Statistical Classification in Anthropology: An Application to Ethnomusicology

LINTON C. FREEMAN
University of Illinois
and
ALAN P. MERRIAM
University of Wisconsin

WHILE a growing interest in the application of statistics to cultural anthropology is indicated by the current literature (Driver 1953; Clements 1954), it is only relatively recently that such processes have received any widespread attention from cultural anthropologists. The earliest application was made by Tylor (1889), but Driver and Kroeber (1932) with their trait-element list were the forerunners of the current trend which has culminated in Murdock's extensive study of the interrelationships among various aspects of social structure (1949). With the development of the Cross Cultural Survey at Yale as well as the Human Relations Area Files, statistical treatment of cultural materials has become relatively common (Anon. 1954a); the importance of the trend is perhaps best indicated by the fact that an entire section of the 1954 meeting of the American Statistical Association was devoted to "The Use of Statistics in Anthropological Studies " (Anon. 1954b).

Most statistical studies in cultural anthropology have been relational. They have attempted to determine the degree of association which exists between two or more cultural variables through application of the techniques of correlation. However, correlational techniques may be extended to deal with problems of classification. It is with problems of this latter sort that the present paper is concerned.

Classification is one of the central problems of cultural anthropology. Types are established and distinctions are made among them, but for the most part both the classifications and the distinctions are intuitively established. Even when mathematical techniques are used, they are usually restricted to simple frequency counts and comparisons of percentages. At best, such procedures afford only crude generalizations and comparisons. The present report is designed primarily as a demonstration of the power of statistical techniques in classifying and comparing.

Statistical techniques provide the investigator with precise estimates of the homogeneity of objects classified together. Furthermore, one may determine the exact degree to which differences obtain between two or more classifications. Comparisons made by statistical means need no longer be based on intuitive judgment nor subject to revision on the basis of "expert" opinion. Accurate statements of probability describe the confidence with which statistical generalizations may be held. In short, statistical techniques can provide system and precision in constructing and comparing classificatory types.

Among those aspects of culture which seem particularly susceptible to statistical analysis, music is perhaps outstanding (Herskovits 1949:573–74). This is partly due to the nature of the way in which material can be gathered in the field, and partly due to the structure of music itself. With reference only to music as a structure, it may be said that laboratory control is more closely approximated in ethnomusicological studies than in most anthropological methodology. Through the use of a recording device, a specific cultural activity can be mirrored in process; the activity may be examined over and over again without change, and the element of subjectivity which enters into observational method can be excluded to a considerable degree. Thus in field work itself, a high degree of control over the materials is achieved with a minimum of effort. Music is also particularly susceptible to objective analysis once collection has been made; through study of its actual structure the investigator can define a musical style without recourse to subjective characteristics based either upon his own preconceptions of what music should be structurally or upon qualities of music which cannot be objectified. Thus in a specific rendition of a song, such qualities as pitch, tonal range, tonal duration, interval usage, modal structure, and others, cannot only be isolated, counted, computed, and analyzed, but can be objectified in terms of the fact that in the rendition at hand they do not change. This does not indicate, of course, that variation should not be taken into account, but it does suggest that if a song A has a variant B, the structure of each can be objectively isolated from a number of rigorous standpoints and then compared in like terms. These characteristics of music, which in the sense discussed above are analogous to those of language, make it a potentially rich area for inquiry by methods of statistics.

So far as is known to the present authors, no statistical applications as such have been made in the field of comparative musicology. Various trigonometric calculations have been used in the study of tonometrics (Hornbostel 1921; Kunst 1950), but these are not statistical in the sense used here. Arithmetical counts and percentages have also figured in comparative musicology, as in the work of Densmore (1929), and these have been extended by Kolinski (1936), for example, until they form the basis for at least one method of research inquiry. The set of problems treated here is derived from the Kolinski method.

One of the techniques used by Kolinski, as well as by other students, is the interval count and the resulting percentages of specific intervals used in a particular body of music; this is usully broken down into ascending intervals, descending intervals, and total intervals. Thus in a brief study made by one of the present authors (Merriam 1949, 1950) of Cheyenne music, it was found that in a groups of songs 33 percent of the total intervals were major seconds, 27 percent minor thirds, 15 percent perfect fourths, and so on; similar percentages expressed the relative use of specific ascending and descending intervals. The purpose of such analysis is to express in objective terms a particular musical measure which will differentiate between groups of songs. Thus it is hopefully assumed that by comparing interval counts, as well as other characteristics, a differentiation could be made between Cheyenne and Iroquois songs, for example; or conversely, that if the percentage of major seconds used in an unknown body of song is 32.66 percent, it has a far greater probability of being Cheyenne than anything else. This is, of course, reducing the problem in the extreme; a single interval count would probably not give such an indication. However, there is a possibility that a group of interval counts might do so, and it is this possibility which is investigated here. The question, then, is whether interval count will actually differentiate between two bodies of music; the problem is to find a statistical procedure which will indicate both the plausibility of the use of intervals or other measures to differentiate bodies of song, and the extent to which it is reliable.

If music can be classified into various types, and distinction made among these types by statistical methods, several problems in comparative musicology can be attacked with considerably greater precision than has heretofore been possible. Restricting ourselves to the interval count as an example, we have already pointed out this method might be used to differentiate between two groups of songs; more specifically, interest lies in whether or not certain groups of percentages in interval usage can be used as a criterion of identifying a body of song. In extension, if the measure proves valid, it should also be possible to trace musical influences which have played upon a specific group or tribe. Further, if it is established that specific interval usages and other structural forms characterize specific cultures or subcultures, a clearer picture of the relationship between music and other elements of culture is obtained, and the individuality of esthetic expression as it is shaped by the customs of a particular group is more sharply established. The present paper does not hope to answer these problems, but rather to suggest the possibilities which may be assayed through the use of statistics. It must also be emphasized that the interval count alone, or for that matter ethnomusicology itself, cannot be expected to indicate past contact or derivations with finality without corroboration from other investigations. The normal precautions taken by anthropologists when postulating contact are equally valid in ethnomusicology; in this sense the study of music is simply contributory as an added technique of anthropological investigations.

The problem of classifying into types and distinguishing among them is not peculiar to the field of ethnomusicology. In any science the investigator is frequently faced with the necessity of differentiating between two groups of objects on the basis of measurements of several characteristics of each. The procedure is simple in those cases in which the measures yield very different values for the two groups, but where the measured values are fairly similar for objects in the two groups there may be considerable overlap in the distribution of any single measure; thus differentiation on the basis of any one of the measurements may be impossible.

The process of distinguishing among classes of objects on the basis of a set of measurements of their properties calls for a specialized statistical technique. Such a technique, the discriminant function, has been developed by R. A. Fisher (1936). This function enables the investigator to make a linear combination of several variables in such a manner that scores on the resulting combined index will possess a minimum of distributional overlap. Thus the possibility of grouping the objects into their proper classes is increased, and the chance of error in any single classification is reduced. In short, the discriminant function provides a technique of weighting scores so that a maximum separation or differentiation between groups is accomplished.

In practice, Fisher's technique may be applied to any multiple measurements on two or more groups. Weights are computed and the original measurements converted into index scores. A critical value of the index is determined, and all cases which fall above it are classed as members of one group while those falling below it belong to the other. New cases may then be transferred into the index classification and assigned to their proper groups with a high degree of success.

Anthropological applications of the discriminant function need not, of course, be limited to studies of ethnomusicology; they may be made wherever successive measures are taken on any object of interest. The technique has been employed at least once in craniometry (Barnard 1935) to describe a progressive trend for a dated series of skulls. It has also been used recently in classifying teeth (Bronowski and Long 1952). Applications are possible in any area of anthropometry, in linguistics, or in almost any study of material culture. Pottery, for example, might be classified, compared, and related through use of the discriminant function. Empirical types could be established and these might be contrasted and their derivations traced in terms of their overall differences. Music was chosen for the present illustrative report, but any of the objects noted above upon which suitable data were available might also have been used.

The music selected for the present study represents New World Negro material derived from Africa—the two groups are the Ketu cult of Bahia, Brazil, and the Rada cult of Trinidad. Ceremony, including music, in the Ketu groups is derived from the religious practices of the Yoruba people of Nigeria (Bascom 1944; Carneiro 1936; Ramos 1943), while that of the Rada stems from Dahomey (Herskovits 1938; Carr 1953). The Ketu songs were recorded by M. J. and F. S. Herskovits in 1941–42 in Bahia (M. J. and F. S. Herskovits 1943), the Rada songs by Andrew Carr in Trinidad in 1953; both groups of songs were analyzed by one of the present writers (Merriam 1951; Merriam, Whinery and Fred 1954). The selection of these materials is based on the special interest in attempting to differentiate between two derived variants of what has been classed as a regional style (Waterman 1952).

It was clear from the outset that a random sample of these song-groups would be impossible to obtain, since the universe of songs in any group was necessarily unknown. While without sampling, statistical results could not be generalized to all of the music of the societies in question, such sampling was considered unnecessary in view of the fact that this application was designed primarily to suggest a technique of analysis.

Considerations of economy, availability of materials, and potential productivity led to the selection of three measures of interval use; these are song length measured in intervals, the proportionate use of major seconds, and the proportionate use of minor thirds. Proportion of use was chosen in order to avoid any systematic bias resulting from differential song length, since one very long or very short song might have completely changed the results.

Twenty songs of each of the two musical forms were selected. Interval counts were made, and the frequencies of use of major seconds, minor thirds, and total intervals were determined for each song. The frequencies of major seconds and minor thirds were divided by their song length, and the resulting proportions were tabulated. Means were computed for each form and differences were calculated. These means and differences are reported in Table 1.

TABLE 1. MEANS AND DIFFERENCES OF UNCORRECTED MEASURES FOR RADA AND KETU

	Major Seconds	*Minor Thirds*	*Total Intervals*
Radu	.2335	.4081	148.9
Ketu	.4240	.2086	134.2
Difference	−.1905	.1995	14.7

A glance at this table indicates that differences between means do obtain on each of the three measures. "Student's" *t*-test was employed in order to evaluate the significance of these differences.[1] Table 2 shows that the mean differences for major seconds and minor thirds were each significant beyond the one percent level of confidence, while the value for totals was not significant at that level.

TABLE 2. SIGNIFICANCE OF DIFFERENCES IN UNCORRECTED MEASURES FOR RADA AND KETU

	Difference	*S.E.*	*t*	*P*
Major Seconds	−.1905	.0297	6.41	<.01
Minor Thirds	.1995	.0307	6.50	<.01
Total Intervals	14.7	17.5779	.836	>.01

In other words, the measured differences between Rada and Ketu may be evaluated as reflecting a real difference in proportionate use of major seconds and of minor thirds. Differences as great as those observed would have arisen by chance in a sample of songs less than one time in a hundred. Total interval use, however, presents a less well-defined result. A difference as great as the one observed would arise through the operation of chance factors alone more than one time in one hundred. It must therefore be concluded that no significant difference in total interval use has been demonstrated.

The results of the *t*-tests indicate the possibility of accurate discrimination between samples of the music of Rada and Ketu either on the basis of use of major seconds or use of minor thirds. There is, however, a large amount of overlap in the distributions of each measure. Any attempt to classify a single song on the basis of either of these intervals would be subject to a high probability of error. Even the distribution of minor thirds, which is the best index of discrimination, contains this wide area of overlap in which it is impossible to distinguish between music of Rada and Ketu. Table 3 illustrates that any song which involves use of minor thirds in a range from .20 to .49 cannot be

TABLE 3. DISTRIBUTIONS OF MINOR THIRDS FOR RADA AND KETU

Score	*Rada*	*Ketu*	
.50–.59	4		
.40–.49	7	2	Area of overlap
.30–.39	6	2	
.20–.29	3	5	
.10–.19		9	
. 0–.09		2	

classified on the basis of this measure. Twenty-five of the forty songs fall into this area of overlap, and it is likely that half of these (12.5) would be misclassified. Thus, thirty-one percent of all the songs in the sample would be incorrectly categorized if use of minor thirds was employed as the only classifying instrument. It therefore seemed advisable to compute a discriminate function on all three of these measures in order to maximize their ability to differentiate.

In employing the discriminant function, maximum differentiation is achieved by computing a set of lambda (λ) scores which are used to weight the various measurements.[2] These scores result from the solution of the following three simultaneous equations in three unknowns:

$$\lambda_2 \sum x_2^2 + \lambda_3 \sum x_2x_3 + \lambda_T \sum x_2x_T = d_2$$

$$\lambda_2 \sum x_2x_3 + \lambda_3 \sum x_3^2 + \lambda_T \sum x_3x_T = d_3$$

$$\lambda_2 \sum x_2x_T + \lambda_3 \sum x_3x_T + \lambda_T \sum x_T^2 = d_T$$

These equations may be solved by substituting the known values and solving for three unknowns.[3] Solving for the three lambdas yielded the following results:

$$\lambda_2 = -.49069$$

$$\lambda_3 = .47958$$

$$\lambda_T = .00045$$

The final step involved computation of the mean weighted score for each group and the point of discrimination between the groups. For this operation the following equations were employed:

For the Ketu,

$$D_K = \lambda_2\bar{X}_{2K} + \lambda_3\bar{X}_{3K} + \lambda_T\bar{X}_{TK}$$

For the Rada,

$$D_R = \lambda_2\bar{X}_{2R} + \lambda_3\bar{X}_{3R} + \lambda_T\bar{X}_{TR}$$

and for the point of discrimination,

$$D = \lambda_2\bar{X}_2 + \lambda_3\bar{X}_3 + \lambda_T\bar{X}_T$$

Substituting means and λ's, the following results were obtained:

$$D_K = (-.49069)(.4240) + (.47958)(.2086) + (.00045)(134.2) = -.0476$$

$$D_R = (-.49069)(.2335) + (.47958)(.4081) + (.00045)(148.9) = .1481$$

$$D = (-.49069)(.3288) + (.47958)(.3084) + (.00045)(141.55) = .0503$$

Thus the mean weighted score for Rada was .1481 while that for Ketu was −.0476. The point of discrimination halfway between these values was .0503. Any weighted score falling above .0503 would be classified as Rada and any falling below as Ketu.

The difference in means of weighted scores was .19576. The average variance of the two musical forms was estimated by dividing this value by the degree of freedom for the two groups (36). This yielded a variance of .005438, the variance of the difference between two means of twenty songs each. For this mean difference the standard error is .0233.

Dividing the mean difference (.19576) by its standard error (.0233) produced a t value of 8.39 which was significant well above the one percent level of confidence. This t value is greater than any of those obtained from differences in individual measurements. Use of the discriminant function, then, has reduced the probability of error in classifying samples of songs.

Here a question may be raised concerning the possible improvement in classifying a single song through application of the discriminant function. An estimate of the proportion of error in classifying single songs is provided by dividing one-half of the difference between means (.09788) by the standard deviation. The variance is .005438, which yields a standard deviation of .0737. The ratio of this number to one-half the mean difference is 1.33. A table of areas of the normal curve indicated that only 9 percent of the cases in a distribution will fall beyond this point. Such cases will be misclassified. When this error of 9 percent is compared with 31 percent, the estimated minimum error before combining scores, a substantial improvement is seen. It is reasonable to believe that further reduction in error may be accomplished merely by adding more variables. In the final analysis it should be possible to reduce error in classification to less than one in one hundred.

The present report has outlined a statistical device which might prove useful in meeting the problem of classifying data in cultural anthropology. As an example, a preliminary application of Fisher's discriminant function was made to data from the field of comparative musicology. Songs of Trinidad Rada and of Brazilian Ketu were compared with respect to use of major seconds, minor thirds and total intervals. Although it was possible to differentiate significantly between the music of these societies on the basis of some of the frequencies of interval use, application of the discriminat function markedly enhanced the probability of correct classification. After combining weighted scores, the probability of misclassifying a single song was reduced from about .31 to about .09.

This study has demonstrated that classification may be systematized and improved through application of a statistical technique. Further studies must employ larger samples and more variables, and in so doing significant measures may be isolated and classificatory problems solved. In this manner errors may be further reduced and questions of relatedness and derivation may be systematically attacked.

NOTES

[1] Descriptions of the *t*-test are available in any standard introductory statistics text. (*t* is not to be confused with T, which symbolizes total interval use in the present report.)

[2] Computational procedures for the discriminant function may be found in Fisher (1946), Garrett (1943), Hoel (1947), and Moroney (1953).

[3] A discussion of the solution of simultaneous equations may be found in any text in college algebra.

BIBLIOGRAPHY

ANONYMOUS
1954a Function and scope of the Human Relations Area Files, Inc. New Haven, Human Relations Area Files.
1954b 114th Annual Meeting of the American Statistical Association. American Statistician 8:32.

BARNARD, MILDRED M.
1935 The secular variations of skull characteristics in four series of Egyptian skulls. Annals of Eugenics 6:352–71.

BASCOM, WILLIAM R.
1944 The sociological role of the Yoruba cult-group. Menasha, American Anthropological Association Memoir 63.

BOAS, FRANZ
1927 Anthropology and statistics. *In* W. F. Ogburn and A. A. Goldenweiser, The social sciences and their interrelations. Boston, Houghton & Mifflin Co.

BRONOWSKI, J. and W. M. LONG
1952 Statistics of classification in anthropology. American Journal of Physical Anthropology 10 n.s.: 385–94.

CARNEIRO, EDISON
1936 Religioes Negras. Rio de Janeiro, Civilizacao Brasileira.

CARR, ANDREW
1953 A Rada community in Trinidad. Caribbean Quarterly 3:35–54.

CLEMENTS, FORREST E.
1954 Use of cluster analysis with anthropological data. American Anthropologist 56:180–99.

DENSMORE, FRANCES
1929 Papago music. Washington, Bureau of American Ethnology Bulletin 90.

DRIVER, H. E.
1953 Statistics in Anthropology. American Anthropologist 55:42–57.
DRIVER, H. E. and A. L. KROEBER
1932 Quantitative expression of cultural relationships. University of California Publications in American Archaeology and Ethnology 31:211–56.
FISHER, R. A.
1936 The use of multiple measurements in taxonomic problems. Annals of Eugenics 7:179–88.
1946 Statistical methods for research workers. New York, Hafner Publishing Company.
GARRETT, HENRY A.
1943 The discriminant function and its use in psychology. Psychometrika 8:65–79.
HERSKOVITS, MELVILLE J.
1938 Dahomey. New York, J. J. Augustin. 2 vols.
1949 Man and his works. New York, Alfred A. Knopf.
HERSKOVITS, MELVILLE J., and FRANCES S. HERSKOVITS
1943 The Negroes of Brazil. Yale Reiview 32:263–79.
HOEL, PAUL G.
1947 Introduction to mathematical statistics. New York, John Wiley and Sons.
HORNBOSTEL, ERICH M. VON
1921 Eine Tafel zur logarithmischen Darstellung von Zahlenverhaltnissen. Zeitschrift für Physik 6:29–34.
KOLINSKI, MIECZYSLAW
1936 Suriname music. *In* Melville J. Herskovits and Frances S. Herskovits, Suriname folklore. New York, Columbia Contributions to Anthropology, Vol. 27, pp. 491–760.
KUNST, JAAP
1950 Musicologica. Amsterdam. Koninklijke Vereeniging Indisch Institut.
MERRIAM, ALAN P.
1949 A transcription and analysis of Cheyenne Indian music. Unpublished MS.
1950 Notes on Cheyenne songs. Journal of the American Musicological Society 3:289–90.
1951 Songs of the Afro-Bahian cults: An ethnomusicological analysis. Unpublished Doctoral Dissertation, Northwestern University.
MERRIAM, ALAN P., SARAH WHINERY, and BENHART FRED
1954 Music of a Rada cult group. (Forthcoming in Anthropos)
MORONEY, M. M.
1953 Facts from figures. Melbourne, Penguin Books.
MURDOCK, GEORGE PETER
1949 Social structure. New York, Macmillan.
RAMOS, ARTHUR
1943 Introducao a antropologia Brasileira. Rio de Janeiro, Colecao Estudos Brasileiros
TYLOR, E. B.
1889 On a method of investigating the development of institutions, applied to laws of marriage and descent. Journal of the Royal Anthropological Institute 18:245–69.
WATERMAN, RICHARD A.
1952 African influence on the music of the Americas. *In* Sol Tax, Acculturation in the Americas. Chicago, University of Chicago Press.