MNK Publication
International Journal of Latest Research in Science and Technology ISSN (Online):2278-5299
Volume 2,Issue 2 :Page No.40-45 , March - April 2013
https://www.mnkpublication.com/journal/ijlrst/index.php

APPLICATION OF GROUP-DEPENDENT TRANSFORMATION OF FISHER'S DISCRIMINANT ANALYSIS IN A UNIVERSITY OF TECHNOLOGY

[1]Audu Isah,[2]Usman Abdullahi
[1](Ph.D), Department of Mathematics/Statistics, Federal University of Technology, Minna-Nigeria.
[2]Academic Planning Unit, Federal University of Technology, Minna-Nigeria.

Abstract* - In this study, an expression for the discriminant rule in the situation of two groups was derived. This was used to identify the relative contributions of the subjects to the separation of the groups. Group dependent Fisher discriminant analysis was employed in classifying students into various departments on the basis of their cumulative results for one year foundation programme (Preliminary Degree Programme in a University of Technology). The discriminant scores for each department was predicted with probability of correct classification of 0.796 and apparent error rate of 0.204.

***Keyword* -** Linear Discriminant Coefficients, Classification Rules, Apparent Error Rate, Group Means, and multivariate.

1. INTRODUCTION

Discriminant analysis is an exploratory multivariate procedure of determining variables and a reduced set of functions called discriminants or discriminant functions. Discriminants that are linear functions of the variables are called linear discriminant functions (LDF). The number of functions required to maintain maximum separation for a subset of the original variables is called the rank or dimensionality of the separation. At the basis of observations with known group membership, the training data, called discriminant functions are constructed aiming at separating the groups as much as possible. These discriminant functions can then be used for classifying new observations to one of the populations. Discriminant analysis is used in situations where the groups are known a priori. For example, in personnel management one may want to discriminate among groups of professionals based upon a skills inventory. In medicine one may want to discriminate among persons who are at high risk or low risk for a specific disease. In a community, the mayor may want to evaluate how far apart several interest groups are on specific issues and to characterize the groups. In industry, one may want to determine when processes are in-control and out-of-control.

A multivariate technique closely associated with discriminant analysis is classification analysis. Classification analysis is concerned with the development of rules for allocating or assigning observations to one or more groups. While one may intuitively expect a good discriminant to also accurately predict group membership for an observation, this may not be the case. A classification rule usually requires more knowledge about the parametric structure of the groups. Because linear discriminant functions are often used to develop classification rules, the goals of the two processes tend to overlap and some authors use the term classification analysis instead of discriminant analysis. Because of the close association between the two procedures we treat them together in this study.

The Fisher's method for two populations leading to linear discriminant functions was adopted to design a classification rule for predicting which course/department a preliminary degree student is most likely to be admitted into at the end of the one year preparatory programme. The main objectives of this study are to find a set of rules, based on the student's results in classifying them into the five departments of the faculty, identify the relative contribution of the variables (subjects) to separation of the groups, and evaluate how well the rule performs in assigning a student to the correct programme.

2 LITERATURE REVIEW

Universities admissions processes often depend on the ability to predict student success. However, the use of a test to help determine admission has traditionally been problematic and continues to be so. Fisher (1938) introduced discriminant analysis as a statistical method for separating two groups of populations. Rao (1948) extended this multivariate technique to multiple populations.Charles and June (1970) carried out a study to determine if a differentiation or separation among students graduating, withdrawing or failing could be identified. Adebayo and Jolayemi (1998, 1999), applied the τ statistic to investigate how predictable the final year result would be using the first year result or Grade Point Average (GPA) of some selected University graduates.

It was cited in (Selingo & Brainard, 2001) that, the chancellor of the University of California called for the end of using testing for admissions to college. This was not a new call: a plethora of research has shown that standardized tests

Publication History
Manuscript Received : 11 April 2013
Manuscript Accepted : 15 April 2013
Revision Received : 17 April 2013
Manuscript Published : 30 April 2013

do not predict success equally well for all groups (Cleary, Humphreys, Kendirick, & Wesman, 1975; Melnick, 1975; Nettles, Thoeny, & Gosman, 1986; Tracey & Sedlacek, 1985) and that standardized tests do not measure what they claim to measure (Riehl, 1994; Sturm & Guinier, 2001). As an alternative to standardized tests, Sturm and Guinier (2001) suggested the use of multiple measures as a better way of deciding entry into law school.

Often, colleges may rely on two tests as a means of using multiple criteria, but if the two tests are highly correlated with each other, there is needless duplication in measuring the same aspect of a construct (Anastasi, 1982). Because the use of standardized tests has been shown to be problematic, multiple selection methods are being used to predict student success (Ebmeir & Schmulbach, 1989). The use of using multiple measures is called triangulation, the goal of which is to "strengthen the validity of the overall findings through congruence and/or complementarity of the results of each method" (Greene & McClintock, 1985, p. 524). This method is used extensively in education for admissions (Markert & Monke, 1990; McNabb, 1990) and involves using a variety of techniques simultaneously to measure a student's knowledge, skills, and values (Ewell, 1987).

Colleges can benefit from combining cognitive and non cognitive variables in predicting student academic success (Young & Sowa, 1992). Because the essence of triangulation is to measure the same construct in independent ways (Greene & McClintock, 1985), the more non-related information gathered, the better the prediction. Triangulation can also minimize or decrease the bias inherent in any particular method by counterbalancing another method and the biases inherent in the other methods (Mathison, 1988). For instance, most researchers rely heavily on survey research; however, the assumptions of survey research (e.g., the survey asked all the pertinent questions in a format the respondent can understand) are usually never questioned as a study is designed (Stage & Russell, 1992) which may lead to incomplete or inaccurate conclusions.

In the California Community Colleges, the required assessment process dictates the use of multiple measures in placing students into courses. Though the use of a test as one of the multiple measures is highly regulated, the use of multiple measures is not – unless using another test. Because of this, most multiple measures are chosen based on anecdotal or gut reactions and rarely on statistical evidence. It is the lack of research-based decisions for using multiple measures that inspired this study.

3. MATERIALS AND METHODS

3.1 Derivation of Expressions for First choice and Second choice and Discriminant Rules in the Setup of the Preliminary Degree Admission.

The preliminary degree final scores of students that applied to five departments of School of Engineering and Engineering Technology data set, X , contains five courses. Let us denote all the students that make first choice and second choice by X_f and X_s, respectively. The corresponding linear combinations are $y = \mathrm{X}a$, $y_f = \mathrm{X}_f a$ and $y_s = X_s a$.

The within-group sum of squares satisfies the relation

$$y_f^T H_f y_f + y_s^T H_s y_s = a^T W a$$

(3.1.1)

where H_f and H_s denote the appropriate centering matrices of dimensions $n_f = n_s = 100$. Observe that

$$a^T W a = a^T (X_f^T H_f X_f + X_s^T H_s X_s) a$$

(3.1.2)

And, hence, the matrix W can be written as:

$$W = X_f^T H_f X_f + X_s^T H_s X_s$$

(3.1.3)

$$= H_f X_f^T H_f X_f + H_s X_s^T H_s X_s$$

$$= n_f S_f + n_s S_s$$

$$= 100(S_f + S_s)$$

where S_f and S_s denote the empirical covariances w. r. t. the f and s .

For the between-group sum of squares we have

$$a^T B a = n_f (\bar{y}_f - \bar{y})^2 + n_s (\bar{y}_s - \bar{y})^2$$

(3.1.4)

where $\bar{y}$, $\bar{y}_f$ and $\bar{y}_s$ denote respectively the sample means y , y_f and y_s . It follows that

$$a^T B a = a^T \{ n_f (\bar{x}_f - \bar{x})(\bar{x}_f - \bar{x})^T + n_s (\bar{x}_s - \bar{x})(\bar{x}_s - \bar{x})^T \} a$$

,

where $\bar{x}$, $\bar{x}_f$ and $\bar{x}_s$ denote respectively the column vectors of sample means of X , X_f and X_s . Hence, we obtain

$$B = n_f (\bar{x}_f - \bar{x})(\bar{x}_f - \bar{x})^T + n_s (\bar{x}_s - \bar{x})(\bar{x}_s - \bar{x})^T$$

(3.1.5)

$$= 100\{ (\bar{x}_f - \bar{x})(\bar{x}_f - \bar{x})^T + (\bar{x}_s - \bar{x})(\bar{x}_s - \bar{x})^T \}$$

$$= 100\{ (\bar{x}_f - \frac{\bar{x}_f + \bar{x}_s}{2})(\bar{x}_f - \frac{\bar{x}_f + \bar{x}_s}{2})^T + (\bar{x}_s - \frac{\bar{x}_f + \bar{x}_s}{2})(\bar{x}_s - \frac{\bar{x}_f + \bar{x}_s}{2})^T \}$$

$$= \frac{100}{4} \{ (\bar{x}_f - \bar{x}_s)(\bar{x}_f - \bar{x}_s)^T + (\bar{x}_s - \bar{x}_f)(\bar{x}_s - \bar{x}_f)^T \}$$

$$= 25(\bar{x}_f - \bar{x}_s)(\bar{x}_f - \bar{x}_s)^T$$

The vector a maximizing the ratio $a^T Ba \big/ a^T Wa$ can be calculated as the eigenvector of $W^{-1}B$ corresponding to the largest eigenvalue. It is easy to see that the matrix $W^{-1}B$

can have at most one none zero eigenvalue since rank $B \le 1$. The nonzero eigenvalues λ_1 can be calculated as

$$\lambda_1 = \sum_{j=1}^{5} \lambda_j = trW^{-1}B = trW^{-1}25(\overline{x}_f - \overline{x}_s)(\overline{x}_f - \overline{x}_s)^T$$

(3.1.6)

$$= 25tr(\overline{x}_f - \overline{x}_s)^T W^{-1}(\overline{x}_f - \overline{x}_s)$$

$$= 25(\overline{x}_f - \overline{x}_s)^T W^{-1}(\overline{x}_f - \overline{x}_s)$$

From the equation (3.1.6)

$$W^{-1}BW^{-1}(\overline{x}_f - \overline{x}_s) = 25(\overline{x}_f - \overline{x}_s)^T W^{-1}(\overline{x}_f - \overline{x}_s)W^{-1}(\overline{x}_f - \overline{x}_s)$$

In the context of the pre-degree result data set, the "between-group-sum of squares" is defined as

$$100\left\{(\overline{y}_f - \overline{y})^2 + (\overline{y}_s - \overline{y})^2\right\} = a^T Ba$$

for some matrix B. Here, $\overline{y}_f$ and $\overline{y}_s$ denote the means for the first choice and second choice and $\overline{y} = 1/2\left(\overline{y}_f + \overline{y}_s\right)$. The "within-group-sum of squares" is

$$\sum_{i=1}^{100}\left\{(\overline{y}_f)_i - \overline{y}_f\right\}^2 + \sum_{i=1}^{100}\left\{(\overline{y}_s)_i - \overline{y}_s\right\}^2 = a^T Wa$$

with $(\overline{y}_f)_i = a^T x_i$ and $(\overline{y}_s)_i = a^T x_{i+100}$

for $i = 1,\ 2,\ \ldots,\ 100$

It follows that the eigenvector of $W^{-1}B$ corresponding to the largest eigenvalue is $a = W^{-1}(\overline{x}_f - \overline{x}_s)$.

The resulting discriminant rule consists of allocating student to the first choice programme if

$$R_f = \{x : (\overline{x}_f - \overline{x}_s)^T W^{-1}(x - \overline{x}) \ge 0\} \quad (3.1.7)$$

and of allocating student to the second choice programme when the opposite is true.

Considering ABE programme and analyse, we get

$$a = (-0.066,\ 0.000,\ 0.171,\ 0.304,\ 0.380)^T$$

Thus, substituting these values of the linear discriminant coefficients, in equation (3.1.7), we get:

$$R_f = -0.066(Eng) + 0.000(Maths) + 0.171(Phy) + 0.304(Chem) + 0.380(Agric)$$

$$R_f = -0.066(46.20) + 0.000(48.00) + 0.171(50.40) + 0.304(51.60) + 0.380(56.86)$$

$$R_f = 43$$

3.2 Test of significance with Multivariate Data

Several tests of significance are useful in conjunction with a discriminant function analysis. In this study, the T^2 test for equality of group means and Box's M test approximation to W using χ^2 were employed for each of the five courses (English Language, Mathematics, Physic, Chemistry, and Agricultural Science).

3.2.1 Equality of Group Means

To test the Hypothesis

$$H_0:\ \mu_f - \mu_s = 0 \text{ against } \mu_f - \mu_s \ne 0$$

Using F-transformation of Hotelling's T^2, as our test statistic.

$$F = (n_f + n_s - p - 1)T^2 / \{(n_f + n_s - 2)p\}$$

where $T^2 = (n_f n_s / n_f + n_s)D^2$ and D^2 is Mahalanobis distance with p and $(n_f + n_s - p - 1)$ d.f.

Reject H_0 if $F_{cal} > F_{p,\, n_f + n_s - p - 1,\, 1-\alpha}$

At the 0.05 level of significance, we rejected the hypothesis of equality of group means. This implies that there exist significant differences between the group means.

3.2.2 Equality of Covariance Matrices

To test the Hypothesis

$$H_0 : \Sigma_f - \Sigma_s = 0 \text{ against } \Sigma_f - \Sigma_s \ne 0$$

The test statistic is Box's M test approximation to W using a χ^2.

$$W = (n-k)Log\left|S\right| - \sum_{i=1}^{k} v_i Log\left|S_i\right|$$

where $S = \sum_{i=1}^{k} v_i S_i / (n-k)$

This approximation is reasonable provided $n_i > 20$ and both p and k are less than 6. Multiplying W by $\rho = 1 - c$ where

$$C = \frac{2p^2 + 3p - 1}{6(p+1)(k-1)}\left[\sum_{i=1}^{k}\frac{1}{v_i} - \frac{1}{n-k}\right]$$

the quantity

$$X^2 = (1-C)W = -2\rho LogM \xrightarrow{d} \chi^2(h)$$

where $h = p(p+1)(k-1)/2$.

Reject H_0 if $X^2 > X^2_{1-\alpha}(h)$

At the 0.05 level of significance, as expected, we accepted the hypothesis of equality of variance-covariance matrices. This analysis shows that the equality of variance assumption required when using Hotelling's T^2 statistics is tenable.

3.3 Classification Rules

To develop a classification rule for classifying an observation y into one or the other population in the two group case requires some new notation. First, we let $g_1(y)$ and $g_2(y)$ represent the probability density functions $(pdfs)$ associated with the random vector X for populations X_f and X_s, respectively. We let p_1 and p_2 be the prior probabilities that y is a member of p_1 and p_2, respectively, where $p_1 + p_2 = 1$. And, we let $c_1 = C(2|1)$ and $c_2 = C(1|2)$ represent the misclassification cost of assigning an observation from X_s to X_f, and from X_f

to X_s, respectively. Then, assuming the $pdfs$ $g_1(y)$ and $g_2(y)$ are known, the total probability of misclassification (TPM) is equal to p_1 times the probability of assigning an observation to X_s given that it is from X_f, $P = C(2|1)$, plus p_2 times the probability that an observation is classified into X_f given that it is from X_s, $P = C(1|2)$. Hence,

$$TPM = p_1 P(2|1) + p_2 P(1|2) \quad (3.3.1)$$

The optimal error rate (OER) is the error rate that minimizes the TPM. Taking costs into account, the average or expected cost of misclassification is defined as

$$ECM = p_1 P(2|1) C(2|1) + p_2 P(1|2) C(1|2) \quad (3.3.2)$$

A reasonable classification rule is to make the *ECM* as small as possible (Neil, 2002). In practice costs of misclassification are usually unknown.

Assuming that $\bar{y}_f > \bar{y}_s$, the classification rule becomes

Assign student to X_f if $T^2 \geq R_f$

Assign student to X_s if $T^2 < R_f$

3.4 Evaluating Classification Rules

Given a classification rule and $g_1(y)$ and $g_2(y)$ are known (along with their associated population parameters), the TPM expression given in (3.3.1) may be evaluated to obtain the actual error rate (AER). Because the specification of $g_1(y)$ and $g_2(y)$ is seldom known one generally cannot obtain the AER, but must be satisfied with an estimate. The simplest nonparametric method is to apply the classification rule to the sample and to generate a classification table. This is called the substitution or resubstitution method.

Then, the observed error rate or apparent error rate $(APER)$ is defined as the ratio of the total number of misclassified observations to the total

$$APER = \frac{n_{1Error} + n_{2Error}}{n_1 + n_2} \quad (3.4.1)$$

3.5 Application

3.5.1 Data collection

The data used for this study is the final score of preliminary degree students of Centre for Preliminary and Extramural Studies (CPES) that applied to school of Engineering and Engineering Technology, obtained from the average of the scores of the 1st and 2nd semesters 2011/ 2012 academic session in a university of technology. There are five courses offered by every preliminary degree student and these courses were used in constructing the discriminant rules. These include: English language, Mathematics, Physics, Chemistry, and Agricultural Science. The school of Engineering and Engineering Technology has five departments namely: Agricultural and Bioresources Engineering (ABE), Chemical Engineering (CHEME), Civil Engineering (CIE), Electrical/Electronic Engineering (EEE), and Mechanical Engineering (MECE)

3.5.2 Method of analysis

The two group Fisher's linear discriminant rule (Fisher, 1936) based on the maximization of the ratio of the between (B) to the within variance (W) of a projection $a^T x$ is used in this study. Every student makes two choices of departments/courses for admission. The first choice (department) is the first group and the second choice (defer from first department) is the second group.

Since the discriminant analysis technique, essentially is used to distinguish between two or more groups using characteristics on which the groups are expected to differ. These groups are expected to be statistically different from each other. This is achieved by forming a linear combination of the discriminating variables (independent variables) the coefficients are estimated so that they are in the best separation between the groups. Normally the first group gives the best discriminating coefficients. However, five different parameters arre used to adjudge which function gives the best discriminating coefficient, these are: wilks lambda, eigenvalue, canonical correlation, *p-value* and percentage variation.

4 RESULTS AND DISCUSSIONS

The data were analyzed using R software. The function used to carry out linear discriminant analysis is available in the MASS library and the results are shown in tables 4.1 to 4.3 below.

Table 4.1: Prior Probabilities of Groups

Programme	X_f (First Choice)	X_s (Second Choice)
ABE	0.546	0.454
CHME	0.523	0.477
CIE	0.500	0.500
EEE	0.613	0.387
MECE	0.540	0.460

Table 4.1 show that prior probabilities for all the departments are higher for the first choice courses as compared to the second choice courses except for civil engineering in which they are equal.

Table 4.2: Group Means

Programme	English	Mathematics	Physics	Chemistry	Agricultural science
ABE	46.20	48.00	50.40	51.60	56.86
CHEME	51.69	62.83	52.35	55.06	51.09
CIE	56.00	52.70	56.10	52.10	45.20
EEE	50.55	54.35	53.20	54.15	51.65
MECE	57.25	56.87	59.37	65.37	49.75

Table 4.2 show the predicted group means for the five subjects and for each department. The group means of chemical engineering and electrical and electronic engineering are higher for the five subjects as compared to the other courses.

Table 4.3: Estimated Coefficients of Linear Discriminants

Cours	ABE		CHEME		CIE		EEE		MECE	
	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
Eng	-0.066	-0.103	-0.299	-0.240	0.312	-0.007	-0.402	-0.737	0.000	0.000
Maths	0.000	0.046	0.836	0.522	0.536	0.029	1.006	0.408	0.384	1.173
Phy	0.171	0.130	0.590	0.153	0.618	0.165	1.422	0.629	0.343	1.153
Chem	0.304	0.074	0.754	0.371		0.214	0.103	0.063	0.490	0.809
Agric	0.380	0.227	-0.808	-0.692	-0.314	-0.033	-0.615	-0.594	-0.059	-1.961
……	………	………	………	………		………	………	………	………	………
Evalu	15.78	6.220	0.846	0.421	-0.058	7.514	3.767	1.510	28.41	9.403
% Va	98.85	60.06	99.94	71.59	………	90.86	99.18	50.57	92.28	35.96
CCor	0.708	0.552	0.667	0.544	11.62	0.736	0.996	0.888	0.983	0.934
Wilks	0.107	0.251	0.187	0.338	93.71	0.293	0.032	0.072	0.234	0.259
χ^2	69.13	42.83	29.55	17.60	0.959	2.226	32.79	13.27	11.06	4.030
P-Val	0.031	0.055	0.002	0.044	0.034	0.130	0.026	0.059	0.003	0.003
					5.109					
					0.053					

Table 4.3 is the estimated discriminant function coefficients for the departments, identifying the relative contribution of the subjects to separation of the groups. ABE department indicates that physics, chemistry, and agricultural science have greater contribution to the discriminant function. Mathematics has zero contribution and English which has negative contribution. This may be due to the weak academic background of the students in English and mathematics. CHEME and EEE department show that English and agriculture science contributed negatively to the prediction of prospective students for chemical and electrical/electronic engineering department. From the functions of CIE and MECE departments, physics has the highest contribution to the discriminating process of the function,of CIE and MECE departments, physics has the highest contribution to the discriminating process of the function, with mathematics and English in that order. Chemistry and agricultural science have negative contributions to these departments.

It could also be observed that the discriminating power is better for the first choice candidates for civil engineering and electrical engineering departments with mechanical engineering having the lowest. For the second choice, electrical engineering also has better discriminating power while chemical engineering has the lowest discriminating power.

Table 4.4: Assignment Rules

Programme	Assign candidate to X_f if $\lambda \geq R_f$	Assign candidate to X_s if $\lambda < R_f$
ABE	$R_f = 43$	$R_f = 43$
CHEME	$R_f = 68$	$R_f = 68$
CIE	$R_f = 61$	$R_f = 61$
EEE	$R_f = 80$	$R_f = 80$
MECE	$R_f = 71$	$R_f = 71$

APER = 0.204. Probability of Correct Classification = 80%. Table 4.4 gives the assignment rules for the programmes. It shows that ABE would admit first choice with at least 43 discriminant score, CHEME would admits first choice with at least 68 discriminant score, CIE requires first choice with at least 61 discriminant score, EEE requires at least 80 discriminant score for first choice, and MECE admits first choice with at least 68 discriminant score.

Evaluating the classification rules, there is misclassification of 39 students into second choice departments and 63 students into first choice departments. The apparent probability of error (APER) is: $(39+63)/500 = 20.4\%$. Hence, if a new student gets admission, he will be correctly classified with probability approximately 80%.

CONCLUSION

The discriminant analysis in which students' admission into programmes of study is characterized by means of discriminant rule was designed for each programme. Among the 500 Pre-Degree students, only 400 students are apparently classified correctly according to the first and second choices of year one degree programmes. The remaining 100 students are misclassified into year one degree programmes. This gives the proportion of correct classification rate to be 0.796 and apparent error rate of 0.204.

REFERENCE

[1] Adebayo S.B., Jolayemi E.T. (1998b). On the Effect of Rare Outcome on some Agreement/Concordance Indices. Nig. J. Pure and Appl. Sci. 13: 718-723.

[2] Adebayo S.B., Jolayemi E.T. (1999). Effect of Rare Outcome on the Measure of Agreement Index. J. Nig. Stat. Assoc. 13:1-10.

[3] Anastasi, A. (1982). Psychological Testing. New York: Macmillan.

[4] Charles B.K., June E.H. (1970). Predicting Graduation Withdrawal and Failure in College by Multiple Discriminant Analysis. J. Edu. Measur. 7(2): 91-95.

[5] Cleary, T.A., Humphreys, C., Kendrick, S.A., & Wesman, A. (1975). Educational Uses of American Psychologist, 30, 15-41.

[6] Ebmeir, H., & Schmulbach, S. (1989). An Examination of the Selection Practices Used in the Talent Search Program. Gifted Child Quarterly, 33, 134-141.

[7] Ewell, P.T. In Diane F. Halpern (Ed.), Student Outcomes Assessment: What Institutions Stand To Gain (pp.9-24). New Directions in Higher Education, No. 59. Vol. XV, number 3. San Francisco: Jossey-Bass.

[8] Fisher. R. A. (1938).The statistical utilization of multiple measurements. *Annals of Eugenics*, 8:376–386,.

[9] Fisher R.A. (1936). The use of Multiple Measurements in Taxonomic problems. Ann. Eug. 7: 179 – 188.

[10] Greene, J., & McClintock, C. (1985). Triangulation in Evaluation: Design and Analysis Issues. Evaluation Review, 9, 523-545.

[11] Markert, L.F., & Monke, R.H. (1990). Changes in Counselor Education Admissions Criteria. Counselor Education and Supervision, 30, 48-57.

[12] Mathison, S. (1998). Why Triangulate? Educational Researcher, 17, 13-17.

[13] McNabb, S.L. (1990). The Uses of "Inaccurate" Data: A Methodological Critique and Applications of Alaska Native Data. American Anthropologist, 92, 116-129.

[14] Melnick, M. (1975). Testing the Disadvantaged. American Psychologist, 30, 944-945.

[15] Nettles, M.T., Thoeny, A.R., & Gosman, E.J. (1986). Comparative and Predictive Analyses of Black and White Students' College Achievement and Experiences. Journal of Higher Education, 57 (3), 289-318.

[16] Neil H. T. (2002). Applied Multivariate Analysis. Springer, Pittsburgh

[17] Rao. C. R. (1948). The utilization of multiple measurements in problems of biological classification. *Journal of the Royal Statistical Society, Series B*, 10:159–203.

[18] Riehl, R.J. (1994). The Academic Preparation, Aspirations, and First-Year Performance of First-Generation Students. College and University, 70 (1), 14-19.

[19] Stage, F.K., & Russell, R.V. (1992). Using Method Triangulation in College Student Research. Journal of College Student Development, 33, 485-491.

[20] Sturm, S., & Guinier, L. (2001). The Future of Affirmative Action. Boston Review, December 2000/January 2001, 4-10

[21] Tracey, T.J., & Sedlacek, W.E. (1985). The relationship of Noncognitive Variables to Academic Success: A Longitudinal Comparison by Race. Journal of College Student Personnel, 26, 405-410.

[22] Young, B.D., & Sowa, C.J. (1992). Predictors of Academic Success for Black Student Athletes. Journal of College Student Development, 33, 318-324.

[23] Selingo, J., & Brainard, J. (2001, March 2). Call to Eliminate SAT Requirement May Reshape Debate on Affirmative Action. The Chronicle of Higher Education, pp. A21.