6-1-1975

# Statistical Methods for Determining Quality of Communication Channels

Scott Guthery

Andres C. Salazar

# Statistical Methods for Testing Quality of Communication Channels

S. B. GUTHERY AND ANDRES C. SALAZAR, MEMBER, IEEE

*Abstract*—**The reliable transmission of information over communication channels depends greatly on the adequacy of the maintenance program provided for the transmission medium. An integral part of such a program is the procedure used for determining the quality of a communication link. In this concise paper, we discuss several statistical methods for ascertaining those impairments—and their respective levels—which reveal the most information about the fidelity of the channel in signal transmission. For the purpose of illustration, an example is presented which demonstrates how a testing program could implement each particular statistical method.**

## I. INTRODUCTION

The transmission medium of a communication system may consist of a variety of links—wire pairs, repeaters, multiplexers, switching equipment, microwave radio paths, etc. Thus, a communication channel or connection set up for the transmission of information between two points in the system may have to be evaluated according to a *variety* of impairments such as background noise, nonlinear distortion, impulse noise, envelope delay distortion, etc. Each impairment, of course, affects the quality of the channel differently according to the type of modulation technique being used.

To delineate clearly the objectives of a maintenance program for communication channels, we must first define what is meant by channel quality. Further, we must be told how the channel can be observed. That is, what equipment is available and how many *measurements* are possible. It is then the objective of the maintenance program engineer to establish the limits on the channel characteristics which permit adequate transmission. Further, he must specify what order the measurements should follow in order to minimize the effort required for ascertaining quality. Perhaps most important of all is the accuracy which the maintenance engineer feels his testing program has in ferreting out transmission problems.

In this concise paper, we will delve into many aspects of the construction of a channel testing program for maintenance purposes. The heart of the program consists of discovering relationships between channel fidelity and channel measurements. Analytical efforts often fail to lead to useful relationships simply because the channel models used are simplistic in nature and often ignore a large number of possible impairments. In addition, mathematical models often do not deal with interrelationships between various channel impariments, and when they do, the mathematical problems which evolve are often intractable.

A useful approach to determining the relevant and statistically significant relationships between channel quality and measurements is through the statistical analysis of a large data base linking the two aspects of a communications system—performance and impairment levels.

There is a myriad of pattern recognition techniques or feature selection algorithms which could be used to advantage in such a statistical analysis. However, the objective here is to construct a channel testing program which requires little or no computation by a craftsman. This criterion delimits severely the number of statistical methods useful for our purpose. We have found two sequential nonparametric methods to be of practical value for channel maintenance programs. They will be discussed along with a parametric statistical procedure (offered as contrast), known as linear discriminant analysis.

## II. STATISTICAL METHODS

We begin describing the statistical approach to solving the channel maintenance problems by first considering a conventional parametric statistical procedure. The purpose to be served by the first procedure is that of considering a large number of measurements simultaneously, which then sets the stage for the sequential methods to follow.

### *A. Linear Discriminant Analysis* (*LDA*) [1]

*1) The Linear Discriminant Method of Classification:* Consider $n$ channel classes, each defined by a specific grade of quality. Let $Q$ denote the number of measurements $(x_1,x_2,\cdots,x_Q)$ which will be made on a channel. The possible outcomes of these measurements can be described as $n$ $Q$-dimensional distribution functions, each expressing the probability that a channel from its class has a vector value $x = (x_1,x_2,x_3,\cdots,x_Q)$. If $R_j$, $j = 1,2,\cdots,n$ denotes that region in $Q$-dimensional space in which a channel will be classified as belonging to class $J$, then the probability of correct channel classification can be determined from the distribution functions.

If the elements of each of the $n$ classes are normally distributed according to $N(\mu_j,\Sigma)$, $j = 1,2,\cdots,n$, with the same covariance matrix but different means, then the minimization of the probability of misclassification occurs when $R_j$, $j = 1,2,\cdots,n$ is the set of points

Paper approved by the Associate Editor for Data Communication Systems of the IEEE Communications Society for publication after presentation at the National Telecommunications Conference, San Diego, Calif., December 1974. Manuscript received January 15, 1975; revised January 22, 1975.

The authors are with Bell Laboratories, Holmdel, N. J. 07733.

$x$ where $u_j(\mathbf{x})$ is maximum over the set $\{u_k(x)\}_{k=1}^n$ and where

$$u_k(x) = 2x^T \Sigma^{-1} \mu_k - \mu_k^T \Sigma^{-1} \mu_k. \quad (1)$$

$u_j(x)$ is called the linear discriminant function for the class indexed by $j$.

The technique can be thought of as a search for a set of hyperplanes of dimension $Q - 1$ (where $n$ is the number of populations and $Q$ is the number of variates, in our case, channel parameters) which partitions exhaustively a hypercube of dimension $Q$ designated by the ranges of the values the variates can take, and whereby the partition separates the observations of any class from those of any other class. The partitioned hypercube then consists of $n$ convex regions, each region being identified with a single population or class.

As an illustration of how linear discriminant analysis (LDA) would work, Fig. 1 depicts the separation of two channel classes with regard to two channel measurements. The hyperplane which divides the classes is line $C$ (one dimensional) since there are only two measurements.

An evaluation of the difference between the populations is obtained from the Mahalanobis statistic $V$ [2] given by

$$V = \sum_{i=1}^{Q} \sum_{j=1}^{Q} \alpha^{ij} \sum_{r=1} N_r(\bar{x}_{ir} - \bar{x}_i)(\bar{x}_{jr} - \bar{x}_j) \quad (2)$$

where

- $\alpha^{ij}$ entries of the matrix $\Sigma^{-1}$
- $N_r$ sample size of population $r$
- $\bar{x}_{ir}$ sample mean of population $r$
- $\bar{x}_i$ pooled sample mean of variate $i$ given by

$$\bar{x}_i = \sum_{r=1}^{n} N_r x_{ir} / \sum_{r=1}^{n} N_r.$$

The statistic can be used as $\chi^2$ with $Q \cdot (n - 1)$ degrees of freedom for testing the null hypothesis. In this case, the null hypothesis would be that the mean values of the $n$ populations of channels designated by the error rates of specific data sets are really estimates of the same mean.

*2) Application of LDA:* We have considered a data base consisting of 39 measurements made on each of 484 channels. Four grades (namely, Classes I, II, III, IV) of quality were established, and each channel was assigned a grade according to the predetermined specifications regarding the quality of the transmission of information it would support. These grades were not subjectively determined. Rather, they follow standard industrial settings which had been used for evaluating this type of channel. The measurement vector consisted of many different types of meter readings, including those measuring ambient noise, impulse noise, envelope delay distortion, harmonic distortion, intermodulation distortion, etc. Again, the measurements selected for characterizing each channel were those measurements which experienced maintenance engineers felt contained valuable information about channel quality.

The identical form of information transmission was used on each channel.

An LDA computer program [3] was used to compute the mean values of individual classes and the respective $V$ statistic values for the data base according to each grouping of the channel classes (I versus II versus III versus IV; I, II, III versus IV; I, II versus III, IV; and I versus II, III, IV). In every case, the $V$ statistic value was high, indicating a distinct difference in the classes.

Channels were classified by evaluating (1). This was carried out by a linear discriminant computer program [3], using the channel data base. The results of this classification method (known as the resubstitution method) are given in Table I.

In the resubstitution method of measuring classification accuracy, the same data are used for evaluating the linear discriminant function and for finding the percentage of misclassification. The results are therefore optimistic. How optimistic they are is not clear, but since the data base is large, the bias introduced by resubstitution should not be great. The degree of bias was examined by another means of estimating the misclassification rate. A recent paper [4] considers the merit of using linear discriminant techniques for estimating misclassification rates, and reported that three less optimistic (and presumably more valid) estimates can be obtained from the $V$ statistic. The three estimates are termed the $D$, $DS$, and the $D^*$, and the actual estimates of the misclassification rate obtained are given by $\Phi(-D/2)$, $\Phi(-DS/2)$, $\Phi(-D^*/2)$ where $\Phi$ is the standardized normal distribution. The parameter $\psi$ denotes the estimate of the probability of misclassification obtained through resubstitution.

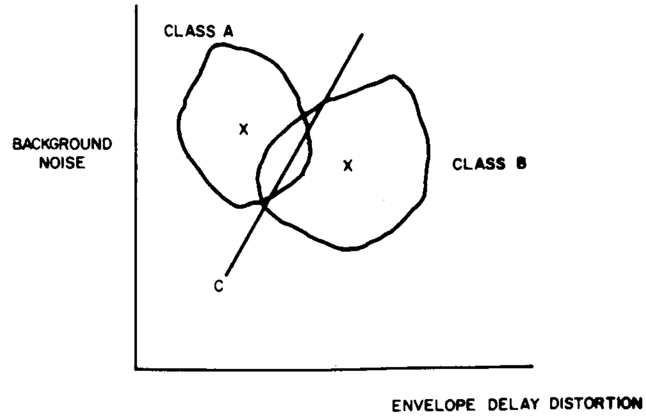



Fig. 1. LDA; example for channel classification. Separation of classes is enacted by a linear combination of channel variables (line $C$).

Estimates of linear discriminant misclassification rates (two classes only) for the channel data obtained through the $D$, $DS$, $D^*$, and $\psi$ methods are given in Table I. It can be seen that in many cases, all estimates are nearly the same. In those cases where they do differ, we can assume that the data were not of multivariate normal character or the population sizes $(N_1, N_2)$ of the classes were rather small compared to the number of dimensions $Q$.

In all cases of channel classification, no *a priori* distributions were assigned to the individual channel classes, nor were misclassification costs assumed.

### B. Partition Analysis

The objective of partition analysis is to examine the effects of a set of interrelated independent variables on a dependent variable. The analysis employs a nonsymmetical branching process based on variance analysis to subdivide the sample into a series of subgroups which maximize one's ability to predict values of the dependent variable. A complete description of the technique is given in [3], so we shall describe its utilization here by way of an example.

*1) An Example of Partition Analysis:* The dependent variable under consideration is

| | Code | Acutal Range |
|---|---|---|
| channel grade $Y$ = | 1 | Excellent (Class I) |
| | 2 | Good (Class II) |
| | 3 | Fair (Class III) |
| | 4 | Poor (Class IV) |

and two exemplary independent variables might be

| | Code | Actual Range |
|---|---|---|
| measurement $X_1$ = | 1 | 0 |
| | 2 | 1–5 units |
| | 3 | 6+ |

TABLE I
LDA Classification Scheme Summary (484 Channels)

| CLASSIFICATION STRATEGY | $N_1$ | $N_2$ | $N_3$ | $N_4$ | V | $D^2$ | $\phi\left(\frac{-D^+}{2}\right)$ | $\phi\left(\frac{-DS}{2}\right)$ | $\phi\left(\frac{-D}{2}\right)$ | $\psi$ | $1-\psi$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| I vs II vs III vs IV | 167 | 142 | 115 | 60 | - | - | - | - | - | .48 | .52 |
| I, II, III vs IV | 424 | 60 | - | - | 128 | 2.44 | .27 | .23 | .22 | .16 | .84 |
| I, II vs III, IV | 309 | 175 | - | - | 179 | 1.60 | .30 | .27 | .26 | .26 | .74 |
| I vs II, III, IV | 167 | 317 | - | - | 153 | 1.40 | .32 | .29 | .28 | .27 | .73 |

$N_K$ - No. of channels in group K

V - Mahalanobis statistic value

$\phi\left(\frac{-D^+}{2}\right)$ - misclassification rate estimated by $D^+$ statistic

$\phi\left(\frac{-DS}{2}\right)$ - misclassification rate estimated by DS statistic

$\phi\left(\frac{-D}{2}\right)$ - misclassification rate estimated by D statistic

$\psi$ - misclassification rate using resubstitution method

and

$$\text{measurement } X_2 = \begin{cases} 1 & \text{0–19 units} \\ 2 & 20+. \end{cases} \qquad \begin{matrix}\text{Code} & \text{Actual Range}\end{matrix}$$

Suppose that the data consist of the following four observations.

| Observation Number | $Y$ | $X_1$ | $X_2$ |
|---|---|---|---|
| 1 | 1 | 1 | 1 |
| 2 | 1 | 2 | 1 |
| 3 | 2 | 3 | 1 |
| 4 | 4 | 2 | 2 |

Now there are four ways we can partition the observations using the variables $X_1$ and $X_2$ individually, namely,

$$X_1 = 1 \quad \text{versus} \quad X_1 = 2,3,$$

$$X_1 = 2 \quad \text{versus} \quad X_1 = 1,3,$$

$$X_1 = 3 \quad \text{versus} \quad X_1 = 1,2,$$

and

$$X_2 = 1 \quad \text{versus} \quad X_2 = 2.$$

In order to determine which partition is best, we must have some measure of how well a partition discriminates between values of the dependent variable. The measure we shall consider is between sum of squares (BSS), which may be described as follows.

Suppose $Y_1,\cdots,Y_N$ is a sample of dependent values which has been partitioned into two subgroups: $Y_{11},\cdots,Y_{1N_1}$ and $Y_{21},\cdots,Y_{2N_2}$. The total sum of squares (TSS) in the sample is given by

$$\text{TSS} = \sum_{j=1}^{N} (Y_j - \bar{Y})^2$$

where $\bar{Y}$ is the overall mean, and the sum of squares within each subgroup (WSS) is given by

$$\text{WSS}_i = \sum_{j=1}^{N_i} (Y_{ij} - \bar{Y}_i)^2$$

where $\bar{Y}$ is the mean within each subgroup. The between sum of squares is given by

$$\text{BSS} = N_1(\bar{Y}_1 - \bar{Y})^2 + N_2(\bar{Y}_2 - \bar{Y})^2$$

and it is easily seen that

$$\text{TSS} = \text{WSS}_1 + \text{WSS}_2 + \text{BSS}.$$

The effect of choosing that partition which maximizes BSS is therefore to maximize the amount of variation accounted for by the partition, while at the same time minimizing the amount of residual variation within the resulting subgroups.

In our example, TSS = 6 and the BSS for each partition is given by the following.

| Partition | Part 1 | Part 2 | $\text{WSS}_1$ | BSS | $\text{WSS}_2$ |
|---|---|---|---|---|---|
| $X_1 = 1$ versus $X_1 = 2,3$ | 1 | 1,2,4 | 0 | 1.3 | 4.7 |
| $X_1 = 2$ versus $X_1 = 1,3$ | 1,4 | 1,2 | 4.5 | 1.0 | 0.5 |
| $X_1 = 3$ versus $X_1 = 1,2$ | 2 | 1,1,4 | 0 | 0 | 6 |
| $X_2 = 1$ versus $X_2 = 2$ | 1,1,2 | 4 | 0.7 | 5.3 | 0 |

Therefore, the partition on $X_2$ best discriminates between values of of the dependent variable and accounts for about 88 percent of its variation. The partition tree associated with this example is shown in Table II.

Once a sample has been partitioned into two subgroups, the same analysis can be applied to each of them, thereby generating a succession of bipartite partitions of the original sample. In the discrimination of channel quality, we see that each partition leads to a group of channels whose quality is nearly the same. Thus, channel measurements and ranges used to generate the partitions lead in an obvious way to the identification of a population of channels whose quality for the transmission of information is nearly uniform.

*2) Partition Analysis of a Channel Data Base:* Forty independent variables were considered at each stage in the partition analysis of a data base consisting of 602 channels. The few instances of missing data were coded as a value of zero with negligible effect. The results of the partition analysis appear in Tables III and IV where we note that only partitions which accounted for more than three percent of the total sum of squares and led to different error classes were considered (although at times not all channel classes are represented at the endpoints). The program used to perform these analyses was written by Sonquist and Morgan of the Institute for Social Research, University of Michigan, Ann Arbor [5]. Table IV shows the independent variable that was used to partition each group, along with the range of the variable associated with the high-quality channel group. The quality class distribution within each final subgroup is also

TABLE II
PARTITION TREE FOR EXAMPLE IN SECTION II-B1)

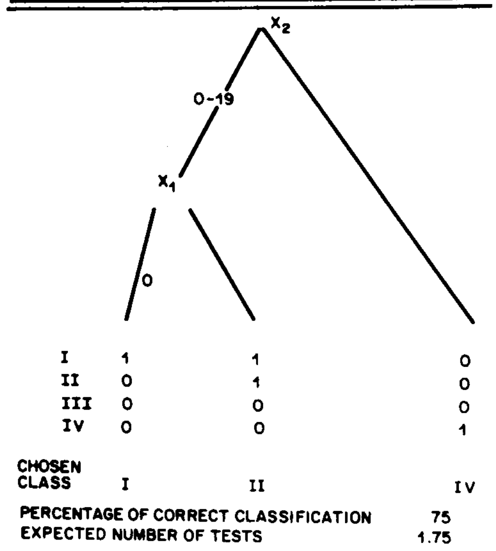
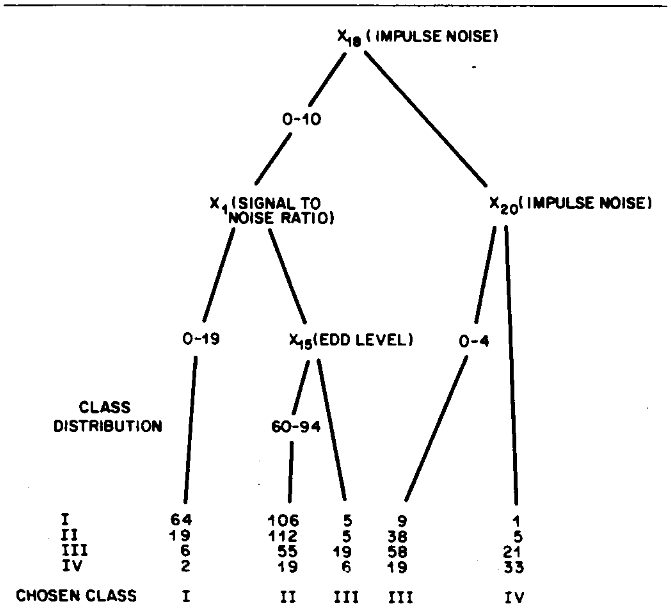



TABLE III
PARTITION ANALYSIS CLASSIFICATION SCHEME SUMMARY

| CLASSIFICATION STRATEGY | PERCENTAGE OF CORRECT CLASSIFICATION |
|---|---|
| I vs II vs III vs IV | 48 |
| I, II, III vs IV | 90 |
| I, II vs III, IV | 76 |
| I vs II, III, IV | 79 |
| | EXPECTED NUMBER OF TESTS |
| I vs II vs III vs IV | 2.5 |
| I, II, III, IV | 2.1 |
| I, II vs III, IV | 1.7 |
| I vs II, III, IV | 1.6 |

TABLE IV
PARTITION TREE—CLASSES I VERSUS II VERSUS III VERSUS IV




given, along with the maximum likelihood estimate of the channel class therein. Program limitations sometimes prevent each final subgroup from being associated with a different channel class, and can also preclude a channel class from being identified with some final subgroup. The percent of correct classification figures and the expected number of tests needed to attain this classification rate is given in Table III.

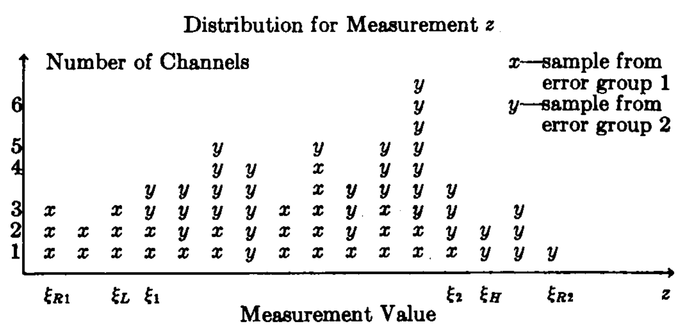
*3) Application of Partition Analysis to the Identification of Data Channel Quality:* The advantage of partition analysis as a channel classifier lies in the fact that it sequentially "selects" the channel feature and associated range which best distinguishes quality in a data channel. The sequential aspect of the technique minimizes the number of tests a craftsman, for example, may need to make to ascertain channel quality. The time and economic considerations of this advantage demonstrate its importance in possible troubleshooting applications.

Of course, no feature or small number of channel features is seriously expected to "explain" the error rates observed in a channel. It is more likely that a great many causes of substandard performance are at work in a channel. Thus, in general, the more measurements we make on a channel, the better we can assess its quality. Further evaluation of any channel quality selector would then require consideration of both the economy and accuracy aspects of using many channel measurements to identify channel quality.

### C. Nonparametric Test by Kendall

Channel quality identification has been found to be a difficult classification problem, especially since channel measurement distribution forms are unknown. Hence, a need for nonparametric testing procedures has been demonstrated. Nonparametric tests (or distribution-free methods) usually involve rank comparisons, and it is this idea [6] which will form the nucleus of the channel classification procedure we will now discuss.

Suppose the following is a sample histogram for a channel measurement $z$ where we have retained the group identity of the sample point.




We have indicated several points ($\xi_L$, $\xi_1$, $\xi_2$, and $\xi_H$) on the $\xi$ axis which are of special interest. Up to point $\xi_L$, measurement $z$ has a distribution portion formed by only members of group 1, namely, $x$'s. A similar situation occurs for the $\xi_H$ point. There is no group 1 channel above $\xi_2$.

Thus, $\xi_L$ and $\xi_H$ can be used to identify correctly eight channels on the left and six channels on the right. Our testing procedure for measurement $z$ then becomes the following.

1) Declare channel to be in group 1 if $\xi < \xi_L$.
2) Declare channel to be in group 2 if $\xi \geq \xi_H$.
3) Make no decision if $\xi_L < \xi < \xi_H$.

Notice that it is possible that $\xi_1 = \xi_{R1}$, namely, that there is a mixture of $x$'s and $y$'s for every $\xi$ value all the way to the left end of the distribution. Also, we envision the possibility of $\xi_2 = \xi_{R2}$. When either of these cases occurs (but not both), it is still possible to gain from our testing procedure with measurement $z$. Of course, the meas-

TABLE V
DISTRIBUTION DECOMPOSITION CLASSIFICATION; CLASSES I, II, III VERSUS IV

| Step No. | Measurement No. | No. of Calls Left Strip | No. of Calls Right Strip | Value of Variable ($\xi_1$) | Value of Variable ($\xi_2$) | No. Remaining Unidentified Calls | Class of Calls in Left Strip | Class of Calls in Right Strip | Type of Measurement |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 19 | 62 | 37 | ( -116, | -84) | 281 | 1 | 1 | Intermod' Distortion |
| 2 | 18 | 63 | 3 | ( -4, | 16) | 215 | 1 | 1 | Intermod' Distrotion |
| 3 | 24 | 1 | 42 | (-1056, | -386) | 172 | 1 | 1 | Amplitude Distortion |
| 4 | 27 | 6 | 26 | ( -19, | 36) | 140 | 1 | 1 | Amplitude Distortion |
| 5 | 20 | 15 | 3 | ( -22, | -8) | 122 | 1 | 1 | Intermod' Distortion |
| 6 | 34 | 4 | 12 | ( -46, | 397) | 106 | 1 | 1 | Delay Distortion |
| 7 | 29 | 1 | 12 | ( -23, | 5) | 93 | 2 | 1 | Amplitude Distortion |
| 8 | 38 | 10 | 3 | ( -101, | 212) | 80 | 1 | 2 | Delay Distortion |
| 9 | 39 | 13 | 1 | ( -45, | 131) | 66 | 1 | 2 | Delay Distortion |
| 10 | 28 | 4 | 5 | ( -14, | 11) | 57 | 1 | 1 | Amplitude Distortion |
| 11 | 35 | 2 | 6 | ( -835, | -433) | 49 | 1 | 1 | Delay Distortion |
| 12 | 4 | 0 | 6 | (*****, | 33) | 43 | 1 | 2 | Impulse Noise |
| 13 | 31 | 6 | 13 | ( -4, | 0) | 24 | 1 | 1 | Amplitude Distortion |
| 14 | 23 | 9 | 0 | ( -55, | *****) | 15 | 1 | 1 | Intermod' Distortion |
| 15 | 15 | 2 | 4 | ( -64, | -51) | 9 | 1 | 1 | Intermod' Distortion |
| 16 | 13 | 3 | 2 | ( 38, | 44) | 4 | 2 | 1 | Harmonic Distortion |
| 17 | 4 | 3 | 1 | ( 24, | 1) | 0 | 1 | 2 | Impulse Noise |

urement is useless for identification purposes if $\xi_1 = \xi_{R_1}$ *and* $\xi_2 = \xi_{R_2}$.

It is possible to go through this procedure for every measurement of the $Q$ available to us, $Z_j, j = 1,2,3,\cdots,Q$. So we have a pair $(\xi_L(K),\xi_H(K))$ and total $N(K)$ of channels identified correctly associated with each measurement, $k_1 = 1,2,\cdots,Q$. Suppose we choose as our first channel measurement that $z_{K_1}$ for which $N(K_1)$ is the maximum, i.e., $N(K_1) = \max_{1\leq j\leq Q} \{N(j)\}$. Having identified $N(K_1)$ channels correctly with a $(\xi_L(K_1),\xi_H(K_1))$ pair, we now turn to the set of channels for which no identity decision was made in this step. We know that $z_{K_1}$ is useless as a possible identifier for the next step, so we do not consider it. But we do redraw all our other distributions for $z_j \neq z_{K_1}, j = 1,2,\cdots Q$ without those $x$'s and $y$'s (channels) which have been identified by $z_{K_1}$. Again we find $z_{K_2}$ which has a maximum $N(K_2) = \max_{1\leq t\leq Q; t\neq K_1} \{N(t)\}$ and has $(\xi_L(K_2), \xi_H(K_2))$ among all redrawn distributions. Now we have $N(K_1) + N(K_2)$ channels correctly identified. Again we discard these identified channels from our set of data and redraw the distributions. However, this time we can consider $z_{K_1}$ as a possible identifier, but not $z_{K_2}$ since a different set of data is now used for drawing the distributions. Our iterative procedure continues until either we classify all channels correctly (i.e., no more work to be done) or we come to an impasse formed by the situation in which *all* channel measurements have $\xi_1(l) = \xi_{R_1}(l)$ *and* $\xi_2(l) = \xi_{R_2}(l)$, $l = 1,2,\cdots Q$. The latter event is unlikely to occur if we are using the smallest cell size possible for the distribution and we have a large number of measurements $(Q)$.

Thus, by listing enough measurements with appropriate thresholds, we can expect to identify all channels correctly. However, a reasonable testing procedure should involve only a few measurements; for example, a maximum of, say, 20. Hence, if our sequential extreme value testing scheme does not identify all the channels in the data base with a few measurements, then little has been gained.

We list in Table V the sequence of measurements found for a data base consisting of 380 channels with 39 measurements for grouping I, II, III versus IV (although other groupings are not shown because of space limitations). A sequence of 89 measurements was required to identify the four classes individually, while the other class groupings required about half this number or fewer. It is clear that there is a slow convergence of the algorithm to identify all the channels correctly.

## III. DISCUSSION AND SUMMARY

Each channel testing program is unique in the sense that maintenance engineers differ in each case as to what is important for quality definition. Quality can be a subjective judgment, even when numerical measurements such as error rates are involved. In this concise paper, we have minimized the attention given to the actual definition of channel quality for the examples given. Also, the actual measurements used in the examples were not detailed, except to say as to what category they pertained. Meters and reference levels differ from problem to problem. In short, measurements differ since tradition and expediency dictate what calibration and measurement equipment is available. We believe what is important is the discovery of relationships between the accepted definition of quality and the available measurements. To date, no serious attention has been given to using statistical methods for uncovering these relationships on a large scale. Our study has involved hundreds of thousands of channel measurement and quality data entries. Such a massive amount of information may not have been available before, thus preventing statistical procedures from being applied to the channel testing problem.

In this concise paper, we have related our experience in constructing a channel testing program for maintenance purposes from a statistical point of view. LDA has been used to offer contrast as a multidimensional parametric method to two sequential nonparametric methods. These latter techniques, termed partition analysis and distribution decomposition, were chosen for their simplicity in the channel testing application. As we became more acquainted with actual channel maintenance program, it was clear that a craftsman usually does not have access to a computation center, nor does he have the background to use it with facility. Thus, the features of simplicity and implementability eliminated many statistical methods which could have been applied to channel testing. To gain insight into the data bases, we have also used contingency tables, standard regression, stepwise regression, and stepwise LDA. The data bases did not exhibit strong linear relationships between the dependent variable and independent variables chosen. This seems to be the case for communication channels where nonlinear interrelationships exist between the measurement values (e.g., toll telephone channels).

The properties of the sequential methods are such that they iterate toward a decision with approximately the same craftsman effort by branch testing, whereas LDA requires all tests to be performed before a decision can be made. Hence, it is understandable that the sequential methods iterated toward a decision faster, and thus were cheaper to implement than the LDA with approximately the same classification rate. We have applied these techniques to dozens of data bases, and we have seen these differences between the routines appear again and again. Of course, the effectiveness of each routine is best measured by the specific application. For example, when few measurements are available, LDA may not be that different from the sequential methods in time to detect since little computation has to be done.

We belive our study demonstrates that statistical methods constitute a valuable tool for constructing channel maintenance programs. It leads engineers to those measurements which reveal the most troubleshooting information in the least amount of time. In addition, it allows programs to be evaluated as to their effectiveness by simulation.

## ACKNOWLEDGMENT

The authors express their gratitude to J. Tischendorf for the many helpful discussions they have had with him concerning the project reported here. In addition, they thank B. Butler and F. J. Brophy for their programming assistance.

## REFERENCES

[1] T. W. Anderson, *Introduction to Multivariate Statistical Analysis.* New York: Wiley, 1958.
[2] C. R. Rao, *Advanced Statistical Methods in Biometric Research.* New York: Wiley, 1952.
[3] W. J. Dixon, Ed., "Biomedical computer programs," Health Sci. Computing Facility, Univ. California, Los Angeles, Programs BMD05M, BMD07M, Sept. 1, 1965.
[4] P. Langebruch and M. Mickey, "Estimation of error rates in discriminant analysis," *Technometrics*, p. 1, Feb. 1968.
[5] J. A. Sonquist, E. L. Baker, and J. N. Morgan, *Searching for Structure*, revised ed. Ann Arbor, Mich.: Inst. for Social Res., Univ. Michigan, 1974.
[6] M. G. Kendall, "Discrimination and classification," in *Multivariate Analysis*, P. R. Krishnaiah, Ed. New York: Academic, 1966, pp. 165–185.