{15}$$

C. Spherically Symmetric Distributions with the Same Mean

For any two spherically symmetric distributions having the same mean, the optimum decision boundary is a plurality of hyperspheres all centered at the common distribution mean. The region associated with one of the categories then consists of the interior of the inner sphere and the space contained within alternate spherical shells; the other category is associated with the remaining regions. Classification consists simply in evaluating r_n and observing in which radial interval it falls.

For two distributions having the same functional form and differing only in the scaling ω, a condition on $f(r)$ (derived in Appendix I) ensuring that a single hypersphere constitute the decision boundary is the following: beginning with that value of r for which the derivative of $f(r)$ is not positive,

$$Q(\omega, r) = \frac{\partial f(\omega r)/\partial r}{f(\omega r)} \tag{16}$$

be a monotonically decreasing function of ω. That is,

$$\frac{\partial Q(\omega, r)}{\partial \omega} < 0 \tag{17}$$

or

$$r\left[f(r)\,\frac{d^2f(r)}{dr^2} - \left(\frac{df(r)}{dr}\right)^2\right] + f(r)\,\frac{df(r)}{dr} < 0. \tag{18}$$

Note that this condition implies that $f(r)$ itself is also monotonically decreasing. Many functional forms $f(r)$ satisfy this condition. By way of example, a general exponential class $e^{-g(r)}$, where $g(r)$ is of a form for which $e^{-g(r)}$ is integrable, satisfies (18) if

$$r\,\frac{d^2g(r)}{dr^2} + \frac{dg(r)}{dr} > 0. \tag{19}$$

Further, if $g(r) = \sum_{v=1}^{N} a_v r^v$, the condition becomes

$$\sum_{v=1}^{N} v^2 a_v r^{v-1} > 0. \tag{20}$$

Inequality (20) can be satisfied with some negative a_v, but a_N must be positive. The condition will be satisfied when all the a_v are positive. Taking single terms of the polynomial $g(r)$ we define a special subclass of exponential density functions,

$$p(\mathbf{x}) = A_{n,m}\,\omega^n e^{-a(\omega r)^m} = A_{n,m}\,\omega^n \exp\left[-a\omega^m\left(\sum_{i=1}^{n}(x_i - u_i)^2\right)^{m/2}\right] \tag{21}$$

and radial distribution

$$p(r) = A_{n,m} S_n \omega^n r^{n-1} e^{-a(\omega r)^m}, \tag{22}$$

where

$$G_{n,m} = A_{n,m}\, S_n = \frac{ma^{n/m}}{\Gamma(n/m)}, \quad \text{and} \quad m > 0, \quad \text{and} \quad a > 0.$$

Special cases of these m-exponential distributions are the normal and Laplace distributions, as are the Gamma distributions for $p(r)$ in n-space through transformation $y = r^m$. For two m-exponential distributions of the form of (22) the radius of the boundary sphere is

$$R_n = \left[\frac{n \log(\omega_2/\omega_1)}{a(\omega_2{}^m - \omega_1{}^m)}\right]^{1/m}. \tag{23}$$

Given M_k known samples of a distribution, ω_k may be estimated

$$\hat{\omega}_k{}^m = \frac{1/am}{[1/n(M_k - 1)] \sum_{j=1}^{M_k} \left[\sum_{i=1}^{n} (z_{ki}^{(j)} - \hat{u}_{ki})^2\right]^{m/2}}. \tag{24}$$

The mean $\mathbf{u}_k$ can be estimated with the sample mean as in (14). Except for the case when $m = 2$, the sample mean is not the maximum-likelihood estimator.

D. Spherically Symmetric Pearson Type II Distributions

We define the spherically symmetric Pearson Type II distributions in terms of a function $h(\mathbf{x})$,

$$h(\mathbf{x}) = A_{n,m}\, \omega^n \left[1 - \omega^2 \sum_{i=1}^{n} (x_i - u_i)^2\right]^m, \tag{25}$$

where $m > 0$, and

$$A_{n,m} = \frac{\Gamma(m + n/2 + 1)}{\pi^{n/2}\Gamma(m + 1)}.$$

The density function is defined

$$p(\mathbf{x}) = \begin{cases} h(\mathbf{x}), & \text{over region } T \\ 0, & \text{elsewhere,} \end{cases} \tag{26}$$

where region T is the interior of the hypersphere

$$\sum_{i=1}^{n} (x_i - u_i)^2 = \frac{1}{\omega^2}.$$

The form of this distribution is dependent upon the parameter m as follows: for m equal to 0 the distribution is uniform, for $m = \frac{1}{2}$ it is an inverted hypersemiellipsoid, for $m = 1$ it is an inverted hyperparaboloid, as m approaches infinity it becomes a normal distribution (and a delta function when the other parameters are held fixed), all of which have spherical symmetry over the region T in n-space. The hypersphere decision boundary has parameters

$$\mathbf{c} = \frac{\mathbf{u}_2 \omega_2^{(n/m)+2} - \mathbf{u}_1 \omega_1^{(n/m)+2}}{\omega_2^{(n/m)+2} - \omega_1^{(n/m)+2}}, \tag{27}$$

$$R_n^2 = \frac{\omega_2^{n/m} - \omega_1^{n/m}}{\omega_2^{(n/m)+2} - \omega_1^{(n/m)+2}} + \frac{(\omega_1 \omega_2)^{(n/m)+2}}{(\omega_2^{(n/m)+2} - \omega_1^{(n/m)+2})^2} \sum_{i=1}^{n} (u_{2i} - u_{1i})^2. \tag{28}$$

If region T_2 is not fully contained within T_1, then the boundary of (27) and (28) is not unique in the region complementary to both T_1 and T_2. This prescribed boundary represents the intersection of two infinite paraboloids. For these distributions, in the region outside of both T_1 and T_2, any boundary is therefore satisfactory, since an unknown vector could not occur there anyway. See Fig. 2.

Maximum-likelihood estimators here are cumbersome, and it is easier to estimate the parameters in terms of the sample moments. In terms of the second central radial sample moment ω may be estimated

$$\hat{\omega}_k^2 = \left(\frac{n}{2m + 2 + n}\right)\left[\frac{1}{(M_k - 1)} \sum_{j=1}^{M_k} \sum_{i=1}^{n} (z_{ki}^{(j)} - \hat{u}_{ki})^2\right]^{-1}, \tag{29}$$

where $\mathbf{u}_k$ is determined as the sample mean.

E. Spherically Symmetric Pearson Type VII Distributions

The spherically symmetric pearson Type VII distributions are of the form

$$p(\mathbf{x}) = A_{n,m}\, \omega^n \left[1 + \omega^2 \sum_{i=1}^{n} (x_i - u_i)^2\right]^{-m}, \tag{30}$$

where $2m > n$, and

$$A_{n,m} = \frac{\Gamma(m)}{\pi^{n/2}\Gamma(m - n/2)}.$$

(Note that for half integer values of m this is equivalent to the t-distribution, a special case of which is the univariate Cauchy distribution, for which $m = 1$).

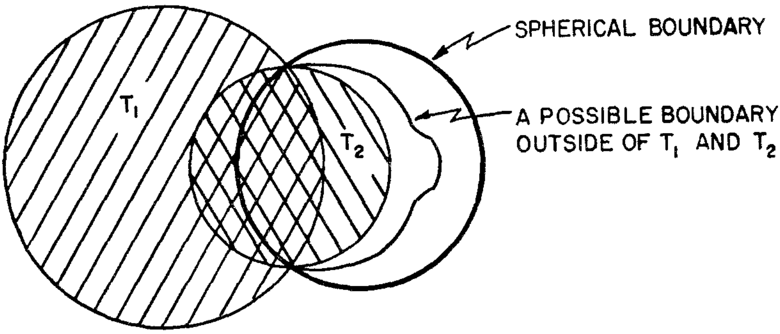

FIG. 2. Regions T_1 and T_2 shown in 2-space

The hypersphere decision boundary for two such distributions has parameters

$$\mathbf{c} = \frac{\mathbf{u}_2\, \omega_2^{2-(n/m)} - \mathbf{u}_1\, \omega_1^{2-(n/m)}}{\omega_2^{2-(n/m)} - \omega_1^{2-(n/m)}}, \tag{31}$$

$$R_n^{\ 2} = \frac{\omega_1^{-(n/m)} - \omega_2^{-(n/m)}}{\omega_2^{2-(n/m)} - \omega_1^{2-(n/m)}} + \frac{(\omega_1\, \omega_2)^{2-(n/m)}}{\left(\omega_2^{2-(n/m)} - \omega_1^{2-(n/m)}\right)^2} \sum_{i=1}^{n} (u_{2i} - u_{1i})^2. \tag{32}$$

The qth central radial moment, which exists only if $2m > n + q$, is

$$\overline{r^q} = \frac{1}{\omega^q} \frac{\Gamma[\frac{1}{2}(q + n)]\, \Gamma\, [m - \frac{1}{2}(q + n)]}{\Gamma(\frac{1}{2}n)\Gamma(m - \frac{1}{2}n)} \tag{33}$$

The parameter ω can be estimated from one of the central radial sample moments, preferably a low one because of both existence and convergence properties. If $2m > n + 2$ so that both the first and second moments exist and the first sample moment converges with increased sample size, an estimate of ω in terms of the first central radial sample moment is

$$\hat{\omega}_k = \frac{\Gamma(\frac{1}{2}n)\Gamma(m - \frac{1}{2}n)}{\Gamma[\frac{1}{2}(n + 1)]\, \Gamma\, [m - \frac{1}{2}(n + 1)]} \left[\frac{1}{(M_k - 1)} \sum_{j=1}^{M_k} \cdot \left(\sum_{i=1}^{n} (z_{ki}^{(j)} - \hat{u}_{ki})^2\right)^{1/2}\right]^{-1}, \tag{34}$$

where $\mathbf{u}_k$ is the sample mean. For m in the range $1 < (2m - n) \leqq 2$, wherein the first radial moment exists but the second does not, the sample mean, which then does not converge, is replaced by the sample median; i.e., $\mathbf{u}_k$ is then taken as the sample median. When m is sufficiently large to ensure satisfactory convergence of the second radial sample moment, ω can be determined through use of (33) for $q = 2$: $\omega^2\overline{r^2} = n/(2m - n - 2)$.

F. Ellipsoidal Distributions

Under certain conditions a hypersphere decision boundary arises for distribution pairs similar to some of those previously discussed but having ellipsoidal constant probability contours and general functional form as shown in Eq. (9). The principal axes of the ellipsoids of both distributions of a pair are identically oriented, and for convenience here these directions will be taken coincident with the coordinate directions. The quantity Δ is a parameter having any value greater than zero which, given the scaling set $\{\omega_{1i}\}$, $i = 1, 2, \cdots, n$, prescribes the values for the members of the set $\{\omega_{2i}\}$ to ensure that the decision boundary be a hypersphere.

Noting that the spherical distributions are included as a special case when $\omega_{1i} = \omega_{1j}$ for $i, j = 1, 2, \cdots, n$, we list some of these ellipsoidal distributions with the associated hypersphere boundary parameters. The constants A_n or $A_{n,m}$ are identical to the corresponding ones listed earlier for the spherical distributions.

Normal Distribution

$$p_k(\mathbf{x}) = A_n \left[\prod_{i=1}^{n} \sigma_{ki}\right]^{-1} \exp\left[-\frac{1}{2}\sum_{i=1}^{n} \frac{1}{\sigma_{ki}^2}(x_i - u_{ki})^2\right]. \qquad (35)$$

The boundary is a sphere if $\Delta = 1/\sigma_{2i}^2 - 1/\sigma_{1i}^2 = \omega_{2i}^2 - \omega_{1i}^2$ for all i. Then $\sigma_{2i}^2 = \sigma_{1i}^2\,(1 + \Delta\sigma_{1i}^2)^{-1}$, and the boundary parameters are

$$c_i = \frac{u_{2i}/\sigma_{2i}^2 - u_{1i}/\sigma_{1i}^2}{\Delta}, \qquad (36)$$

$$R_n^{\,2} = \frac{1}{\Delta}\sum_{i=1}^{n}\left\{\log\,(1 + \Delta\sigma_{1i}^2) + [\Delta(\sigma_{1i}\,\sigma_{2i})^2]^{-1}(u_{2i} - u_{1i})^2\right\}. \qquad (37)$$

Pearson Type II Distributions

$$p_k(\mathbf{x}) = A_{n,m}\left[\prod_{i=1}^{n} \omega_{ki}\right]\left[1 - \sum_{i=1}^{n} \omega_{ki}^2(x_i - u_{ki})^2\right]^m. \qquad (37)$$

Taking $Q_k = \prod_{i=1}^{m} \omega_{ki}$, the condition is

$$\Delta = Q_2^{1/m}\omega_{2i}^2 - Q_1^{1/m}\omega_{1i}^2,$$

or

$$\omega_{2i}^2 = Q_2^{-1/m}(\Delta + Q_1^{1/m}\omega_{1i}^2).$$

Then

$$c_i = \frac{Q_2^{1/m}\,\omega_{2i}^2\,u_{2i} - Q_1^{1/m}\,\omega_{1i}^2\,u_{1i}}{\Delta}, \tag{39}$$

$$R_n^{\,2} = \frac{Q_2^{1/m} - Q_1^{1/m}}{\Delta} + \frac{(Q_1\,Q_2)^{1/m}}{\Delta^2}\sum_{i=1}^{n}(\omega_{1i}\,\omega_{2i})^2(u_{2i} - u_{1i})^2. \tag{40}$$

Pearson Type VII Distributions

$$p_k(\mathbf{x}) = A_{n,m}\left[\prod_{i=1}^{n}\omega_{ki}\right]\left[1 + \sum_{i=1}^{n}\omega_{ki}^2(x_i - u_{ki})^2\right]^{-m}. \tag{41}$$

The condition is

$$\Delta = Q_2^{-1/m}\omega_{2i}^2 - Q_1^{-1/m}\omega_{1i}^2,$$

or

$$\omega_{2i}^2 = Q_2^{1/m}(\Delta + Q_1^{-1/m}\omega_{1i}^2).$$

Then

$$c_i = \frac{Q_2^{-1/m}\omega_{2i}^2 u_{2i} - Q_1^{-1/m}\omega_{1i}^2 u_{1i}}{\Delta}. \tag{42}$$

$$R_n^{\,2} = \frac{Q_1^{-1/m} - Q_2^{-1/m}}{\Delta} + \frac{(Q_1Q_2)^{-1/m}}{\Delta^2}\sum_{i=1}^{n}(\omega_{1i}\omega_{2i})^2(u_{2i} - u_{1i})^2. \tag{43}$$

G. Other Decision Criteria

Although the decision boundaries described have been obtained for the criterion that the total probability of misclassification be minimal, simple modification can account for a wider range of criteria wherein the likelihood ratio is compared with a threshold. As is well known, Bayes, Minimax, and Neyman-Pearson decision criteria are satisfied by the decision rule:

$$\text{If} \quad \frac{p_2(\mathbf{x})}{p_1(\mathbf{x})} \gtreqless L, \quad \begin{array}{l}\text{decide on category 2,}\\ \text{otherwise, category 1,}\end{array} \tag{44}$$

where the threshold L is determined from the particular criterion chosen and the associated specifications; e.g. the Bayes criterion with specified prior probabilities and misclassification costs. The minimal probability criterion corresponds to an L of unity. For the distributions here considered introduction of L does not alter the spherical form of the boundary, but only its parameters. In observing the behavior of the spherical boundary as a function of L it is apparent that the sphere shrinks with increasing L and eventually degenerates to a point. For this point the value of L is denoted L_m, where $L_m > 1$. For any values of L greater than L_m the sphere does not exist, the expressions for the radius are not real, and an unknown $\mathbf{x}$ is then always assigned to category 1 (by our convention, the region of category 2 having disappeared with the sphere). The classification procedure then becomes trivial and the optimal solution then calls for no measurements on the unknown but merely depends upon a prescribed a priori decision that all points be assigned to category 1. It is instructive to note that as L approaches L_m, the center of the sphere approaches the point at which $p_2(\mathbf{x})/p_1(\mathbf{x})$ attains its maximum value.

Modifications in the boundary, determined by taking the equality in (44), follow. For the normal distribution add to the expression for R^2_n the term $[-2\sigma_1^2\sigma_2^2(\sigma_1^2 - \sigma_2^2)^{-1} \log L]$ in Eq. (12) and $[-(2/\Delta) \log L]$ in Eq. (37). We observe that the sphere center is here not dependent upon L. By setting $R_n^2 = 0$, L_m is obtained. For the m-exponentials add $\log L$ to the numerator within the bracket in Eq. (23). There $L_m = (\omega_2/\omega_1)^n$. For the ellipsoidal Type II and Type VII distributions, replace Q_1 with LQ_1, noting that these also enter in the expression for Δ. Equations thereby affected are (39), (40), (42), and (43). For the spherically symmetric Type II and Type VII distributions replace ω_1 by $(L^{1/n}\omega_1)$ in all places where ω_1 is explicitly raised to the (n/m) or $(-n/m)$ power, but not elsewhere; e.g., $\omega_1^{(n/m)+2} = \omega_1^{(n/m)}\omega_1^2$ becomes $(L^{1/n}\omega_1)^{(n/m)}\omega_1^2 = L^{1/m}\omega_1^{(n/m)+2}$. Equations thereby affected are (27), (28), (31), and (32).

H. Other Distributions

Some other classes of distributions for which the hypersphere is an optimum boundary deserve mention. For any distribution pair giving rise to a hyperellipsoid decision boundary, a linear transformation of the hyperellipsoid into a hypersphere serves to generate a distribution pair for which the optimum boundary is a hypersphere. Actually, the transformation need not be linear provided that the interior and exterior regions of the ellipsoid map, respectively, into the interior and exterior of the sphere.

There are a number of geometric surfaces for which the density functions having these forms intersect in a hyperellipsoid, and therefore in a hypersphere for appropriate relationships among the parameters. As already discussed, two of these are the Type II for $m = \frac{1}{2}$ and $m = 1$, which are, respectively, in the form of an $(n + 1)$-dimensional semiellipsoid and a paraboloid. These density functions are, of course, $(n + 1)$-dimensional figures defined on an n-dimensional space, and the intersecting hypersphere boundary is an n-sphere. Two density functions each in the form of a hyperellipsoidal segment, less than a semihyperellipsoid, also intersect in a hyperellipsoid, so, for example, do two density functions one of which is in the form of an ellipsoidal hypercone radiating downward from its vertex and the other in the form of a skew hyperplane of appropriate orientation. A spherical boundary also arises when one density function is uniform over a hypersphere and over which it is everywhere higher than the second density function, which can have any functional form. A uniform distribution defined over a region containing a subregion over which is defined a second density function which is spherically symmetric and bounded also produces a spherical boundary, or boundaries.

Two density functions which are identically distributed in a k-dimensional subspace and which in the remaining $(n - k)$-dimensional subspace satisfy any of the previously described conditions for intersecting in a spherical boundary give rise to a decision boundary which is a hypersphere in the $(n - k)$-space, and which in n-space is a spherical infinite hypercylinder having an $(n - k)$-dimensional base. (Of course all of the discriminatory information is contained within the $(n - k)$-subspace).

Disjoint sets can be perfectly (error-free) separated by a boundary of appropriate form. In some cases any of a number of boundary forms are perfect; for example, the hypersphere and hyperplane are both perfect for separating two disjoint convex sets. However, there are many pairs of nonconvex disjoint sets which are perfectly separable by a sphere but not by a plane.

These examples cited in this section are samples of the various forms of distributions giving rise to a hypersphere boundary. The cases listed in the last few paragraphs do not exhaust the field and are included only to be representative of some of the more specialized forms. Some of them were described in terms of an ellipsoidal boundary, but in all cases the relationships among the parameters required to ensure that the boundary be a sphere are easily determined, and will not be further dwelt upon here.

IV. SIGNIFICANT COORDINATE DIRECTIONS

For these cases where the hypersphere is the optimum decision boundary, the ordering of the coordinate directions according to their relative importance in contributing to the decision is relatively simple. For spherically symmetric distributions having the same mean, no direction is more important than any other one. In choosing the v best coordinate direction for the purpose of reducing the dismensionality, all sets of v coordinate directions are equally effective from a mathematical point of view, (and, therefore, practical considerations pertaining to the relative difficulties in measuring various attributes could then dictate the choice of coordinate directions). Of course an n-space representation is here better than a v-space one, where $v < n$.

For ellipsoidal distributions having the same mean, the coordinate directions are ordered in decreasing importance as $\{\sigma_{1i}/\sigma_{2i}\}$ or $\{\omega_{2i}/\omega_{1i}\}$ is ordered in decreasing magnitude, where $i = 1, 2, \cdots, n$, corresponding to the n coordinate directions. Making use of the relationships between ω_{1i}, ω_{2i}, and Δ, this is equivalent to ordering the coordinate directions in decreasing importance as the set $\{\sigma_{1i}\}$ is also ordered in decreasing magnitude, or $\{\omega_{1i}\}$ in increasing magnitude.

For the spherically symmetric distributions—normal, Type II, and Type VII—having different means, the one most important direction is the one connecting the two distribution means. Taking any orthonormal basis in the $(n - 1)$-dimensional hyperspace perpendicular to the line of means, all $(n - 1)$ directions defined by the vectors in that basis are equally important, but less so than the aforementioned principal direction. Given a prescribed orthonormal basis in n-space $\{\mathbf{e}_i\}$, $i = 1, 2, \cdots, n$, these coordinate directions are ordered in decreasing importance according as the magnitude of their direction cosines with the line of means are ordered in decreasing magnitude; i.e., as the set

$$\{ \mid \mathbf{e}_i \cdot (\mathbf{u}_2 - \mathbf{u}_1) \mid \}$$

is ordered in decreasing value. When the mean vectors are expressed in terms of this basis, the bracket reduces to $\{ \mid u_{2i} - u_{1i} \mid \}$.

V. MULTIPLE-CATEGORY

Discriminating among K categories all of which pairwise satisfy the conditions of Section III can be achieved with spherical boundaries between each pair of categories, and these hyperspheres can each be determined as in the 2-category classification problem. Such a procedure requires $\frac{1}{2}K(K - 1)$ hyperspheres. However, only K discrimination functions are required if distance is measured directly to the mean of each category, instead. Accordingly, the unknown $\mathbf{x}$ is assigned to that category k for which $F_k(\mathbf{x})$ is least or greatest, depending upon the distribution form. Choose the greatest $F_k(\mathbf{x})$ for the Type II distributions, and the least for the remaining distributions.

$$\text{Spherical normal: } F_k(\mathbf{x}) = \frac{1}{\sigma_k^2} \sum_{i=1}^{n} (x_i - u_{ki})^2 + 2n \log \sigma_k \tag{45}$$

$$\text{Ellipsoidal normal: } F_k(\mathbf{x}) = \sum_{i=1}^{n} \left\{ \frac{1}{\sigma_{ki}^2} (x_i - u_{ki})^2 + 2 \log \sigma_{ki} \right\} \tag{46}$$

$$\text{Spherical Type II: } F_k(\mathbf{x}) = \omega_k^{n/m} \left[1 - \omega_k^2 \sum_{i=1}^{n} (x_i - u_{ki})^2 \right] \tag{47}$$

$$\text{Ellipsoidal Type II: } F_k(\mathbf{x}) = \left[\prod_{i=1}^{n} \omega_{ki} \right]^{1/m} \left[1 - \sum_{i=1}^{n} \omega_{ki}^2 (x_i - u_{ki})^2 \right] \tag{48}$$

$$\text{Spherical Type VII: } F_k(\mathbf{x}) = \omega_k^{-n/m} \left[1 + \omega_k^2 \sum_{i=1}^{n} (x_i - u_{ki})^2 \right] \tag{49}$$

$$\text{Ellipsoidal Type VII: } F_k(\mathbf{x}) = \left[\prod_{i=1}^{n} \omega_{ki} \right]^{-1/m} \left[1 + \sum_{i=1}^{n} \omega_{ki}^2 (x_i - u_{ki})^2 \right] \tag{50}$$

$$\text{Spherical } m\text{-exponential: } F_k(\mathbf{x}) = \omega_k^m \left[\sum_{i=1}^{n} (x_i - u_{ki}^2) \right]^{m/2} - (n/a) \log \omega_k \tag{51}$$

Equation (51) is suitable for use with any set of K spherical m-exponential distributions (defined in Eq. (21)) having different means. However, except for $m = 2$, the boundary between any two categories with spherical m-exponential distributions is a sphere only when the two means are identical.

For any K distributions having an identical mean and the same distribution form which satisfies inequality (18), the space is divided with $(K - 1)$ concentric hyperspheres into K regions consisting of $(K - 2)$ spherical shells containing an interior sphere and all contained within an infinite exterior region. Beginning in the outer region and moving toward the center, these regions are associated with the k categories with increasing index k. That is, labeling the outermost hypersphere boundary "1," and the innermost "$K - 1$," an unknown $\mathbf{x}$ is classified into the kth category if $R_n^{(k-1)} \geqq r_n > R_n^{(k)}$, where $R_n^{(0)} = \infty$ and $R_n^{(K)} = 0$, and $r_n^2 = \sum_{i=1}^{n} (x_i - u_i)^2$.

VI. APPROXIMATIONS

There is a wide range of situations for which the hypersphere serves as a good approximation to the actual optimum boundary; for example, where the actual distributions are not much different from the ones described in Section III. Of course, when the categories are widely separated and relatively tightly clustered, the hypersphere (or the hyperplane or almost any dividing boundary) is an excellent decision boundary . . .) the degradation in its performance from the optimum can be negligible.

When two categories are each described by unimodal distributions of similar functional form which display spherical symmetry, the hypersphere can generally be a good approximate decision boundary. The spherical boundary is especially useful when the variances of the two distributions differ significantly; although when they are of comparable value the boundary sphere becomes very large and can be replaced with a hyperplane, particularly in the limit when the variances are equal. When one category is tightly clustered relative to the other one, a hypersphere decision boundary about the tighter distribution can generally be good, regardless of the actual forms of the two distributions.

When the distributions are spherically symmetric having one of the forms described in Section III giving rise to a spherical boundary, but where it is not known specifically which form is correct, it is tempting to assume normality and use the boundary sphere parameters prescribed in Eqs. (11)–(13). These equations can be expressed in terms of the variance σ^2 for the Type VII, the Type II, or the m-exponentials, for each of which the variance is, respectively,

$$\begin{gathered}\sigma^2 = 1/\omega^2(2m - n - 2), \qquad \sigma^2 = 1/\omega^2\,(2m + n + 2), \\ \sigma^2 = (1/n\omega^2\, a^{2/m})\,\frac{\Gamma(n+2)/m}{\Gamma(n/m)}\,.\end{gathered} \tag{52}$$

The degradation incurred in using the boundary obtained under the assumption of normality from using the appropriate boundary for the true distribution can then be calculated with the error probability expressions of Appendix II. This comparison is relatively easy to make for the special case when the means of the two category distributions are identical. (We might note that for the m-exponentials this comparison is of interest only when the two means are the same.)

VII. CONCLUDING REMARKS

In reviewing distribution pairs for which a hypersphere is the optimum boundary, we observe that as the two distributions become comparable in variance the hypersphere boundary grows and becomes a hyperplane in the limit. The sphere is most useful when the two category distributions have variances which differ. When the variances are close, classification can be more easily implemented by comparison of weighted direct distance measurements from the unknown to each category mean, as prescribed in Section V.

By appropriate linear transformations, the cases where spherical boundaries prevail can be transformed into ellipsoidal ones, and vice versa. Therefore, any of the results developed herein can be applied through linear transformation to distributions leading to ellipsoidal boundaries.

Relative orderings of coordinate directions according to their importance may perhaps not be of as major concern with spherical boundaries as with other boundary forms because of the relative ease with which spherical boundaries can be implemented in a space of any dimension (the complexity is linearly proportional to dimension), and therefore there is less need for dimensionality reduction.

Because of the paucity of parameters required to define the spherical distributions, estimation ("learning") of parameters for them can begin with only two known samples from each category.

APPENDIX I. DERIVATION FOR SINGLE BOUNDARY

To prove conditions (16) to (18):

In n-space, we have two radial distributions of the same functional form:

$$p_1(r) = r^{n-1}f(r)$$

and

$$p_2(r) = \omega p_1(\omega r) = \omega^n r^{n-1} f(\omega r). \tag{53}$$

In the range from $r = 0$ to the lowest value of r for which $f(r)$ first becomes a nonincreasing function of r, there are no restrictions on $f(r)$. Beginning by taking $\omega > 1$, for which $p_2(r) > p_1(r)$ for $0 \leqq r < r_0$, where r_0 is the lowest value of r for which the two density functions intersect, we seek a condition that for $r > r_0$, $p_2(r) < p_1(r)$. To require that $f(r)$ be monotonically decreasing is not sufficient, whereas the condition that $df(r)/dr$ be a monotonically decreasing function of r is too strict. Beginning with r_0, our fundamental requirement is satisfied if

$$\frac{dp_2(r)/dr}{p_2(r)} < \frac{dp_1(r)/dr}{p_1(r)}, \tag{54}$$

which falls in strictness between the two possible conditions of the previous sentence. Note that (54) allows $dp_2(r)/dr$ to be greater than $dp_1(r)/dr$, i.e., it allows $|dp_2(r)/dr| < |dp_1(r)/dr|$, the more so when $p_2(r)$ is much less than $p_1(r)$. Substitution of (53) in (54) gives

$$Q(\omega, r) = \frac{\partial f(\omega r)/\partial r}{f(\omega r)} < \frac{\partial f(r)/\partial r}{f(r)}. \tag{55}$$

Inequality (55) is good for $\omega < 1$ also, since the roles of $p_2(r)$ and $p_1(r)$ could have been reversed in (53). Inequality (55) is equivalent to requiring that $Q(\omega, r)$ be a monotonically decreasing function of ω, i.e., $\partial Q(\omega, r)/\partial\omega < 0$, which upon expansion and with substitution of r for ωr leads to inequality (18).

APPENDIX II. ERROR PROBABILITIES

The error probabilities associated with certain of the spherically symmetric distributions are expressed relatively easily in closed form or in terms of tabulated functions; particularly the Incomplete Gamma and Beta Functions and the normal probability integral. Although greater interest centers on the case where the means of the two category distributions differ, we here list primarily the error probabilities when the means are the same, since it is the latter case which lends itself to simple expression in closed form. In 2-category classification the total probability of misclassification is the sum of the probabilities that a member of category 1 falls within the decision hypersphere and a member of category 2 falls without the sphere. We denote with $P_n(R)$ the probability that a member of a category having a spherically symmetric distribution in n-space will fall within a hypersphere of radius R centered at the distribution mean. That is, taking $p(r)$ as the radial density function,

$$P_n(R) = \int_0^R p(r)\,dr. \tag{56}$$

Normal Distribution

In terms of the Error Integral (N. B. S., 1953),

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} e^{-(1/2)t^2}\,dt, \tag{57}$$

$$P_n(R) = 1 - e^{-(1/2)(R/\sigma)^2} \sum_{j=0}^{\frac{n}{2}-1} \frac{(R^2/2\sigma^2)^j}{\Gamma(j+1)}, \qquad \text{for } n \text{ even}, \tag{58}$$

$$P_n(R) = \phi(R/\sigma) - e^{-(1/2)(R/\sigma)^2} \sum_{j=1}^{\frac{n-1}{2}} \frac{(R^2/2\sigma^2)^{j-\frac{1}{2}}}{\Gamma(j+\frac{1}{2})}, \qquad \text{for } n \text{ odd}, \tag{59}$$

where the summation is taken to be zero for $n = 1$.

m-Exponentials

In terms of Eq. (22) and the Incomplete Gamma Function (Bateman, 1953; Pearson, 1922),

$$\gamma(\alpha, x) = \int_0^x r^{\alpha-1} e^{-r}\, dr, \qquad \text{for } \alpha > 0; \tag{60}$$

$$P_n(R) = 1/\Gamma(n/m)\gamma(n/m, a\omega^m R^m), \tag{61}$$

Utilizing the recurrence relation,

$$\gamma(\alpha, x) = (\alpha - 1)\gamma(\alpha - 1, x) - x^{\alpha-1}e^{-x}, \tag{62}$$

and the fact that

$$\gamma(\tfrac{1}{2}, x) = \sqrt{\pi}\phi(\sqrt{2x}),$$

we can express (60) and (61) in either closed form or in terms of the Error Integral for integer and odd half integer values of α, respectively:

For $\alpha = 1, 2, 3, 4, \cdots$;

$$\gamma(\alpha, x) = \Gamma(\alpha)\left\{1 - e^{-x}\sum_{j=0}^{\alpha-1} \frac{x^j}{\Gamma(j+1)}\right\} \tag{64}$$

For $\alpha = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \cdots$;

$$\gamma(\alpha, x) = \Gamma(\alpha)\left\{\phi(\sqrt{2x}) - e^{-x}\sum_{j=1}^{\alpha-\frac{1}{2}} \frac{x^{j-\frac{1}{2}}}{\Gamma(j+\frac{1}{2})}\right\}, \tag{65}$$

where the summation is taken to be zero for $\alpha = \frac{1}{2}$.

Type II

In terms of Eqs. (25) and (26) and the Incomplete Beta Function (Pearson, 1934; Pearson and Hartley, 1958),

$$B_x(\alpha, \beta) = \int_0^x r^{\alpha-1}(1 - r)^{\beta-1}\, dr, \qquad x \leqq 1; \tag{66}$$

$$P_n(R) = \frac{\Gamma(m + \frac{1}{2}n + 1)}{\Gamma(\frac{1}{2}n)\Gamma(m + 1)} B_x(n/2, m + 1), \qquad \text{where } x = \omega^2 R^2. \tag{67}$$

For integer values of β, Eq. (66) can be simply evaluated through direct integration. A useful recurrence relation is

$$B_x(\alpha, \beta) = x^\alpha(1-x)^{\beta-1}(1/\alpha) + (1/\alpha)(\beta - 1)B_x(\alpha + 1, \beta - 1). \quad (68)$$

Type VII

In terms of Eq. (30) and

$$I_x(n, m) = \frac{2\Gamma(m)}{\Gamma(n/2)} \int_0^x \frac{r^{n-1}\,dr}{(1+r^2)^m}; \quad (69)$$

$$P_n(R) = [1/\Gamma(m - \tfrac{1}{2}n)]I_{\omega R}(n, m). \quad (70)$$

Utilizing the recurrence relation

$$I_x(n, m) = I_x(n-2, m-1) - \frac{x^{n-2}\Gamma(m-1)}{(1+x^2)^{n-1}\Gamma(n/2)}, \quad (71)$$

we obtain through its repeated use,

$$I_x(n, m) = \left\{ \frac{\Gamma(m - \frac{1}{2}n + W)}{\Gamma(1+W)}\psi - \sum_{j=1}^{(n/2)-W} \frac{x^{n-2j}}{(1+x^2)^{m-j}} \frac{\Gamma(m-j)}{\Gamma(\frac{1}{2}n + 1 - j)} \right\}, \quad (72)$$

where the summation is taken to be zero for $n = 1$, where

$$W = \begin{cases} 0, & \text{for } n \text{ even,} \\ \frac{1}{2}, & \text{for } n \text{ odd} \end{cases}$$

and

$$\psi = \begin{cases} 1, & \text{for } n \text{ even} \\ \dfrac{\sqrt{\pi}\Gamma(m - \frac{1}{2}n)}{\Gamma[m - \frac{1}{2}(n-1)]}\left[P(x\sqrt{2m-n}; 2m-n) - \dfrac{1}{2}\right], & \text{for } n \text{ odd,} \end{cases}$$

and

$$P(t; v) = \frac{\Gamma[(v+1)/2]}{\sqrt{v\pi}\Gamma(v/2)} \int_{-\infty}^t \frac{du}{(1+u^2/v)^{(v+1)/2}}, \quad (73)$$

where $P(t; v)$ is tabulated (Pearson and Hartley, 1958) as the univariate t-distribution for half integer values of v. Intermediate values can be obtained through interpolation.

Normal Distributions with Different Means

Consider a spherically symmetric normal distribution in n-space with mean displaced from the n-sphere center by a distance d. The probability

$\mathcal{P}_n(R)$ of falling within the sphere is then

$$\mathcal{P}_n(R) = \frac{1}{\sigma\sqrt{2\pi}} \int_{d-R}^{d+R} e^{-(1/2)(x/\sigma)^2} P_{n-1}(\rho(x))\, dx, \tag{74}$$

where

$$\rho(x) = [R^2 - (x - d)^2]^{-1/2},$$

and $P_{n-1}(y)$ is obtained from Eqs. (58) or (59). Equation (74) arises from the fact that, taking x in the direction defined by the mean and the sphere center, the marginal probability distribution indicating the probability of falling within the n-sphere for a particular value of x is simply the probability of falling within an $(n - 1)$-sphere of radius $\rho(x)$; and, the normal distribution is independently distributed in the x direction from the $(n - 1)$-space perpendicular to this x-direction.

Approximation for Normal Distribution

A fairly good lower bound for $P_n(R)$ for the spherically symmetric normal distribution is obtained by evaluating the integral of the density function over a hypercube centered at the mean and having a volume equal to the volume of the hypersphere of radius R. Taking a cube edge to be $2b$, the equal volume condition dictates that $b = (\frac{1}{2})R\sqrt{\pi} \cdot[\Gamma(\frac{1}{2}n + 1)]^{-1/n}$, and

$$P_n(R) > [\phi(b/\sigma)]^n \tag{75}$$

Received: July 2, 1962

References

Anderson, T. W., (1958), "An Introduction to Multivariate Statistical Analysis." Wiley, New York.

Bateman, H., (1953), "Higher Transcendental Functions," Vol. II. McGraw Hill, New York.

Cooper, P. W., (1961a), Classification by statistical methods (pattern recognition). Melpar Technical Note 61/2; Melpar Applied Science Division, Watertown, Mass.

Cooper, P. W., (1961b), The hyperplane in pattern recognition. Melpar Technical Note 61/6; Melpar Applied Science Division, Watertown, Mass.

Kendall, M. G., (1960), "The Advanced Theory of Statistics," Vols. I and II. Hafner, New York.

Middleton, D., and VanMeter, D., (1955 and 1956), Detection and extraction of signals in noise from the point of view of statistical decision theory, Parts I and II, *J. Soc. Indust. and Appl. Math., Vol.* **3**, 192–253; **4**, 86–1.19

National Bureau of Standards, (1953), "Tables of Normal Probability Functions."

Applied Mathematics Series 23, U. S. Gov't Printing Office, Washington, D. C.

PEARSON, E. S., AND HARTLEY, H. O., eds. (1958), "Biometrika Tables for Statisticians, Vol. I, Cambridge Univ. Press.

PEARSON, K., ed. (1922), "Tables of the Incomplete Gamma Function." Cambridge Univ. Press.

PEARSON, K., ed. (1934), "Tables of the Incomplete Beta Function." Cambridge Univ. Press.

SHANNON, C., (1949), Communications in the presence of noise. *Proc. IRE*, **37**, 10–21.

SOMMERVILLE, D. M. Y., (1958), "An Introduction to the Geometry of N Dimensions." Dover, New York.