DOCUMENT RESUME

ED 400 280 TM 025 548

AUTHOR	Meshbane, Alice; Morris, John D.
TITLE	Predictive Discriminant Analysis Versus Logistic Regression in Two-Group Classification Problems.
PUB DATE	8 Apr 96
NOTE	25p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, April 8-12, 1996).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Classification; *Comparative Analysis; *Group Membership; *Predictor Variables
IDENTIFIERS	Covariance Matrices; Cross Validation; *Logistic Regression; *Predictive Discriminant Analysis

ABSTRACT

A method for comparing the cross-validated classification accuracies of predictive discriminant analysis and logistic regression classification models is presented under varying data conditions for the two-group classification problem. With this method, separate-group, as well as total-sample proportions of the correct classifications, can be compared for the two models. The test for contrasting correlated proportions developed by Q. McNemar (1947) is used in the statistical comparisons of the separate-group data and total-sample proportions. The method is illustrated with 32 real data sets that varied in number of cases, relative group sizes, number of predictor variables, degree of group separation, and equality of group covariance matrices. (Contains 1 table and 24 references.) (SLD)

Predictive Discriminant Analysis Versus Logistic Regression in Two-Group Classification Problems

Alice Meshbane and John D. Morris

Florida Atlantic University

Paper presented at the annual meeting of the American Educational Research Association, April 8, 1996, New York, NY

Predictive Discriminant Analysis Versus Logistic Regression in Two-Group Classification Problems

ABSTRACT. A method for comparing the cross-validated classification accuracies of predictive discriminant analysis and logistic regression classification models is presented under varying data conditions for the two-group classification problem. With this method, separate-group as well as total-sample proportions of correct classifications can be compared for the two models. McNemar's test for contrasting correlated proportions is used in the statistical comparisons of the separate-group and total-sample proportions. The method is illustrated with 32 real data sets.

Among the methods used for solving two-group classification problems, logistic regression (LR) and predictive discriminant analysis (PDA) are two of the most popular (Yarnold, Hart, & Soltysik, 1994, p. 73). Unlike PDA, LR captures the probabilistic distribution embedded in dichotomous measures, avoids violations to the assumption of homogeneity of variance, and does not require strict multivariate normality (e.g., see Aldrich & Nelson, 1984; Cox & Snell, 1989). Therefore, when PDA assumptions are violated, we might theoretically expect higher cross-validated classification hit rate accuracy with LR than PDA.

Although several studies have compared the classification accuracy of LR and PDA, the results have been inconsistent. For example, results of three simulation studies (Barón, 1991; Bayne, Beauchamp, Kane, & McCabe, 1983; Crawley, 1979) suggest that LR is more accurate than PDA for nonnormal data. However, several researchers (e.g., Cleary & Angel, 1984; Dey & Astin, 1993; Knoke, 1982; Krzanowski, 1975; Press & Wilson, 1978) using nonnormal real data found little or no difference in the accuracy of the two techniques. Findings are also inconsistent for degree of group separation. Bayne, et al. (1993) found that larger group separation favored PDA, while Crawley (1979) found this condition to favor LR. Sample size is yet another data condition yielding inconsistent results. In a simulation study, Harrell and Lee (1985) found that PDA was more accurate than LR for small samples. By contrast, in a study by Johnston and Seshia (1992) using real data, LR worked better than PDA for small samples.

Given these inconsistencies in the literature, it is not clear which of the two methods will work better for a given data set. Consequently, a method for comparing the accuracy of PDA and LR for a specific data set will enable researchers to select the optimal classification procedure for that data set. In this paper we describe a method for determining the superior classification model for a specific data set, regardless of data conditions. In addition, a computer program that accomplishes the method is introduced and demonstrated.

Method

Data Sources. We used 32 classification data sets varying in number of cases, relative group sizes, number of predictor variables, degree of group separation, and equality of group covariance matrices to illustrate the method. To bolster validity, all were taken from real classification studies. The sources include journal articles, paper presentations and research texts.

Procedure. For PDA, we built linear classification functions based on assumptions of multivariate normality, equal covariance matrices, and equal prior probabilities of group membership. We classified cases into groups using Tatsuoka's (1988) minimum chi square rule. For LR, we used the elegant IMSL subroutine CTGLM, conveniently available with the powerful new 32-bit Microsoft Fortran v4.0 Powerstation, to obtain model coefficients. The CTGLM routine uses a standard nonlinear approximation technique (Newton-Raphson) to determine maximum likelihood estimates of model coefficients. We classified each case into the group with the highest log-likelihood probability.

In comparing the predictive accuracy of the PDA model to that of the LR model, we considered external rather than internal results. Results of an internal classification analysis are those obtained when measures for the individuals on whom the statistics were based are resubstituted to obtain the predicted classification scores. In an external classification analysis statistics based on one set of individuals are used in classifying new individuals. An external analysis is appropriate for making inferences about the discriminatory power of the predictors for a new set of data (Huberty, 1994).

We estimated external, or cross-validated, hit-rate accuracy using the leave-one-out procedure. A case is classified by applying the model derived from all cases except the one being classified. This process was repeated round-robin for each case with a count of the overall classification accuracy used to estimate the cross-validated accuracy. This procedure has a relatively wide following in the discriminant analysis literature (see, for example, Huberty, 1994; Huberty & Mourad, 1980; Lachenbruch, 1967; Mosteller & Tukey, 1968).

We compared separate-group as well as total-sample proportions of correct classification for the PDA and LR models. We used McNemar's (1947) test for contrasting correlated proportions in the statistical comparisons between PDA and LR models for the separate-group and total-sample proportions. This method was previously suggested for comparing full and reduced classification methods (Morris & Huberty, 1995; Morris & Meshbane, 1995) as well as for comparing linear and quadratic classification models (Meshbane & Morris, 1995), but is equally applicable in comparing PDA and LR models. (See Looney, 1988, for a method of comparing classification results of more than two models.) Because the calculation of the McNemar correlated proportion statistic requires the joint distribution of hits and misses for both the PDA and LR models, no statistical package will accomplish the method. Therefore, we wrote a FORTRAN computer program to provide this information.

We used the Box test for testing the assumption of homogeneity of covariance structures. This test is sensitive to multivariate normality, and the outcome is therefore confounded with the homogeneity of dispersion issue. Nevertheless, the Box test is routinely used for testing the homogeneity of dispersion assumption and is even the default in some statistical packages. Notwithstanding concerns over this test, once could argue that, theoretically, a logistic classification model is more likely to be appropriate when the Box test indicates that the covariance structures are unequal.

Results and Discussion

For each of the data sets, Table 1 gives a short description, the degree of group separation (D), the number of cases in group 1 (n_1), the number of cases in group 2 (n_2), an index of disproportionality of the group sizes (I) calculated as (n_i * 100) / n_j, where n_i is the larger of the two groups, the number of predictor variables (p), results of the Box test for homogeneity of covariance structures, and a comparison of the leave-one-out performance of the PDA and LR models for each group separately and for the total sample. We compared the performance of the two classification models, displayed as the hit-rate percentage obtained by the p predictor variables, via McNemar's test for contrasting correlated proportions.

To make an inferential decision using this method, the researcher must, as is customary, choose an alpha level; the choice of alpha level results in an associated critical z statistic. To illustrate the method for these data sets, we used the .01 alpha level with the associated z of 2.58.

In the first three data sets, maximum likelihood estimates of logistic regression parameters could not be calculated due to complete separation of the data (see Table 1). Among the remaining 29 data sets, differences between the PDA and LR models in classifying the total sample were not statistically significant, with the exception of

Data Set 26 (see Table 1). Here, the LR model yielded a significantly higher total hit rate.

Insert Table 1 about here

Differences between the two classification models in separate-group hit rates were statistically significant in nine of the 32 data sets (9, 16, 17, 23, 24, 26, 30, 31, and 32). In eight of these data sets, superior performance of the LR model in classifying the larger group was offset by superior performance of the PDA model in classifying the smaller group; the only exception was Data Set 31, in which the significant advantage of the LR model for the larger group was offset by a nonsignificant advantage of the PDA model for the smaller group.

Statistically significant differences in separate-group hit rates were found in data sets with moderate to relatively large discrepancies in sample sizes and small to moderate group separation, but not in any data set with similar sample sizes (I < 118) nor in any data set with relatively large group separation (D > 2.0). This indicates that CTGLM uses sample sizes as estimates of population sizes when generating maximum likelihood estimates of LR model parameters. Using sample sizes as estimates of population sizes is inappropriate, however, when population sizes are unknown or when sample sizes are not proportional to population sizes (Huberty, 1994, p. 65). Consequently, we decided to force the assumption of equal population sizes by adjusting the LR model by a constant. We determined the value of the constant by referring to the FORTRAN program LOGDIS (Albert & Harris, 1987), which uses an iterative Newton-Raphson procedure to obtain maximum likelihood estimates of LR model parameters for the k-group classification problem. Under the assumption of equal population sizes, there were no statistically significant differences in total-sample or separate-group hit rates between LR and PDA for any of the 29 data sets for which maximum likelihood estimates of LR model parameters could be calculated (results available on request).

Therefore, from the perspective of these data sets, which were selected to portray a wider range of characteristics than were previously available, some evidence can be tentatively gleaned. For total-group accuracy, hit rates for the LR and PDA models were the same in 28 of the 29 data sets for which LR model estimates could be calculated. Neither theoretical nor data-based considerations were helpful in predicting which of the two models would work better.

In reference to separate-group accuracy the results are a bit more complicated. For separate-group hit rate to be of interest the researcher must make an a priori decision that accuracy in one group is more important than in the other. For example, the researcher may decide that, in predicting high school dropouts, to be correct for the dropouts is more important than for a persister group. These data suggest that when the size of one population is much larger than the other, the researcher may improve separate-group hit rate by choosing the LR model if interest is in classification accuracy of the larger group, and by choosing the PDA model if interest is in classification accuracy of the smaller group.

It certainly may be that there are other data sets in which LR or PDA would be judged significantly superior for total-group as well as separate-group hit rates; from the results of these analyses, it seems quite possible that there may be data that manifest a PDA rule that is significantly superior for one group and an LR rule that is significantly superior for the other. In this case, if the researcher has interest in separate group accuracy, knowledge of these results would allow selection of a rule depending on which group is of highest superiority. Use of the method and computer program demonstrated herein will allow such decisions to be made based on explicit cross-validated classification accuracies.

Note

For a copy of the FORTRAN program that accomplishes the method, send a returnable diskette and diskette mailer to Alice Meshbane, College of Education, Florida Atlantic University, P.O. Box 3091, Boca Raton, FL 33431-0991. Internet: Meshbane@acc.fau.edu.

References

Albert, A., & Harris, E. K. (1987). Multivariate interpretation of clinical laboratory data. New York: Marcel Dekker, Inc.

Aldrich, J. H. & Nelson, F. D. (1984). Linear probability, logit, and probit models. Newbury Park, CA: Sage.

Barón, A. E. (1991). Misclassification among methods used for multiple group discrimination - The effects of distributional properties. Statistics in Medicine, 10, 757-766.

Bayne, C. K., Beauchamp, J. J., Kane, V. E., & McCabe, G. P. (1983). Assessment of Fisher and logistic linear and quadratic discrimination models. Computational Statistics and Data Analysis , 1, 257-273.

Cleary, P. D. & Angel, R. (1984). The analysis of relationships involving dichotomous dependent variables. Journal of Health and Social Behavior, 25, 334-348.

Cox, D. R. & Snell, E. J. (1989). The analysis of binary data (2nd ed). London: Chapman and Hall.

Crawley, D. R. (1979). Logistic discriminant analysis as an alternative to Fisher's linear discriminant function. New Zealand Statistics, 14(2), 21-25.

Dey, E. L. & Astin, A. W. (1993). Statistical alternatives for studying college student retention: A comparative analysis of logit, probit, and linear regression. Research in Higher Education, 34, 569-581.

Harrell, F. E. Jr. & Lee, K. L. (1985). A comparison of the discrimination of discriminant analysis and logistic regression under multivariate normality. In P. K. Sen (Ed.), Biostatistics: Statistics in biomedical, public health and environmental sciences (pp. 333-343). Amsterdam: North Holland.

Huberty, C. J. (1994). Applied discriminant analysis. New York: Wiley.

Huberty, C. J & Mourad, S. A. (1980). Estimation in multiple correlation/prediction. Educational and Psychological Measurement, 40, 101- 112.

Johnston, B. & Seshia, S. S. (1992). Discriminant analysis when all variables are ordered. Statistics in Medicine, 11, 1023-1032.

Knoke, J. D. (1982). Discriminant analysis with discrete and continuous variables. Biometrics, 38, 191-200.

Krzanowski, W. J. (1975). Discrimination and classification using both binary and continuous variables. Journal of the American Statistical Association, 70, 782-790.

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

Looney, S. W. (1988). A statistical technique for comparing the accuracies of several classifiers. Pattern Recognition Letters, 8, 5-9.

McNemar, Q. (1947). Note on the sampling error of the differences between correlated proportions or percentages. Psychometrika, 12, 153-157.

Meshbane, A. & Morris, J. D. (1995). A method for selecting between linear and quadratic classification models in discriminant analysis. The Journal of Experimental Education, 63, 263-273.

Morris, J. D. & Huberty, C. J. (1995). Full versus restricted model testing in discriminant analysis. Journal of Experimental Education, 63, 161-165.

Morris, J. D. & Meshbane, A. (1995). Selecting predictor variables in two-group classification problems. Educational and Psychological Measurement, 55, 438-441.

Mosteller, F. & Tukey, J. W. (1968). Data analysis, including statistics. In G. Lindzey & E. Aronson (Eds.), Handbook of Social Psychology: Vol. 2. (pp. 80-203). Reading, MA: Addison-Wesley.

Press, S. J. & Wilson, S. (1978). Choosing between logistic regression and discriminant analysis. Journal of the American Statistical Association, 73, 699-705.

Tatsuoka, M. J. (1988). Multivariate analysis: Techniques for educational and psychological research (2nd ed.). New York: Macmillan.

Yarnold, P. R., Hart, L. A., & Soltysik, R. C. (1994). Optimizing the classification performance of logistic regression and Fisher's discriminant analysis. Educational and Psychological Measurement, 54, 73-85.

Table 1

Data Set Description, Results of Box M Test for Equality of Covariance Matrices, and Comparison of Hit Rate Percents for PDA and LR Models

#	Data Set Description	D	I	n_1	n_2	p	Results of Box M Test for Equal Covariance Matrices	Model Used	L-O-O Hit Rate % Total	GR 1	GR 2
1	Fisher Data - Groups 1 & 3	13.97	100	50	50	4	$\chi^2 = 6.9057$, $p < .0001$	PDA	100	100	100
							$df = 10$	LR	*	*	*
								McNemar's z			
2	Fisher Data - Groups 1 & 2	10.16	100	50	50	4	$\chi^2 = 5.0455$, $p < .0001$	PDA	100	100	100
							$df = 10$	LR	*	*	*
								McNemar's z			
3	Bisbey Data - Groups 1 & 3	5.12	106	35	37	13	$\chi^2 = .9939$, $p = .5013$	PDA	97	94	100
							$df = 91$	LR	*	*	*
								McNemar's z			
4	Fisher Data - Groups 2 & 3	3.77	100	50	50	4	$\chi^2 = .7148$, $p = .7125$	PDA	93	92	94
							$df = 10$	LR	91	92	90
								McNemar's z	.82	.00	1.41

* Due to complete separation of the data, maximum likelihood estimates of LR model parameters could not be calculated.

Table 1, cont.

Data Set Description, Results of Box M Test for Equality of Covariance Matrices, and Comparison of Hit Rate Percents for PDA and LR Models

#	Data Set Description	D	I	n_1	n_2	p	Results of Box M Test for Equal Covariance Matrices	Model Used	L-O-O Hit Rate % Total	GR 1	GR 2
5	Junior Faculty Performance Data	3.11	152	27	41	4	$\chi^2 = 2.2172$, $p = .0150$	PDA	91	93	90
							$df = 10$	LR	93	89	95
								McNemar's z	-.58	1.00	-1.41
6	Rulon Data - Groups 1 & 3	2.93	129	85	66	4	$\chi^2 = 3.4973$, $p = .0003$	PDA	93	94	91
							$df = 10$	LR	91	93	89
								McNemar's z	1.41	1.00	1.00
7	Bisbey Data - Groups 1 & 2	2.89	231	35	81	13	$\chi^2 = 1.0021$, $p = .4777$	PDA	89	89	89
							$df = 91$	LR	88	83	91
								McNemar's z	.58	1.41	-1.00
8	Bisbey Data - Groups 2 & 3	2.41	219	81	37	13	$\chi^2 = 1.2131$, $p = .0929$	PDA	84	83	87
							$df = 91$	LR	86	89	78
								McNemar's z	-.63	-1.89	1.73

Table 1, cont.

Data Set Description, Results of Box M Test for Equality of Covariance Matrices, and Comparison of Hit Rate Percents for PDA and LR Models

#	Data Set Description	D	I	n_1	n_2	p	Results of Box M Test for Equal Covariance Matrices	Model Used	L-O-O Hit Rate % Total	GR 1	GR 2
9	Talent Data - Groups 3 & 5	1.97	285	33	94	14	$\chi^2 = 1.1086$, $p = .2238$	PDA	77	79	77
							$df = 105$	LR	79	58	86
								McNemar's z	-.50	2.65	-3.00
10	Demographic # 2 - Body Char	1.88	129	157	122	8	$\chi^2 = 6.6870$, $p < .0001$	PDA	82	83	82
							$df = 36$	LR	81	83	77
								McNemar's z	1.89	-1.00	2.45
11	Rulon Data - Groups 2 & 3	1.87	141	93	66	4	$\chi^2 = 3.4962$, $p = .0003$	PDA	83	85	80
							$df = 10$	LR	82	87	76
								McNemar's z	.45	-1.41	1.73
12	Rulon Data - Groups 1 & 2	1.74	109	85	93	4	$\chi^2 = 4.9003$, $p < .0001$	PDA	81	84	79
							$df = 10$	LR	80	80	81
								McNemar's z	.45	1.73	-1.41

Table 1, cont.

Data Set Description, Results of Box M Test for Equality of Covariance Matrices, and Comparison of Hit Rate Percents for PDA and LR Models

#	Data Set Description	D	I	n_1	n_2	p	Results of Box M Test for Equal Covariance Matrices	Model Used	L-O-O Hit Rate % Total	GR 1	GR 2
13	Talent Data - Groups 1 & 5	1.72	113	83	94	14	χ^2 = 1.5086, p = .0014	PDA	75	74	76
							df = 105	LR	74	74	75
								McNemar's z	1.00	.00	-1.00
14	Demographic # 3 - Body Char	1.36	104	142	137	8	χ^2 = 5.2724, p < .0001	PDA	73	70	76
							df = 36	LR	73	72	74
								McNemar's z	.38	-1.41	1.34
15	Press & Wilson - 50 States	1.31	100	25	25	5	χ^2 = 2.6544, p = .0008	PDA	64	68	60
							df = 15	LR	70	72	68
								McNemar's z	-1.13	-.58	-1.00
16	Business Sch Perf Data - 1990	.89	188	147	78	7	χ^2 = 1.5719, p = .0300	PDA	64	61	69
							df = 28	LR	67	83	37
								McNemar's z	-1.05	-5.74	5.00

Table 1, cont.

Data Set Description, Results of Box M Test for Equality of Covariance Matrices, and Comparison of Hit Rate Percents for PDA and LR Models

#	Data Set Description	D	I	n_1	n_2	p	Results of Box M Test for Equal Covariance Matrices	Model Used	L-O-O Hit Rate % Total	GR 1	GR 2
17	Talent Data - Groups 1 & 3	.89	252	83	33	14	$\chi^2 = .9401$, $p = .6493$	PDA	58	64	42
							$df = 105$	LR	62	81	15
								McNemar's z	-1.04	-3.74	3.00
18	Block Data - Groups 3 & 4	.85	100	38	38	4	$\chi^2 = 4.3098$, $p < .0001$	PDA	68	74	63
							$df = 10$	LR	68	71	66
								McNemar's z	.00	.58	-1.00
19	Block Data - Groups 1 & 2	.84	105	40	38	4	$\chi^2 = 2.2028$, $p = .0157$	PDA	68	60	76
							$df = 10$	LR	69	65	74
								McNemar's z	-.58	-1.41	1.00
20	Block Data - Groups 1 & 4	.81	105	40	38	4	$\chi^2 = 1.5500$, $p = .1163$	PDA	60	58	63
							$df = 10$	LR	58	58	58
								McNemar's z	1.41	.00	1.41

Table 1, cont.

Data Set Description, Results of Box M Test for Equality of Covariance Matrices, and Comparison of Hit Rate Percents for PDA and LR Models

#	Data Set Description	D	I	n_1	n_2	p	Results of Box M Test for Equal Covariance Matrices	Model Used	L-O-O Hit Rate % Total	GR 1	GR 2
21	Higher Ed Persistence Data	.76	102	263	269	11	χ^2 = .9860, p = .5101	PDA	62	62	61
							df = 10	LR	61	61	61
								McNemar's z	1.13	2.00	-.58
22	Block Data - Groups 1 & 3	.74	103	40	39	4	χ^2 = 5.3857, p < .0001	PDA	65	58	72
							df = 10	LR	62	58	67
								McNemar's z	1.41	.00	1.41
23	MBA Success Data - 1986	.72	144	64	92	7	χ^2 = 1.2640, p = .1632	PDA	64	66	62
							df = 28	LR	58	36	73
								McNemar's z	1.67	4.36	-3.16
24	Warncke Data - Groups 1 & 3	.69	163	65	40	10	χ^2 = 1.5335, p = .0086	PDA	57	62	50
							df = 55	LR	60	79	30
								McNemar's z	-.69	-3.32	2.83

Table 1, cont.

Data Set Description, Results of Box M Test for Equality of Covariance Matrices, and Comparison of Hit Rate Percents for PDA and LR Models

#	Data Set Description	D	I	n_1	n_2	p	Results of Box M Test for Equal Covariance Matrices	Model Used	L-O-O Hit Rate % Total	GR 1	GR 2
25	MBA Success Data - 1989	.69	114	66	75	7	$\chi^2 = .6058$, $p = .9480$	PDA	60	61	59
							$df = 10$	LR	59	53	64
								McNemar's z	.33	2.23	-2.00
26	Business Sch Perf Data - 1989	.66	246	160	65	7	$\chi^2 = .9730$, $p = .5060$	PDA	56	54	63
							$df = 28$	LR	69	94	08
								McNemar's z	-2.80	-8.00	6.00
27	Block Data - Groups 2 & 3	.64	105	37	39	4	$\chi^2 = 4.5542$, $p < .0001$	PDA	55	57	54
							$df = 10$	LR	57	54	59
								McNemar's z	-.58	1.00	-1.41
28	Block Data - Groups 2 & 4	.52	103	37	38	4	$\chi^2 = 1.2033$, $p = .2838$	PDA	59	62	55
							$df = 10$	LR	57	65	50
								McNemar's z	.00	.00	.00

Table 1, cont.

Data Set Description, Results of Box M Test for Equality of Covariance Matrices, and Comparison of Hit Rate Percents for PDA and LR Models

#	Data Set Description	D	I	n_1	n_2	p	Results of Box M Test for Equal Covariance Matrices	Model Used	L-O-O Hit Rate % Total	GR 1	GR 2
29	Demographic # 1 - Body Char	.50	104	137	142	8	$\chi^2 = 5.4808$, $p < .0001$	PDA	58	61	55
							$df = 36$	LR	58	58	58
								McNemar's z	.00	2.00	-2.00
30	Warncke Data - Groups 1 & 2	.48	138	65	47	10	$\chi^2 = 1.0593$, $p = .3611$	PDA	48	51	45
							$df = 10$	LR	45	68	13
								McNemar's z	.78	-3.32	3.87
31	Warncke Data - Groups 2 & 3	.45	118	47	40	10	$\chi^2 = 1.2556$, $p = .1039$	PDA	40	45	35
							$df = 55$	LR	43	60	23
								McNemar's z	-.58	-2.65	2.24
32	Enrollment Data	.39	157	272	426	4	$\chi^2 = 83.2884$, $p = .0000$	PDA	60	58	61
							$df = 10$	LR	62	28	84
								McNemar's z	-1.28	8.94	-9.85

TM 025570

AERA April 8-12, 1996

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC®

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Predictive Discriminant Analysis Versus Logistic Regression in Two-Group Classification Problems	
Author(s): Alice Meshbane & John D. Morris	
Corporate Source: Florida Atlantic University	Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

☑ **Sample sticker to be affixed to document**

Check here
Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Sample
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 1

Sample sticker to be affixed to document ☐

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY
Sample
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 2

or here
Permitting reproduction in other than paper copy.

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: Alice Meshbane	Position: Statistical Consultant
Printed Name: Alice Meshbane	Organization: Self-Employed
Address: P.O. Box 2847 Boca Raton, FL 33427	Telephone Number: (407) 392-4499
	Date: 4/8/96

CUA

THE CATHOLIC UNIVERSITY OF AMERICA
Department of Education, O'Boyle Hall
Washington, DC 20064
202 319-5120

February 27, 1996

Dear AERA Presenter,

Congratulations on being a presenter at AERA[1]. The ERIC Clearinghouse on Assessment and Evaluation invites you to contribute to the ERIC database by providing us with a written copy of your presentation.

Abstracts of papers accepted by ERIC appear in *Resources in Education (RIE)* and are announced to over 5,000 organizations. The inclusion of your work makes it readily available to other researchers, provides a permanent archive, and enhances the quality of *RIE*. Abstracts of your contribution will be accessible through the printed and electronic versions of *RIE*. The paper will be available through the microfiche collections that are housed at libraries around the world and through the ERIC Document Reproduction Service.

We are gathering all the papers from the AERA Conference. We will route your paper to the appropriate clearinghouse. You will be notified if your paper meets ERIC's criteria for inclusion in *RIE*: contribution to education, timeliness, relevance, methodology, effectiveness of presentation, and reproduction quality.

Please sign the Reproduction Release Form on the back of this letter and include it with **two** copies of your paper. The Release Form gives ERIC permission to make and distribute copies of your paper. It does not preclude you from publishing your work. You can drop off the copies of your paper and Reproduction Release Form at the **ERIC booth (23)** or mail to our attention at the address below. Please feel free to copy the form for future or additional submissions.

Mail to: AERA 1996/ERIC Acquisitions
The Catholic University of America
O'Boyle Hall, Room 210
Washington, DC 20064

This year ERIC/AE is making a **Searchable Conference Program** available on the AERA web page (http://tikkun.ed.asu.edu/aera/). Check it out!

Sincerely,

Lawrence M. Rudner, Ph.D.
Director, ERIC/AE

[1]If you are an AERA chair or discussant, please save this form for future use.