DOCUMENT RESUME

ED 125 896 SE 020 873

AUTHOR Huberty, Carl J.; Perry, Marcia D.
TITLE Using Discriminant Analysis in the Homogeneous Grouping of High School Mathematics Students.
PUB DATE Apr 76
NOTE 14p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, California, April 19-23, 1976); Not available in hard copy due to marginal legibility of original document

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Achievement Tests; *Grouping (Instructional Purposes); *Homogeneous Grouping; Intelligence Tests; Mathematics Education; Predictive Validity; *Predictor Variables; *Research; Secondary Education; *Secondary School Mathematics
IDENTIFIERS Research Reports

ABSTRACT

Nine predictor measures available from student files were used in classifying seventh- and eighth-grade students in mathematics classes into one of three or four levels of instruction in the following year. These measures were scores on tests of language intelligence, nonlanguage intelligence, general intelligence, reading vocabulary, reading comprehension, arithmetic concepts, arithmetic applications, and teacher recommendation. At the end of the year, teachers were asked to judge the appropriateness of students placement. The sample consisted of 505 students from a large suburban junior high school in the southeast. A quadratic multivariate classification rule was used for both internal and external analyses. Internal discriminatory power was high; external power dropped for across-groups classification but not appreciably for separate large groups. One predictor measure, based on teacher recommendation, proved to be about as accurate as all nine, in an external sense. (Author/SD)

Using Discriminant Analysis in the Homogeneous Grouping of High School Mathematics Students

Carl J Huberty and Marcia D. Perry

University of Georgia

Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April, 1976.

Abstract

Nine predictor measures were used in classifying seventh and eighth grade students in mathematics classes into one of three or four levels of instruction in the following year. The sample consisted of 505 students from a large suburban junior high school in the southeast. A quadratic multivariate classification rule was used for both internal and external analyses. Internal discriminatory power was high: external power dropped for across-groups classification but not appreciably for separate large groups. One predictor measure, based on teacher recommendation, proved to be about as accurate as all nine, in an external sense.

Using Discriminant Analysis in the Homogeneous Grouping of High School Mathematics Students

In the practice of ability (or homogeneous) grouping, placement of students at different levels of instruction has typically been accomplished on the basis of three available measures, either singly or in combination. These are measures of achievement, measures of aptitude (generally intelligence test scores), and teacher recommendations. (See Findley, 1973, and Wick & Beggs, 1971, for elaboration.) If these measures are used in combination, traditional practice has been to base the weights for the measure on subjective judgment.

The purpose of this paper was to develop a formal quantitative method of determining differential weights of some typically available measures for assigning seventh and eighth grade mathematics students to one of four levels of instruction in the succeeding grades. Real data were used to assess the effectiveness of the method.

Method

Subjects

The children considered in this study attended a large junior high in a suburb of Athens, Georgia. Four classes of students were used: Class I completed seventh grade mathematics in the spring of 1970 and continued in eighth grade mathematics during the 1970-71 school year. This class had a total of 160 children (72 boys and 88 girls) of which 134 were white. Class II completed eighth grade mathematics in the spring of 1970 and continued in ninth grade mathematics during the 1970-71 school year. This class had a total of 89 children (36 boys and 53 girls) of which 83 were white. Class III completed seventh grade mathematics in the spring of 1971 and continued in eighth grade mathematics during the 1971-72 school year. This class had a total of 137 children (62 boys and 75 girls) of which 111 were white. Class IV completed eighth grade mathematics in the spring of 1971 and continued in ninth grade mathematics during the 1971-72 school year. This class had a total of 134 children (47 boys and 87 girls) of which 89 were white.

Variables

Prior to the collection of data, potential predictors of membership in one of the four levels of mathematics instruction were specified. Measures on the following variables were available from student files: language intelligence (LIQ), nonlanguage intelligence (NLIQ), general intelligence (GIQ), reading vocabulary (ACOM), arithmetic concepts (ACPT), and arithmetic application (AAPP). Intelligence measures were based on the California Test of Mental Maturity. The achievement tests used were the Comprehensive Tests of Basic Skills. At the end of the school year (spring 1970 for classes I and II, and spring 1971 for classes III and IV) each seventh and eighth grade teacher was asked to make a recommendation, based on her insight into the mathematical talents of each of her students, on the level of mathematics instruction that each student ought to receive the following school year. Their response became the measure of the placement variable (PLCM). This gave a total of nine predictors: PLCM, LIQ, NLIQ, GIQ, RVCB, RCOM, ACOM, ACPT, and AAPP.

At the end of the following school year (spring 1971 for classes I and II, and spring 1972 for classes III and IV), each student's eighth or ninth grade teacher was asked to make a judgment regarding the appropriateness for the student of the level of course instruction in mathematics which he had pursued during the year just ending. This judgment was made by taking into account all of the experiences accumulated as a result of working with the students during the year. Four recommended levels of instruction determined the criterion groups where level I was the lowest level.

Data Analyses

There were only seven children in the second level of instruction for class II and only eight in the fourth level for class III. Since subsequent analyses required that the number of subjects in each level be at least one more than the number of predictors (nine), these 15 children were excluded from further consideration.

It was decided that four sets of analyses, one for each class would be carried out. Preliminary univariate analyses of variance with either three or four groups were conducted to identify measures which would not show any promise of contributing ($F<1.00$) to multivariate separation of the groups corresponding to the four levels of instruction. Univariate F values for all nine measures for the four classes were all greater than 6.87; hence all nine measures were retained for the final analyses.

Assuming multivarate normality of the nine predictor measures in the four populations, the condition of equal population covariance matrices was assessed via both a chi-square and an F statistic. Wilks's lambda statistic was used as an index of group separation. Multivariate distance measures were used to assess pairwise group proximity. The "ordering" of the groups in terms of distances was used to detect "second-order" misclassifications -- where a student was classified into a level of instruction nonadjacent to his actual level. An attempt was also made to sort out the best and poorest predictors, in terms of contribution to group separation. (See Huberty, 1975).

Classification procedures were used to assess the predictive accuracy of the total set and subsets of discriminators. Both "internal" and "external" classification results were considered. Results of an internal classification analysis are those obtained when measures for the students, on whom the basic statistics (mean vectors and covariance matrices) were determined, are resubstituted to obtain the values for the classification rules. In an external classification analysis statistics based on one set of students are used in classifying "new" students. The external classification method used in this study is an extension of that proposed by Lachenbruch (1967). The procedure for the Lachenbruch method is as follows: Compute the statistics for each of the possible total samples of size $\Sigma N_k - 1$ obtained by omitting one student's vector of measures from the original total sample, and record for each computation whether the omitted student is misclassified. (N_k = number of students in k th criterion group.)

The computer program used was one developed by the first author. This program yields linear and quadratic classification results--both internal and external analyses--as well as the usual values of means, covariance matrices, distances, test statistics, and indices for discrimination.

Results

Means, standard deviations, univariate ANOVA mean-square ratios, and within-groups predictor intercorrelations were determined for each class. Tables of these values are available upon request.

An F statistic used to test the equality of population covariance matrices yielded significance ($p<.05$) for all four classes. The four values of Wilks's lambda statistic, in light of the apparent inequality of covariance matrices, were fairly low (.24, .36, .33, and .15), while all univariate F-values were fairly large. This information lead us to believe that the criterion groups were separated to some degree. Distances-like measures -- Mahalanobis D^2 values adjusted for unequal covariances -- supported the ordering of the levels of instruction.

Since the equal covariances condition was judged untenable, a quadratic classification rule was used (see Cooley & Lohnes, 1971, p. 268). Percentages of correct classification yielded by both the internal and external analyses using all nine variables are given in Table I.

Insert Table 1 about here

For the internal analysis the across-groups percentage of correct classifications is very respectable for each class -- average percentage greater than 84. Although these percentages drop for the external analyses, the average percentage is still nearly 64. Separate group internal classification accuracy was quite high in only three levels was the percentage of correct classifications less than 75: level 3 of class III (72), level 3 of class IV (67), and level 2 of class I (50). The separate group accuracy is still quite high for the external analyses on the larger groups of each class: (a) the percentages (81 and 84) for levels 1 and 3 in class I where the group size was 57, (b) the percentage (71) for level 3 of class II where the group size was 41, (c) the percentages (70 and 65) for levels 2 and 3 of class III and (d) the percentage (76) for level 2 of class IV.

As further evidence of the respectability of the classification results of the second-order misclassifications were examined. The internal analysis yielded only seven (4%) such misclassifications, and 13 (8%) by the external analysis for class I. For class II the corresponding misclassifications were 2 (3%) and 9 (11%); for class III the corresponding misclassifications were 0 to 1 (less than 1%); for class IV the corresponding misclassifications were 1 (1%) and 5 (4%).

The nine variables were clustered for the purpose of assessing the relative discriminatory power of the subsets. The placement variable (PLCM) was considered singly; the two reading variables, RVCB and RCOM, comprised a second cluster; the three arithmetic variables, ACOM, ACPT, and AAPP, comprised a third cluster; and the three intelligence variables, LIQ, NLIQ, and GIQ, comprised the fourth cluster. The classification accuracy of the various sets of variables is indicated in Table 2.

Across all levels of instruction and across the four classes, external classification accuracy yielded by PLCM was about as high,

---

Insert Table 2 about here

---

if not higher, than any of the other three clusters of variables as well as that yielded by all nine variables. For classes III and IV, however, the arithmetic variables, and, for class III the intelligence variables, performed about as well as PLCM alone. A cluster comprised of all variables except for PLCM also performed quite well for classes III and IV.

For separate-group external classification accuracy, PLCM again performed about as well, if not better, than any other cluster, including all nine variables, with two exceptions. Both exceptions were for groups with the smallest frequencies: level 2 of class I and level 3 of class IV.

For neither across-group nor separate-group external classification was there a clear cut ordering of the clusters of variables in terms of relative discriminatory power. The 18 (four across-group plus 14 separate groups), possible sets of rankings yielded a coefficient of concordance of only 0.31. Overall, however, as might be expected, the arithmetic cluster was judged to be the best cluster, second to PLCM; the reading cluster was judged to be the worst.

A complete study of the rank-ordering of the nine variables was not undertaken. Although various ordering methods have been used and proposed for the linear case (Huberty & Smith, 1976) (i.e., with identical underlying covariance structure across populations), little work has been done for the quadratic case (Lachenbruch, 1975). One approach for the current nine-variable situation would be to perform analyses for $2^9 - 2 = 510$ proper subsets of variables. Even with the use of a computer it would be very expensive to obtain complete external classification results.

Discussion

It was encouraging to find that the power of the nine (one teacher judgment, three intelligence, two reading, and three arithmetic) measures considered in this study to discriminate groups of children in different levels of instruction in matematics was respectable. It might appear that nearly all of the important measurable attributes of seventh and eighth grade mathematics students were included. For all of the classes inv investigated, the teacher judgment measure (PLCM) was one of the best, if not the best, predictor. When classification was carried out with PLCM as the only predictor, it was found that the internal analysis results were somewhat poorer than when all nine measures were used. However, the percentages of correct classifications yielded by the external analysis, using PLCM alone were about the same as those yielded by the internal analysis, and, interestingly, a little better than the external results using all nine measures. That the PLCM measure turned out to be such a good predictor by itself it not surprising since the correlations between it and each of the other eight measures were low to moderate and were all positive (Cochran, 1964). This finding supports the conjecture of Huberty and Curry (1975) that reducing the number of predictors to include only the "good" ones will increase the external classification accuracy. An implication of this reduction to a single predictor may be that we should have rocognized all along that teacher intuition and judgment is as good as, if not better than, the use of multiple objective test scores in making predictions about student performance and benefits from different levels of instruction. In the absence of teacher recommendations for placement, a combination of arithmetic measures such as those employed here could be used.

References

Cochran, W. G. On the performance of the linear discriminant function. Technometrics, 1964, 6, 179-190.

Cooley, W. W., & Lohnes, P. R. Multivariate data analysis. New York: Wiley, 1971.

Findley, W. G. How ability grouping fails. Inequality in Education, Number 14, July 1973, pp. 37-40.

Huberty C. J. Discriminant Analysis. Review of Educational Research, 1975, 45, 543-598.

Huberty, C. J, & Curry, A. R. Linear versus Multivariate Classification. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, April 1975.

Huberty, C. J, & Smith, D. U. Variable contribution in discriminant analysis. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April 1976.

Lachenbruch, P. A. An almost unbiased method of obtaining confidence interval for the probability of misclassification in discriminat analysis. Biometrics 1967, 23, 639-645.

Lachenbruch, P. A. Some unsolved problems in discriminant analysis. Institute of Statistics, University of North Carolina, Mimeo Series No. 1050, December 1975.

Wick, J. W., & Beggs, D. L. Evaluation for decision-making in the schools. Boston: Houghton-Mifflin, 1971.

Table 1

Percentages of Correct Classifications
Using All Nine Variables

Class	Instructional Level: [illegible]	2	3	4	Total
I	88/81 (57)	50/06 (18)	95/84 (57)	89/50 (28)	86/68 (160)
II	90/62 (21)	----	88/71 (41)	75/40 (20)	85/61 (82)
III	100/27 (15)	89/70 (57)	72/65 (57)	----	83/63 (129)
IV	86/70 (37)	84/76 (49)	67/19 (21)	93/67 (27)	84/63 (134)

Note. Internal percentages are to left of tilted line, external to right. N-values are given in parentheses.

TABLE 2

PERCENTAGES OF CORRECT CLASSIFICATIONS

USING VARIOUS VARIABLE SETS

	Level 1 (Int.)	Level 1 (Ext.)	Level 2 (Int.)	Level 2 (Ext.)	Level 3 (Int.)	Level 3 (Ext.)	Level 4 (Int.)	Level 4 (Ext.)	Total (Int.)	Total (Ext.)
INSTRUCTIONAL LEVELS	1		2		3		4		TOTAL	
CLASS I	75	75	00	00	82	82	68	68	68	68
	60	53	00	00	51	40	61	33	50	46
	79	77	06	00	56	54	57	57	59	57
	65	54	17	17	56	51	57	50	55	48
	84	72	33	00	72	54	75	57	73	55
	88	81	50	06	95	84	89	50	86	68
	N = 57		N = 18		N = 57		N = 28		N = 160	
CLASS II	86	86			89	89	40	40	76	76
	43	38			66	63	25	20	51	46
	62	43			66	51	60	40	63	48
	57	38			66	61	55	35	61	49
	81	57			68	57	65	35	71	51
	90	62			88	71	75	40	85	61
	N = 21				N = 41		N = 20		N = 82	
CLASS III	93	93	74	74	60	60			70	70
	53	53	67	63	54	51			60	57
	67	60	81	75	60	56			70	65
	60	53	66	61	70	68			67	64
	100	33	86	70	75	61			83	62
	100	27	89	70	72	65			83	63
	N = 15		N = 57		N = 57				N = 129	
CLASS IV	86	86	53	53	00	00	89	89	61	61
	51	49	53	51	24	24	85	85	54	53
	73	68	78	76	10	10	81	81	66	65
	62	59	51	45	14	00	77	63	54	46
	78	65	84	76	57	19	93	70	80	63
	86	70	84	76	67	19	93	67	84	63
	N = 37		N = 49		N = 21		N = 27		N = 134	

NOTE. IN EACH CELL THE PROPORTIONS TO THE LEFT ARE FROM THE INTERNAL ANALYSIS; THOSE TO THE RIGHT, FROM THE EXTERNAL ANALYSIS. THE PROPORTIONS FROM TOP TO BOTTOM ARE BASED ON PLACEMENT ONLY, READING VARIABLES (RCVM & RCOM), ARITHMETIC VARIABLES (ACOM, ACPT, & AAPP), INTELLIGENCE VARIABLES (LIQ, NLIQ, & GIQ), ALL VARIABLES EXCEPT PLCM, AND ALL VARIABLES.