Iowa State University
Digital Repository

Retrospective Theses and Dissertations

Iowa State University Capstones, Theses and Dissertations

1982

# A comparison of: four multiple prediction selection techniques, mathematical and empirical estimation of weight validity, and the quality of prediction of three different combinations of a real life data set

Ahmad Suleiman Audeh
*Iowa State University*

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Educational Assessment, Evaluation, and Research Commons

Recommended Citation

Audeh, Ahmad Suleiman, "A comparison of: four multiple prediction selection techniques, mathematical and empirical estimation of weight validity, and the quality of prediction of three different combinations of a real life data set " (1982). *Retrospective Theses and Dissertations*. 7492.
https://lib.dr.iastate.edu/rtd/7492

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

المنارة للاستشارات

www.manaraa.com

## INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University Microfilms International

300 N. Zeeb Road
Ann Arbor, MI 48106

8224210

**Audeh, Ahmad Suleiman**

A COMPARISON OF: FOUR MULTIPLE PREDICTION SELECTION TECHNIQUES, MATHEMATICAL AND EMPIRICAL ESTIMATION OF WEIGHT VALIDITY, AND THE QUALITY OF PREDICTION OF THREE DIFFERENT COMBINATIONS OF A REAL LIFE DATA SET

*Iowa State University* PH.D. 1982

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark __√__.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages _____
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print __✓__
11. Page(s) __________ lacking when material received, and not available from school or author.
12. Page(s) __________ seem to be missing in numbering only as text follows.
13. Two pages numbered __________. Text follows.
14. Curling and wrinkled pages _____
15. Other ______________________________________________

University
Microfilms
International

# A comparison of: Four multiple prediction selection techniques, mathematical and empirical estimation of weight validity, and the quality of prediction of three different combinations of a real life data set

by

Ahmad Suleiman Audeh

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Professional Studies in Education
Major: Education (Research and Evaluation)

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1982

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

Page

# DEDICATION

To the reader

# CHAPTER I. INTRODUCTION

## Background

Doing everything possible to prevent the occurrence of student failure is far better than looking for the remedy to the failure after its occurrence, because the harm that may be done to students could be irreparable. Prediction studies can be considered as a step toward a more positive solution.

The problem in prediction studies is: What is(are) the best predictor(s) of student success out of an available pool of variables? Utilizing all the available variables to be in the prediction equation is not economical and it will consume inordinate amount of time and effort. Thus, the effective prediction equation is that equation which gives the best prediction with a minimum number of predictors.

There are several statistical methods for selecting variables in a prediction equation and there is no unique method for selecting the best prediction equation which fits all situations. Forward, Backward and Stepwise are three developed methods which are traditionally used (Blum and Naylor, 1968).

The selected prediction equation is method-dependent, that is, the different techniques may lead to different outputs for the same problem.

Draper and Smith (1966) explained the basic steps of selecting the best prediction equation by each of the three techniques utilizing the same data set. The basic steps of each selection procedure can be summarized as follows:

The basic steps of the backward procedure

1. A regression equation containing all the available variables is computed.
2. The partial F-test value is calculated for each variable treated as if it had been added to the regression equation after all other variables had been included.
3. The variable providing the lowest partial F-test value is removed if the F value is not significant; for this purpose a significance level ($\alpha$) must be preselected. Then Steps 1 and 2 are repeated using the remaining variables.
4. The procedure is terminated when the lowest partial F-test value is significant. The regression equation with that variable and all remaining variables is adopted.

The basic steps of the forward procedure

1. The variable having the highest zero-order correlation with the criterion is the first variable which enters the equation.
2. The partial correlation coefficients of the remaining variables with the criterion variable partialling out the variable(s) in the equation are calculated, the variable of highest partial correlation coefficient is entered if its partial F-test value is significant.
3. Step 2 is repeated. If the F-test value of the highest partial correlation coefficient of the remaining variables is not significant, the procedure is terminated and the equation selected prior to this step is adopted.

The basic steps of the stepwise procedure

1. The variable of highest zero-order correlation is entered in the equation first.
2. The second variable of the highest partial correlation coefficient is entered if it has significant F-value.
3. The contribution of the variable selected in Step 1 is recalculated as if it was entered after the variable selected in Step 2. If this partial F is not significant, the variable is deleted from the equation. These three steps are repeated for every variable, that is, F-to enter and F-to remove are calculated for each variable. The procedure is terminated when no variable not in the equation has a significant F-to enter.

Looking at the selection procedure using the proportion of the criterion variance explained by the independent variables in the equation, one can say that the selection of predictors in the forward selection procedure is computed by adding the variables one by one according to specific rules until further addition fails to increase the validity coefficients.

In backward procedure, variables are eliminated one by one from the equation, which initially contained all the variables, until a significant drop occurs in the validity coefficient. Stepwise procedure combines features of the forward selection and backward procedure at each step. In addition to entering one variable on the basis of its having the highest partial correlation with the criterion, with the effect of the included predictors removed, it also examines the possible elimination of each of the included predictors. If the elimination of any of the included predictors does not lead to a significant drop in the validity coefficient, that variable is eliminated.

Stepwise is, traditionally, the most frequently used method for selecting the prediction equation which accounts for about as much of the criterion's variance as do all the available variables. This method is probably the best according to Mallows, 1964 (as cited in Daniel and Wood, 1980) and according to Draper and Smith, 1966.

The problem arising with use of traditional methods is that the selected equation by one method may not be the same as the equation selected by another method, and the independent variables selected to be in the equation depend on the significance level of entering and/or removing a variable, which is itself a matter of subjective judgment. Another problem arising with the use of the traditional methods is that, in some of these methods, if a certain variable is removed at any step within that method (removed because, in the company of the most recently entered variable, it no longer contributes significantly to the validity coefficient), this may lead to a "snowballing" effect where it makes a difference which variable is entered and which is removed in the next and subsequent steps, hence order of entry becomes of prime importance.

Regression analysis is usually designed in such a way that it yields the highest possible correlation between the predicted scores based on the regression equation and the observed criterion scores. If the regression coefficients derived in one sample (screening sample) are applied to the independent variables of a holdout sample (calibration sample), the correlation between the predicted scores and the observed criterion scores in the holdout sample will be smaller than that in the original sample, because the zero-order correlations are usually treated as if they are error-free when obtaining the maximum validity coefficient (Lord and Novick, 1968). This means that the validity coefficient of the original sample is overestimated (biased upward). This biasness depends upon two ratios:

1. The number of selected variables (p) to the number of the available variables (P).
2. The number of subjects in the sample (N) to the number of the selected variables (p).

In general, the smaller the first ratio (p/P), the bigger the shrinkage in the validity coefficient, while the larger the second ratio (N/p) the smaller the shrinkage in the validity coefficient (Nunnally, 1978). It is recommended that the ratio N/p should be at least 30 subjects per variable (Kerlinger and Pedhazur, 1973).

Burket (1964) indicated that the typically used estimate of the shrunken validity coefficient is not an estimate of the accuracy of prediction to be expected on application to subsequent samples, but instead estimate of the validity coefficient that would have existed if the theoretical population regression weights had been used rather than the sample generated least squares estimates.

In prediction, the weight validity coefficient is of more practicality than the validity coefficient because interest is usually focused on how well the prediction equation holds up in subsequent samples or on the accuracy to be expected in future samples for which the criterion measure is unknown.

Tatsuoka (1973) suggested that the weight validity coefficient is the ultimate criterion for selecting the best set of predictors from the available pool of variables. His suggestion was recommended by the American Psychological Association (1974). Hence the prediction equation obtained by one of the traditional methods should be cross-validated to obtain an estimate of the degree of relationship to be expected in subsequent applications. Therefore, the subset of predictors that yields the prediction equation revealing the least shrinkage in the validity coefficient would then be favored.

Morris (1976) described a different regression procedure and provided an accompanying computer program which would select a prediction equation with the least shrinkage of the validity coefficient by optimizing the weight validity of double cross-validation and also yielding weight validity estimates and their significances to be expected in future applications. Later (1977) he modified this method and program version to allow the researcher to mandate the inclusion of certain variables in the selected prediction equation. This he described as an updated method of maximizing weight validity (MWV), claiming its superiority over the traditional methods of selecting the best prediction equation and he recommended this method in crucial situations.

Borg and Gall (1979) indicated that a new regression procedure may produce favorable results in one situation, but that it is dangerous to draw conclusions about the superiority of that procedure over other procedures because this superiority may become questionable after a more extensive assessment of its value in different situations extending over a period of time.

Thus, from a scientific point of view any method can be generalized if it is valid in different situations, and the MWV is one of the multiple regression methods that must be investigated.

Considering the prediction equation, validity coefficient and weight validity coefficient together, it can be said that the population validity coefficient is overestimated by the sample validity coefficient obtained by any of the traditional multiple regression selection methods. The shrinkage in the validity coefficient which one can expect in the long run, when the weights developed in one sample are applied to subsequent samples from the same population, could be obtained by estimating the population validity coefficient.

The population validity coefficient can be estimated by cross-validation as an empirical approach or by mathematical formulas constructed for this purpose as an alternative to the empirical approach.

Wherry (1931) developed a formula for predicting the shrinkage of the observed validity coefficient. This formula is

$$\bar{R}_w^2 = \frac{(N-1)R^2 - (p-1)}{N-p}$$

where

$\bar{R}_w$ = estimated validity coefficient

$R$ = observed validity coefficient

$p$ = number of predictors in the equation

$N$ = sample size.

In real life situations, the population prediction equation can never be known, and one is more interested in how effective the sample prediction equation is in subsequent samples. The sample cross-validity coefficient is a measure of this effectiveness. This coefficient will vary from one sample cross-validation to another, but the average sample cross-validity will be approximately equal to the validity coefficient when the sample regression equation is used in the population. Nicholson (1960) developed a mathematical formula, originally attributable to Lord (1950), for estimating the population cross-validity coefficient. The form of the Lord-Nicholson formula is

$$\bar{R}^2_{L-N} = 1 - \frac{N-1}{N-p-1} \times \frac{N+p-1}{N} (1-R^2)$$

Darlington (1968) presented another formula originally attributable to Stein (1960), for the same purpose, assuming random selection of predictor scores from a multivariate normal distribution. The form of the Stein-Darlington formula is

$$\bar{R}^2_{S-D} = 1 - \frac{N-1}{N-p-1} \times \frac{N-2}{N-p-2} \times \frac{N+1}{N} (1-R^2).$$

It is clear that there is no unique mathematical formula which can be used as an alternative to the empirical method for estimating the population validity coefficient, and there is an uncertainty that the same mathematical formula can be an alternative to different empirical methods.

Assuming that the best prediction equation was selected by one of the previous methods, another problem commonly encountered in predicting the success of students (e.g. college freshmen) is the validity of the equation over time. On the other hand, collecting data needed to revise the prediction equation may be expensive and time consuming. However, it is important to know the rate of decline in the predictive accuracy of a prediction equation selected by a specific method, and to test the sensitivity of different methods to the relative instability of the teaching-learning situation from time to time.

In some real life situations one cannot ignore the existence of males and females as two different populations. The variables selected by any procedure to be entered in the prediction equations may be different if the data sets of males and females are treated separately. The relationship between the variables may be obscured by treating the two sexes as a homogeneous group.

It is well known that a set of available variables may cluster together to form factors. Factor scores may be used to select the best factors as predictors instead of using the original data for selecting the best variables as predictors. Cattell (1966) has suggested that the use of factors as predictors in the prediction equation may be more accurate for generalizing to subsequent samples. This idea was enhanced by Gurtin and Bailey (1970).

The MWV method reveals that the traditional rule in multiple regression, which says that the criterion could be better predicted by including more variables in the prediction equation, may be violated. This may motivate the researchers to use factors as predictors in certain situations.

High school Grade Point Average (HGPA) is frequently used in many institutions, from an economic point of view, as the best single predictor; but from a scientific point of view it may not be the most economical in the long run if it is contaminated by some variables irrelevant to the criterion.

HGPA usually includes a person's grades in several curriculum areas, such as literature and science. Since it is likely that different aptitudes, skills and interests are required for success in each curriculum area, grades for each probably should be predicted separately in order to obtain the maximum validity coefficient or weight validity coefficient (Wayne and William, 1968).

The variables used as predictors must have common characteristics with the criterion. A subset of variables out of all the available variables may be selected by an exploratory analysis. The data from this subset of variables may be analyzed by any of the regression procedure. This methodology is typically used for prediction studies. Webb (1957) selected five variables out of sixteen available variables, by an exploratory analysis, from this subset three predictors eventually contributed significantly to the criterion.

## Definition of Terms

### Validity coefficient

Validity coefficient is the multiple correlation coefficient between the independent variables and the criterion in the original sample.

### Weight validity

Weight validity is the multiple correlation correlation coefficient between the independent variables and the criterion in subsequent samples.

### Partial HGPA

Partial HGPA is the high school grade point average of the variables that cluster together, out of an available pool of variables, to form one factor relevant to the criterion.

Contaminated HGPA

Contaminated HGPA is the high school grade point average of the available variables which are, in this study, all the academic subjects.

Full-word memory

Full-word memory refers to the integer array values in the range $(-2^{31} + 1)$ through $(2^{31} - 1)$.

Half-word memory

Half-word memory refers to the integer array values in the range $(-2^{15} + 1)$ through $(2^{15} - 1)$.

## Purpose of the Study

The purpose of this study was to compare the performance of the prediction equations selected by four different methods using a data set from a typical life situation. The best prediction variables selected out of an available pool of variables utilizing the "superior" method may be of practical use in that real life situation.

The comparison between different regression methods using large sample size is important in promoting a better understanding of the method used and the conditions under which minimum shrinkage in the validity coefficient is achieved.

The MWV plays a major role in this study, because it was developed to provide a minimum shrinkage in the validity coefficient for subsequent samples. This study was designed to ascertain if any of the mathematical shrinkage formulas can be an alternative to MWV as an empirical approach.

The variables selected by the "superior" method as best predictors may cluster together to form one factor. If so it will be possible to factor analyze the data to use partial HGPA instead of total HGPA as the best single predictor. Therefore, one of the purposes of this study was to compare the validity coefficients obtained by the "superior" method using factor scores with that obtained by the same method using variable scores.

The current criterion of selecting the potentially successful college student in the selected situation utilized in this study is a contaminated HGPA and it is currently assumed that this criterion is the best single criterion. Hence, another purpose of this study was to test the efficiency of this predictor, and such other alternatives by comparing its performance with the performance of the predictors as factors or as separate variables.

## Problem of the Study

The following information may provide an insight to better understand the problem.

1. It is well-known that the larger the ratio N/p the smaller the shrinkage in the validity coefficient. A recommended minimum ratio in multiple regression is 30 subjects/variable (Kerlinger and Pedhazur, 1973). Morris utilizing the MWV method claimed that this method is superior to the traditional methods. It remains to be seen if this method is superior under all conditions.

2. The mathematical formula used to estimate the shrinkage of the validity coefficient, as an alternative to an empirical method, is a function of the utilized selection method, the purpose of the estimation and on the extent to which the assumptions of the multiple regression selection method are violated. The amount of shrinkage in the validity coefficient is a function of the method used. The MWV is a method which provides a direct estimation of the weight validity in subsequent samples. The weight validity obtained by this regression technique is maximized and it provides the minimum shrinkage when compared with the shrinkage obtained by any of the traditional methods calculated by the mathematical shrinkage formulas. So, whether any of the known formulas can be used as an alternative to the empirical estimation approach is still open to question.

3. The variables selected as the best depend upon the method used. The validity coefficient and the variables selected by a method may be different if the data of subgroups, for instance, males/females and the total group were treated separately.

4. The validity coefficient may change from time to time in the same population. The stability of the prediction equation may be tested by calculating the validity coefficient from the same population over a period of time using the same regression procedure.

5. The total set of the available variables may cluster together to form one or more factors. Using factor variables as predictors may provide a better quality of prediction as contrasted with using the composite scores of those variables as single predictors and the performance of the selected variables employing any of the regression techniques.

This study was conducted to answer the following questions:

1. Using the subset of variables which have common characteristics with the criterion, and a life data set:

   A. With a fixed ratio, is(are) the same variable(s) selected as the best predictors by the four regression techniques?

   B. With a fixed ratio, are the four techniques different in their quality of predicting the criterion, measured by the proportion of the explained variance of the criterion by all the selected predictors as well as the criterion of the number of times the weight validity coefficient of separated samples is largest?

   C. With a fixed method, are the selected variables, the proportion of the explained variance and the number of times the weight validity coefficient is highest dependent on the ratio?

2. Can any one of the three shrinkage formulas, as a mathematical approach, be an alternative of the MWV, as an empirical approach for estimating the weight validity coefficient in subsequent samples?

3. A. With a fixed method and ratio, is(are) the selected variable(s) dependent upon sex?

   B. With a fixed method and ratio, is the proportion of the explained variance dependent upon the sex of the respondent?

4. With a fixed method, ratio and sex, is the prediction stable over a two year period?

5. Comparing the quality of the prediction of the criterion, which is the Freshman College Grade Point Average (FGPA), from the original variables combined or analyzed in the following manner:

First: the best variables as predictors

Second: the appropriate number of factors as predictors

Third: the single variable HGPA, which is the current criterion of selection used by the admission and registration department of the educational institution from which the data were collected.

Which one of the three analysis described above provides the best quality of prediction which can be recommended for practical application?

## Limitations of the Study

The data used in this study belong to one of a large number of true life situations, that may be selected by the researcher to compare the performance of the four mentioned selection methods in addition to the possibility of using theoretical data for the same purpose. In this study, the situation is the prediction of freshman success in the college of science at Yarmouk University, Irbid - Jordan, reflected by the FGPA utilizing high school achievement reflected by student grades within different subjects taken within the third secondary class (scientific track). The grades for each student on each subject were determined by the General Secondary Education Certificate Examination (GSECE). Any student who did not have a grade for each of the eight required subjects or who did not take the GSECE of Jordan was not selected for the study. Therefore, the results of this study will not be definitive for all populations or all times in deciding the superiority of any of the four multiple regression selection methods. However, the results of this study may contribute partial theoretical additions to what is known about the performance of the four methods, and thus provide practical additions to what is used in the selected life situations.

The academic subjects on the official transcript were the only available variables that could be used in this study.

The number of males, in the population of the study, was greater than the number of females. This fact increased the difficulty of treating the two sexes together as a single group, which would have led to obscuring the relationships between the variables. This limitation is avoided when the two sexes were treated separately.

The current GSECE system was a relatively new system in comparison with one utilized three years earlier. This fact made the testing of replication beyond two years not possible.

## Assumptions of the Study

1. The sample for this study consisted of all the freshmen students, who met the criterion of selection, enrolled in the College of Science at Yarmouk University for the academic year of 1978-1979 and 1979-1980. Therefore, an imaginary infinite population was assumed. The concept of assuming an imaginary infinite population when the sample consists of all known subjects in a

particular situation was introduced by Fisher (cited in Marriott, 1975).

2. It was assumed that the high school grades of these sampled accurately reflect the student high school achievement, and the FGPA accurately reflects the freshman college achievement.
3. It was assumed that the independent variables possess a multivariate normal distribution, but even if this assumption is violated into some extent, that the multiple regression technique is extremely robust with respect to this assumption (Daws and Corrigan, 1974).
4. It was assumed that the criterion (FGPA) was normally distributed in the original population.
5. It was assumed that the zero-order intercorrelations of the independent variables are not collinear. The presence of extreme collinearity leads to sizable standard errors of the estimated regression coefficients (Asher, 1976).

# CHAPTER II. REVIEW OF LITERATURE

The preceding chapter stated that different multiple regression selection methods may provide different results, and the superiority of each method may be situational specific. It is dangerous to draw conclusions about the superiority of a new procedure even if it produces favorable results in specific situations. The contribution of prediction research to educational practice will be more effective if superior method(s) of higher external validity is(are) used. The stability of a prediction equation over time is desired. Prediction may be more effective if the effect of moderator variable(s) (if any) is(are) manipulated, and if the variables or their combination into factors have common characteristics with the criterion. This chapter cites literature pertinent to the problems of this study. Five categories appear relevant: (1) research on the comparison of the predictor-variable selection techniques, (2) research concerning mathematical and empirical estimation of the shrinkage of validity coefficient, (3) research concerning the validity of a selected prediction equation over time, (4) research concerning sex as a moderator variable, and (5) research concerning factors as predictors.

## Research on the Comparison of the Predictor-variable Selection Techniques

The MWV is a newly developed regression technique. The traditional techniques (forward, backward and stepwise) have been frequently used (mainly the stepwise). Many studies are required to provide enough information about the efficiency of the MWV technique in comparison with the traditional techniques; thus, the current literature on this topic is very limited.

Morris et al. (1980) compared the performance of the MWV technique with that of the traditionally used techniques using the standard scores of 83 tenth grade students (black and white, male and female) on three subtests, language, work-study skills and mathematics as well as the composite score of the Iowa Test of Basic Skills, which were administered to these students in the eighth grade and utilized as predictors of their grade point average following the tenth grade.

The variables selected by the MWV to be in the prediction equation were the subtest scores of language and mathematics, while only the composite scores of the Iowa Test of Basic Skills were selected by the three traditional techniques as comprising the best predictor(s) when the same significance level used in all techniques for inclusion or/and exclusion of a variable. The traditional techniques selected the variable of maximum zero-order correlation with the criterion while the MWV selected the two variables of lowest zero-order correlations with the criterion. There were no common variables when comparing the variables chosen by the traditional techniques and by the MWV technique.

This study presented a new idea which differed from what was traditionally known about prediction equation selection techniques. It was known that the typical objective of these techniques was that of finding the subset of variables that would account for about as much of the criterion's variance as do all the variables, and the validity coefficient usually increased as one added predictors to the equation. According to the results of this comparison study, the equation with four predictors performed relatively poorly with respect to many of the alternatives contained within the 15 possible equations. This meant that the variable subset of which one would expect the greatest accuracy of prediction on application to subsequent or similar samples is determined by the MWV technique, irrespective of the number of possible predictors.

The data set used in that comparison study had some certain characteristic which must be mentioned.

1. The number of subjects per variable was considered an intermediate ratio (21 subjects per variable)
2. The data were assumed to be from a homogenous group, while in fact it may have had two moderator variables (sex and race).
3. The zero-order correlations between the four variables and the criterion were rounded to three decimal point accuracy. The maximum difference between any two of these correlations was .071, and the minimum difference was .001.
4. The correlation between language and mathematics was the smallest of the zero-order intercorrelations.
5. The zero-order correlations between the predictors revealed extreme multicollinearity (two-thirds of the intercorrelations were $>.80$).

The results of this comparison study indicated that the prediction equation selected by the traditional techniques were not valid. Morris et al. recommended that the validity of a regression equation selected by any of the traditional techniques could be tested by the MWV technique.

## Research Concerning Mathematical and Empirical Estimation of Shrinkage in the Validity Coefficient

The overestimation of the population multiple correlation by the validity coefficient, due to sampling error, has created considerable concern. Mosier (1951) extended the methodology of cross-validation to double cross-validation as the best approximation of the population multiple correlation coefficient. Researchers have typically employed this empirical estimation procedure since 1951.

Schmitt et al. (1976) indicated that the employment of this methodology in many practical situations posed a dilemma in their own words was as follows:

> The practical researcher with a small sample is faced with the dilemma of deciding between two incompatable goals. If he/she uses all of the data to make the best determination of regression weights, he/she cannot arrive at an estimate of the effectiveness of the prediction equation in subsequent samples. However, if the data are split into estimation and holdout samples in order to obtain such an estimate, one must necessarily settle for less than the most stable weights that could be obtained from the total data set.

The solution to this dilemma, as they indicated, was by employing the appropriate shrinkage formula. In their study they illustrated empirically, using theoretical data, that using the appropriate formula provided an accurate estimate of the population multiple correlation and the amount of shrinkage that could be expected when the weights developed in one sample were applied to subsequent samples from the same population. The following values were calculated to illustrate the relative efficiency of the three aforementioned shrinkage formulas found in Chapter I (Wherry, Lord-Nicholson and Stein-Darlington) and the weight validity as estimates of the population multiple correlation coefficients:

$R$: Validity coefficient when all the variables are in the equation

$R_S$: Stepwise validity coefficient

$R_C$: Cross-validation multiple correlation coefficient

$\bar{R}_W$: Wherry multiple correlation coefficient

$\bar{R}_{L-N}$: Lord-Nicholson multiple correlation coefficient

$\bar{R}_{S-D}$: Stein-Darlington multiple correlation coefficient

$R_P$: Population multiple correlation coefficient.

The calculations were repeated with different sample size with a differing number of predictors included in the prediction equation. When all the variables are in the equations the results indicated that:

1. $R^2 - R_C^2$ (shrinkage occuring upon cross-validation) was the largest shrinkage for all levels of sample size, and the Wherry formula $(R^2 - \bar{R}_W^2)$ consistently underestimated this shrinkage. The Lord-Nicholson $(R^2 - \bar{R}_{L-N}^2)$ and Stein-Darlington $(R^2 - \bar{R}_{S-D}^2)$ formulas overestimated it. These overestimations were negligible. However, the formulas yielded conservative results.
2. The Stein-Darlington estimates were more conservative than the Lord-Nicholson estimations, but the discrepancies were minor and probably due to the extent to which the assumption of multivariate normality was violated.

When the stepwise procedures were used the results indicated that:

1. Wherry formula underestimated the shrinkage occurs upon cross validation $(R_S^2 - R_{CS}^2)$ specially for small sample size. Stein-Darlington and Lord-Nicholson formulas underestimated it; this

underestimation may be of considerable amount in case of Lord-Nicholson and small sample size.

Strictly speaking, the Lord-Nicholson formula provided accurate estimates of how one can expect a sample regression equation to perform in the long run ($R_p^2$) when all predictors were included in the regression equation, while for most practical purposes, the Stein-Darlington formula provided an accurate estimate of cross-validation expectations when traditional techniques (traditional methods) were used.

### Research Concerning the Validity of a Selected Prediction Equation Over Time

A problem commonly encountered in predicting college freshman grades from high school grades was establishing the validity of the prediction equation over time. Changes over time in instructors' grading policies, in the distribution of ability among entering student and other factors may contribute to make old prediction equation inaccurate.

Hills et al. (1965) compared the predicted and the observed grades of students at seven colleges in Georgia over a two academic year period; a very small decline in the validity coefficients (.64, .63) was observed.

Bowers and Loeb (1972) found that the prediction equation for freshmen at the University of Illinois was unstable over a five-year period. The regression weights of the used predictors, American College Test battery were unstable. A significant decline in the validity coefficient was observed.

Perrin and Whitney (1976) studied the records of students who participated in the American College Testing (ACT)'s predictive research services. They found very few differences in the accuracy of the expectancy table over three years as an alternative index of evaluating the validity of a developed prediction equation.

Sawyer and Maxey (1979) used the records of students who participated in ACT program (males and females) and found that the accuracy (using the absolute mean error as an index of accuracy) of the prediction equations of males or females and the total group was stable over the four-year period studied, but the predictions for females were somewhat more accurate than those for males.

## Research Concerning Sex as a Moderator Variable

Earlier studies of predicting the success of college freshmen, using different predicting variables (Graduate Record Examination, Scholastic Aptitude test, High school Grade Point Average and other standardized or local tests) indicated that the validity coefficients were substantially higher for females than the corresponding coefficients for males (Durflinger, 1943; Jackson, 1955; Scannell, 1960). In these studies, the numbers of males and females were different (e.g., in Scannell study the ratio was 2 male/1 female).

McDonald and Gawkoski (1979) used approximately equal numbers of males and females from the sample. The results indicated that there was a significant difference between the validity coefficients in favor of females.

Sawyer and Maxey (1979) indicated that the prediction equation of females showed better stability over time than that of males, and a considerable decline in the weight validity coefficient of males was observed.

## Research Concerning Factors as Predictors

It was mentioned in Chapter one that a single predictor performance, used by any educational institution may be contaminated. Billeh et al. (1974) referred to this kind of contamination. Using factors, extracted from the original variables, as predictors may be an appropriate methodology of reducing the effect of this contamination. At the same time, better predictability and more stability of the validity coefficients hopefully may be attained.

Morris and Guertin (1977) used the theoretical data made up of six hypothetical variables. Two varimax-rotated factors were extracted from the variables. The original variables and the extracted factors, at different communality, were used separately to predict a hypothetical criterion. The validity coefficients were calculated and the shrinkage of these coefficients were estimated. Their results indicated that the validity coefficients for factors and the original variables were equal, while the shrinkage in the validity coefficient for factors was smaller than the shrinkage in the validity coefficient for the original variables.

## Summary

A survey of the literature in the general area of prediction of academic achievement revealed the availability of a wealth of studies. However, those which investigated the quality of the prediction utilized were few. The important issues related to this study which were investigated within the related literatures were: (1) comparing the quality of prediction of MWV and traditional regression techniques utilizing a data set, with specified characteristics, resulting from a practical educational situation. The MWV had better quality of prediction in that specific situation. It produced a new index of testing the quality of prediction of a set of predictors and yielded a new concept about the inclusion of variables in the reduced prediction equation. (2) Comparing the shrinkage in the validity coefficient utilizing mathematical shrinkage formulas and the technique of cross-validation. The shrinkage formulas differentially estimated the shrinkage of the validity coefficient of a theoretical data set for different sample sizes. (3) Researching the validity of a calculated prediction equation over time. The related research revealed that the stability of a prediction equation was situational specific, that is, it depends on the educational system of that situation. (4) Investigating the quality of prediction utilizing sex as a moderator variable. A few of the studies ignored this factor, but most indicated that females were more predictable than males. (5) Comparing the quality of prediction of variables with prediction using factors extracted from these variables. The findings of the few available studies indicated that the different indices of quality of prediction favored the factor predictors. In general, the empirical evidence of comparing the predictability of variables and factors seems to be absent from the literature, because factor scores are not usually used at present by most researchers, but factors may be recommended because a trait or a construct is more stable than its constituents.

# CHAPTER III. METHOD OF PROCEDURE

The purpose of this study was to compare the performance of the prediction equations selected by the four regression techniques, and to find the predictors (factors or variables) which provide maximum predictivity as well as stability for males and females treated separately and for both treated as a single group. The performance of the selected variables was then compared with that of the current single predictor currently used in the real life situation of this study. The contribution of this study could be used in related research and in direct applications, because it had theoretical and practical values.

## Sample and Population

The records of 344 students (220 males and 124 females) who were freshmen in the academic year 1979-1980 and the records of 275 students (198 male and 77 female) who were freshmen in the academic year 1978-1979 in the college of science at Yarmouk University, Irbid, Jordan, were studied. Each student was officially admitted as a full-time student. High school Grade Point Average (HGPA) of the student who passed the General Secondary Education Certificate Examination (GSECE) in the scientific track was the criterion of selection. Students who did not have GSECE or a complete transcript or who were admitted conditionally were deleted from the study. Records of students who withdrew or transferred to another college, such as Art or Administration, or terminated for academic reasons, were retained in the sample.

## Variables and Data Collection

The eight academic subjects contained in the transcript of grades of the student within the scientific track were used as independent variables. These subjects are weighted differentially in the admission process. The eight subjects with their abstracts and their maximum points accompanied with the minimum passing grades acceptable are: Islamic Education (IE, 100, 40); Arabic Language (AL, 200, 100); English Language (EL, 200, 80); Mathematics (MA, 300, 120); Physics (PH, 100, 40); Chemistry (CH, 100, 40); Biology (BI, 100, 40); and Arabic Society and Palestine Problem (AP, 100, 40). The HGPA of each student was calculated by adding the points obtained by the student in the first four subjects listed to the highest two points of the four remaining subjects and then divided by ten. This gives the student a HGPA in the range of 42 to 100.

The Freshman Grade Point Average (FGPA) was the dependent variable (criterion), assuming that it reflected the student's academic scientific achievement in the College of Science. All the required data were available in the records of the Admission and Registration Department of Yarmouk University.

## Data Analysis

As was previously stated, the variables which may be used as predictors must be relevant to the criterion and a subset of the available variables may be selected by an exploratory analysis which also investigated the homogeneity of the HGPA variance for males and females, the assumptions of collinearity and normality of the dependent variable.

The exploratory analysis was conducted in the following manner:

A. The intercorrelations of all the variables for males, females and the total group of the academic year 1979-1980 were calculated utilizing the SPSS Pearson correlation subprogram.

B. The unique contribution of each variable to the criterion was tested utilizing the General Linear Model (GLM).

C. The criterion data were tested for normality, because an extreme violation of this assumption would affect the validity of the F-test, which was used for selecting variables by any regression technique.

The variables selected by the exploratory analysis were used to compare the performance of the prediction equations selected by the four multiple regression methods. To achieve this, the following analyses were done:

1. The data of the total group were analyzed by the four selection techniques (forward, backward, stepwise and MWV). For each technique, the best regression equation was calculated for each of the following sample sizes.

   A. A small sample size of 10 subjects per variable. However, up to 12 subjects per variable was considered small sample size (Drehmer and Morris, 1981).

   B. An intermediate sample size of 20 subjects per variable.

   C. A large sample size of 40 subjects per variable.

2. The data of males and females were treated separately by the same procedure outlined in Step 1. The purpose of this analysis was to compare the performance of the prediction equations of males to that of females.

The shrinkage in the validity coefficient was calculated by estimating the validity coefficient of subsequent samples using the MWV, as an empirical approach, and by the three mathematical shrinkage formulas.

For selecting the prediction equation using factors as predictors, the procedure was as follows:

1. The total data set was factor analyzed using the following steps:
   A. First, the principal component method (SPSS-PA2) which automatically replaced the main diagonal elements of the correlation matrix with communality estimates.
   B. Next, Kaiser's criterion was utilized, which stated that if the associated eigenvalue of a factor was less than 1.0, it was difficult to assign any meaning or any positive generalizability to that factor.
   C. Finally, a varimix rotation was performed. This was an orthogonal method of rotating factors which maximized the variance of the squared loadings in each column of the factor matrix.
2. The factor scores of all the students were calculated.
3. The validity coefficient was computed by the stepwise procedure.
4. Steps 1 to 3 were repeated for the subset of males and the subset of females.

The prediction equation from the second set of data (78/79) was computed utilizing the stepwise procedure. The performance of this equation and the variables selected were compared with that of the first set of data (79/80).

The composite scores of the current single predictor (HGPA) were calculated. Then, the Pearson correlation coefficient with FGPA was calculated. The purpose of this step was to compare the performance of the currently used predictor with the performance the other derived alternatives (variables or factors).

## Computer Programs

The regression, correlation and factor analysis subprograms of the Statistical packages, Statistical Analysis System (SAS) and Statistical Package for Social Science (SPSS), were used for part of the analysis. The MWV has its own program prepared by Morris (1977). The MWV program was written in the Fortran language. The MWV was explained in the words of Morris as follows:

> The input sample with p predictors is randomly split in half and least square-estimated regression weights are calculated for all $2^p-1$ possible regression equations for each sample. The weight validities for each equation over both subsamples are then calculated. The equation manifesting the highest weight validity is noted. This procedure is repeated either an input number of times or until a regression equation shows itself superior. The superiority of the regression equation is judged in terms of the binomial probability that the number of times the superior equation showed the highest weight validity is equivalent to the number of times the next best . . . equation was highest. If the probability thus generated is smaller than the user input probability, . . . the superior equation is chosen as best, and iteration stops.

Unlike typical variable selection procedures, normality assumptions are unnecessary. The possibility that a cluster of equally superior equations arise is also accommodated. Such a cluster would be defined by no significant differences, again through the binomial expansion, between the equations within the cluster and each equation being significantly superior to those equations not in the cluster. This incorporates the fact that there may be more than one "best" equation in terms of weight validity. The selection decision from such a set of equations would normally be made on the economic and/or subjective merits of the measurement process involved with each of the variables. Lacking subjective reasons for preferring certain variables over others, the equation within the superior set with the smallest number of variables would probably be selected.

## Cost Analysis

It is difficult to confirm whether the cost of a computational procedure is expensive or not when this cost is compared with the direct and/or indirect (explicit and/or implicit) benefits of the results over the long run, especially in the behavioral sciences. The cost of the computation procedure of MWV was difficult to estimate before this study was undertaken. Previously the program of MWV was run only in analyzing data of a study belonging to Morris. The cost of analysis of that study was not estimated because he received a free-time university computer account. Looking at the data analysis of the current study, it was found that the average cost was relatively high compared with that of any of the three mentioned multiple regression techniques. The different computer runs utilizing the MWV method revealed that the cost of this program depended on the following factors:

1. The starting value (see number) for use in a random number generator which produces a sequence according to a specific algorithm. There was no direct way of estimating the cost due to this factor. A Monte-Carlo study is suggested to answer this question.
2. The probability level specified for stopping the iteration. In general, the more conservative the level of significance the higher the cost. Again, there is no graphical relationship or a mathematical formula which could be used to estimate the cost of the computation associated with a given probability level, assuming other factors constant.
3. Number of subjects per variable. It can be said that the relationship between the cost and any of these two factors is unspecified. An empirical approach may provide results of permanent use.

The original program for MWV was designed to handle up to 10 independent variables. The reserved memory utilized for these variables maximized the cost of computations, but this cost could be minimized simply by modifying the original program. The computation time was minimized by modifying the READ and WRITE instructions. After modification, the cost range for the different runs of the current study was between \$.61 and \$13.86.

The original program was designed for a Control Data Cooperation (CDC) Fortran compiler. Therefore, it was modified to fit an International Business Machine (IBM) OS/360 compatible machine which is available in the computation center of Iowa State University. The primary cost of modification was \$485.00, but any further modification by other researchers would be inconsequential. A complete copy of the modified program of the MWV regression technique is found in the Appendix.

The original program was modified in the following manner:

1. The memory requirement was reduced by changing the integer array used to specify all possible models from Full-word to Half-word.
2. The memory for all arrays was reduced by changing the maximum number of independent variables from the original number (10) to the number used in this study (5).
3. The READ and WRITE instructions, which were used to store and retrieve the data, were changed from FORMATED INPUT-OUTPUT into UNFORMATED INPUT-OUTPUT. This was accompanied with a corresponding change in the Job Control Language.
4. The subroutine for generating the (0,1) distribution from Cooley and Lohenes Library did not work on IBM OS/360, therefore, the subroutine was changed into IMSL Random Number Generator (GGUBFS).
5. Double precision was used for all floating point computations.

# CHAPTER IV. RESULTS AND DISCUSSION

## Results of Exploratory Analysis

Intercorrelations of the eight predictor variables, the criterion and the moderator variable (sex) are presented for the total sample in Table 1. From the entries in this table, the validity coefficients (which are the zero-order correlations) of each of the predictor variables in descending order of magnitude were as follows: CH, .557 ($p < .01$); PH, .490 ($p < .01$); BI, .390 ($p < .01$); MA, .382 ($p < .01$); EL, .370 ($p < .01$); AL, .352 ($p < .01$); IE, .166 ($p < .01$) and AP, .164 ($p < .01$).

Intercorrelations of the eight predictor variables and the criterion are presented for the subsample of males in Table 2. From the entries of this table, the validity coefficients of each of the predictor variables in descending order of magnitude were as follows: CH, .431 ($p < .01$); PH, .369 ($p < .01$); EL, .279 ($p < .01$); BI, .265 ($p < .01$); MA, .239 ($p < .01$); AL, .190 ($p < .01$); AP, .118 ($p < .05$) and IE, .025 ($p < .05$).

Intercorrelations of the eight predictor variables and the criterion are presented for the subsample of females in Table 3. From the entries of this table, the validity coefficients of each of the predictor variables in descending order of magnitude were as follows: CH, .710 ($p < .01$); PH, .662 ($p < .01$); MA, .641 ($p < .01$); BI, .611 ($p < .01$); AL, .554 ($p < .01$); EL, .412 ($p < .01$); IE, .351 ($p < .01$) and AP, .291 ($p < .01$).

Table 1. Intercorrelations[a] of eight predictors, a criterion and a moderator variable for the total group (N = 344)

| Variables | FGPA | IE | AL | EL | MA | PH | CH | BI | AP |
|---|---|---|---|---|---|---|---|---|---|
| SEX | .244 | .214 | .147 | .267 | (-.021) | .113* | .176 | .110* | (-.032) |
| FGPA | | .166 | .352 | .370 | .382 | .490 | .557 | .390 | .164 |
| IE | | | .396 | .132 | .235 | .205 | .164 | .219 | .273 |
| AL | | | | .306 | .370 | .406 | .410 | .419 | .430 |
| EL | | | | | .188 | .238 | .286 | .330 | .201 |
| PH | | | | | | .540 | .447 | .423 | .187 |
| CH | | | | | | | .474 | .326 | .101* |
| BI | | | | | | | | .500 | .126 |
| | | | | | | | | | .309 |

FGPA: Freshman Grade Point Average
AL : Arabic Language
MA : Mathematics
CH : Chemistry
AP : Arabic society and Palestine problem

IE : Islamic Education
EL : English Language
PH : Physics
BI : Biology

[a]( ) p > .05; *p < .05; others p < .01.

Table 2. Intercorrelations[a] of the eight predictor variables and the criterion of the subgroup of males (N = 220)

| Variables | IE | AL | EL | MA | PH | CH | BI | AP |
|---|---|---|---|---|---|---|---|---|
| FGPA | (.025) | .190 | .279 | .239 | .369 | .431 | .265 | .118* |
| IE | | .402 | (.067) | .218 | .135* | .101* | .188 | .305 |
| AL | | | .293 | .326 | .316 | .285 | .385 | .459 |
| EL | | | | .188 | .159 | .248 | .290 | .290 |
| MA | | | | | .494 | .375 | .390 | .160 |
| PH | | | | | | .376 | .235 | (.070) |
| CH | | | | | | | .479 | (.078) |
| BI | | | | | | | | .346 |

[a]( ) $p > .05$; *$p < .05$; others $p < .01$.

Table 3. Intercorrelations[a] of the eight predictor variables and the criterion for the subgroup of females (N = 124)

| Variables | IE | AL | EL | MA | PH | CH | BI | AP |
|---|---|---|---|---|---|---|---|---|
| FGPA | .351 | .554 | .412 | .641 | .662 | .710 | .611 | .291 |
| IE | | .337 | .116 | .323 | .322 | .210 | .246 | .239 |
| AL | | | .257 | .457 | .530 | .572 | .467 | .407 |
| EL | | | | .226 | .327 | .262 | .370 | .060 |
| MA | | | | | .628 | .582 | .515 | .238 |
| PH | | | | | | .609 | .496 | .156 |
| CH | | | | | | | .527 | .242 |
| BI | | | | | | | | .240 |

[a]All the intercorrelations are of $p < .01$.

Contrasting the entries (validity coefficients) of Tables 1, 2 and 3, the rank of the validity coefficient of each variable is presented in Table 4. The best variables related to the criterion are CH, PH, BI and MA which are considered mainly as scientific subjects. MA is classified as a scientific subject that requires scientific ability, and is considered as a language of science. There is a small gap in the total rank between BI and MA, but these ranks appear to be dependent on sex. The predictor variable (EL), which has the fifth total rank, clarified the overlapping ranks and led to a difficulty in determining a cut-off point which divided the eight predictor variables into two sets, a scientific set which was closely related to the criterion and a literate set which may reveal a significant contribution to the same criterion. The rank table presented the following points related to the EL variable.

1. It had the maximum rank variability.
2. Overlapping of its ranks with the ranks of the variables preceding it was more clear than that with the ranks of the variables that came after it.
3. The gap between its total rank and the total rank of the next variable (AL) was larger than the gap between its total rank and the total rank of the variable which precede it.
4. The intercorrelations of this variable with the four predictor variables (CH, PH, BI and MA) were less than the corresponding intercorrelations of the variable (AL) for all the groups (total group and the subgroups), this made the contribution of the EL to the criterion, when the four mentioned variables were

Table 4. Rank of the validity coefficient of each predictor variable for the total group and the two subgroups (males and females)

| Variable | Rank of validity coefficient for the total group | Rank of validity coefficient for the males subgroup | Rank of validity coefficient for the females subgroup | Total of the ranks |
|---|---|---|---|---|
| CH | 1 | 1 | 1 | 3 |
| PH | 2 | 2 | 2 | 6 |
| BI | 3 | 4 | 4 | 11 |
| MA | 4 | 5 | 3 | 12 |
| EL | 5 | 3 | 6 | 14 |
| AL | 6 | 6 | 5 | 17 |
| IE | 7 | 8 | 7 | 22 |
| AP | 8 | 7 | 8 | 23 |

already entered, more than the contribution of AL under the same circumstances (see Tables 1, 2 and 3).

5. EL was considered a literary subject, but it was also a scientific tool. This led to a dilemma of whether to classify EL with the scientific set or with the literary set.

The first step of the exploratory analysis which provided the aforementioned dilemma led to further exploratory analysis.

The results of GLM analysis for the total group and the two subgroups (males and females) are displayed in Tables 5, 6 and 7, respectively. The predictor variables which had significant unique contribution to the criterion for the total group are: EL ($p < .0001$), PH ($p < .0001$), and CH ($p < .0001$), while none of the remaining variables had significant unique contribution, even at the .30 level.

Table 5. The F values and probability of the unique contribution of each variable to the criterion (FGPA) for the total group (N = 344) utilizing the General Linear Model (GLM) procedure

| Variable | F value | Probability |
|---|---|---|
| IE | .01 | .9086 |
| AL | .02 | .8910 |
| EL | 15.77 | .0001 |
| MA | .35 | .5564 |
| PH | 20.77 | .0001 |
| CH | 40.56 | .0001 |
| BI | .97 | .3266 |
| AP | .47 | .4922 |

Table 6. The F values and probability of the unique contribution of each variable to the criterion (FGPA) for males (N = 220) utilizing the General Linear Model (GLM) procedure

| Variable | F value | Probability |
|---|---|---|
| IE | .71 | .4009 |
| AL | .11 | .7412 |
| EL | 5.83 | .0162 |
| MA | .13 | .7160 |
| PH | 12.21 | .0006 |
| CH | 17.83 | .0001 |
| BI | .13 | .7229 |
| AP | .59 | .4421 |

Table 7. The F values and probability of the unique contribution of each variable to the criterion (FGPA) for females (N = 124) utilizing the General Linear Model (GLM) procedure

| Variable | F value | Probability |
|---|---|---|
| IE | 2.49 | .1173 |
| AL | .12 | .7246 |
| EL | 6.80 | .0103 |
| MA | 5.02 | .0270 |
| PH | 4.92 | .0286 |
| CH | 19.56 | .0001 |
| BI | 5.29 | .0232 |
| AP | 1.07 | .3022 |

The predictor variables which had significant unique contribution to the criterion for the subgroup of males were: CH ($p < .0001$), PH ($p < .001$) and EL ($p < .05$), while none of the remaining variables had a significant unique contribution, even at .40 level.

The predictor variables which had significant unique contribution to the criterion for the subgroup at females are: CH ($p < .0001$), EL ($p < .01$), BI ($p < .05$), MA ($p < .05$) and PH ($p < .05$). None of the remaining variables has a significant unique contribution, even at .10 level.

None of the three predictor variables, IE, AL and AP had a significant unique contribution to the criterion for the total group or for any of the subgroups.

The contribution of this analysis with the previous results led the researcher to the selection of EL, MA, PH, CH and BI as the variables which had common characteristics or related to the criterion. These variables were used in comparing the performance of the four multiple regression techniques.

## Results of Comparison Analysis

The scores on the five variables selected by the exploratory analysis were used for comparing the four regression methods. The best prediction equation was computed by each method for three different sample sizes or ratios. The three different ratios used were labeled small, intermediate and large. The sample size of the first ratio (small) was 50 subjects (18 females + 32 males), selected from the total sample by a stratified random sampling procedure. The sample size of the intermediate ratio was 100 subjects (36 females + 64 males) selected by the same sampling procedure. The sample size of the large ratio was 200 subjects (72 females + 128 males) selected once again with the same sampling procedure.

Results for the small ratio

Maximizing weight validity method Table 8 presents the number of times each of the 31 possible equations was superior in each one of the applications of the MWV method for varying significance levels and seed numbers. Table 9 presents the average (Fisher Z transformed) weight validity coefficient, the associated test of significance and the number of iterations for which the weight validity coefficient was highest for the best and the next best equations. The sample-generated least square estimates of the regression weights of each selected variable, the validity coefficient (R) of the best equation on the total sample and the amount of explained variance ($R^2$) of the criterion (FGPA) by the selected variables were as follows:

$$\widehat{FGPA} = 35.6857 + .1727\,(PH) + .2454\,(CH),$$

$$R = .6675 \text{ and } R^2 = .4456$$

Traditional methods All three traditional methods (forward, backward and stepwise) gave similar results. Table 10 presents the selected variables, the regression weights of these variables and their significances. The analysis yielded the prediction equation,

Table 8. A frequency of superiority for the 31 possible regression equations run at different significance levels (α) and seed numbers (SN) for small sample size utilizing the MWV technique

| Variables in[a] the equation | Run 1[b] | Run 2[c] | Run 3[d] | Total |
|---|---|---|---|---|
| 5 | 0 | 0 | 0 | 0 |
| 4 | 4 | 5 | 3 | 12 |
| 4 5 | 1 | 1 | 2 | 4 |
| 3 | 2 | 1 | 1 | 4 |
| 3 5 | 0 | 0 | 2 | 2 |
| 3 4 | 51[e] | 25[e] | 31[e] | 107[e] |
| 3 4 5 | 19 | 9 | 16 | 44 |
| 2 | 0 | 0 | 0 | 0 |
| 2 5 | 0 | 0 | 0 | 0 |
| 2 4 | 10 | 7 | 4 | 21 |
| 2 4 5 | 5 | 3 | 4 | 12 |
| 2 3 | 1 | 0 | 1 | 2 |
| 2 3 5 | 0 | 0 | 0 | 0 |
| 2 3 4 | 10 | 3 | 2 | 15 |
| 2 3 4 5 | 3 | 4 | 4 | 11 |
| 1 | 0 | 0 | 0 | 0 |
| 1 5 | 0 | 0 | 0 | 0 |
| 1 4 | 11 | 7 | 6 | 24 |
| 1 4 5 | 6 | 2 | 5 | 13 |
| 1 3 | 3 | 4 | 2 | 9 |
| 1 3 5 | 0 | 0 | 0 | 0 |
| 1 3 4 | 27 | 8 | 10 | 45 |
| 1 3 4 5 | 11 | 3 | 8 | 22 |
| 1 2 | 0 | 0 | 0 | 0 |
| 1 2 5 | 0 | 0 | 0 | 0 |
| 1 2 4 | 5 | 7 | 7 | 19 |
| 1 2 4 5 | 6 | 6 | 6 | 18 |
| 1 2 3 | 0 | 0 | 1 | 1 |
| 1 2 3 5 | 0 | 0 | 0 | 0 |
| 1 2 3 4 | 7 | 1 | 1 | 9 |
| 1 2 3 4 5 | 1 | 0 | 1 | 2 |

[a]EL (1), MA (2), PH (3), CH (4) and BI (5).

[b]SN = 10652319 and α = .01.

[c]SN = 22176865 and α = .01.

[d]SN = 22176865 and α = .05.

[e]Selected as the best predictors to be in the prediction equation.

Table 9. The average weight validity coefficient, probability and frequency of superiority of the best and the next best equations for small sample size utilizing the MWV technique

| | Equation | Variables | Average weight validity coefficient ($\bar{R}_C$) | Probability | Frequency of superiority |
|---|---|---|---|---|---|
| Run 1 | best | 3 and 4 | .5924 | .0001 | 51 |
| | next best | 1, 3 and 4 | .5769 | .0001 | 27 |
| Run 2 | best | 3 and 4 | .5823 | .0001 | 25 |
| | next best | 1, 3 and 4 | .5611 | .0001 | 8 |
| Run 3 | best | 3 and 4 | .5842 | .0001 | 31 |
| | next best | 1, 3 and 4 | .5663 | .0001 | 10 |

Table 10. Analysis of traditional multiple regression selection methods using the scores of five predictor variables and the criterion under the condition of small sample size

| Source of variation | Degrees of freedom | Mean square | F value | Probability |
| --- | --- | --- | --- | --- |
| Regression | 2 | 682.169 | 18.890 | .0001 |
| Error | 47 | 36.111 | | |
| Total | 49 | | | |

| | B value | F value | Probability |
| --- | --- | --- | --- |
| Intercept | 35.6798 | | |
| PH | .1727 | 6.14 | .0169 |
| CH | .2455 | 8.80 | .0047 |

$$\hat{FGPA} = 35.6798 + .1727\ (PH) + .2455\ (CH),$$

the validity coefficient (R) = .6675 and $R^2$ = .446.

The results revealed the equality of the performance of the MWV and the traditional methods when applied to a small sample size.

Results for intermediate ratio

Maximizing weight validity method Table 11 presents the number of times each of the 31 possible regression equation was superior in each of the applications of the MWV method for varying seed numbers. Table 12 presents the average (Fisher Z transformed) weight validity coefficient, the associated test of significance and the number of iterations for which the weight validity coefficient was highest for the best and the next best equations. The sample-generated least square estimates of regression weights of each selected variable, the validity coefficient (R) of the best equation for the total sample and the amount of explained variance of the criterion (FGPA) by the selected variables were as follows:

$$\hat{FGPA} = 25.0062 + .0598\ (EL) + .1201\ (PH) + .3302\ (CH),$$

$$R = .6350 \text{ and } R^2 = .404.$$

Traditional methods Table 13 presents the variables, the regression coefficients of these variables and their significances. The following regression equation, R and $R^2$ were the same by the three traditional techniques,

$$\hat{FGPA} = 31.3349 + .1381\ (PH) + .3411\ (CH),$$

$$R = .6240 \text{ and } R^2 = .389.$$

Table 11. A frequency of superiority for the 31 possible regression equations run at different seed numbers (SN) and fixed significance level ($\alpha$) for intermediate sample size utilizing the MWV technique

| Variables | Run 1[a] | Run 2[b] | Total |
|---|---|---|---|
| 5 | 0 | 0 | 0 |
| 4 | 4 | 8 | 12 |
| 4 5 | 3 | 3 | 6 |
| 3 | 0 | 0 | 0 |
| 3 5 | 0 | 0 | 0 |
| 3 4 | 56 | 102 | 158 |
| 3 4 5 | 20 | 27 | 47 |
| 2 | 0 | 0 | 0 |
| 2 5 | 0 | 0 | 0 |
| 2 4 | 1 | 2 | 3 |
| 2 4 5 | 0 | 0 | 0 |
| 2 3 | 0 | 0 | 0 |
| 2 3 5 | 0 | 0 | 0 |
| 2 3 4 | 5 | 8 | 13 |
| 2 3 4 5 | 6 | 9 | 15 |
| 1 | 0 | 0 | 0 |
| 1 5 | 0 | 0 | 0 |
| 1 4 | 25 | 16 | 41 |
| 1 4 5 | 5 | 2 | 7 |
| 1 3 | 0 | 0 | 0 |
| 1 3 5 | 0 | 0 | 0 |
| 1 3 4 | 80[c] | 133[c] | 213[c] |
| 1 3 4 5 | 6 | 14 | 20 |
| 1 2 | 0 | 0 | 0 |
| 1 2 5 | 0 | 0 | 0 |
| 1 2 4 | 1 | 9 | 10 |
| 1 2 4 5 | 0 | 0 | 0 |
| 1 2 3 | 0 | 0 | 0 |
| 1 2 3 5 | 0 | 0 | 0 |
| 1 2 3 4 | 6 | 14 | 20 |
| 1 2 3 4 5 | 2 | 5 | 7 |

[a]SN = 22176868 and $\alpha$ = .05.

[b]SN = 10652319 and $\alpha$ = .05.

[c]Selected as the best predictors to be in the prediction equation.

Table 12. The average weight validity, probability and frequency of superiority of the best and the next best equations for intermediate sample size utilizing the MWV technique

| | Equation | Variables | Average weight validity coefficient ($\bar{R}_C$) | Probability | Frequency of superiority |
|---|---|---|---|---|---|
| Run 1 | best | 1, 3 and 4 | .5776 | .0001 | 80 |
| | next best | 3 and 4 | .5843 | .0001 | 56 |
| Run 2 | best | 1, 3 and 4 | .5781 | .0001 | 133 |
| | next best | 3 and 4 | .5839 | .0001 | 102 |

Table 13. Analysis of traditional multiple regression selection methods using the scores of five predictor variables and the criterion under the condition of intermediate sample size

| Source of variation | Degrees of freedom | Mean square | F value | Probability |
|---|---|---|---|---|
| Regression | 2 | 1500.378 | 30.90 | .0001 |
| Error | 97 | 48.559 | | |
| Total | 99 | | | |

| | B value | F value | Probability |
|---|---|---|---|
| Intercept | 31.3349 | | |
| PH | .1381 | 3.80 | .0179 |
| CH | .3409 | 28.22 | .0001 |

The variables in the equation selected by the traditional methods were the same variables in the next best equations selected by MWV. The total number of times highest, of the next best equation, over two runs, was 158 while the total number of times highest of the best equation selected by MWV was 213, which referred to the superiority of the equation selected by the MWV method.

Results for the large ratio

Maximizing weight validity method  Table 14 presents the number of times each of the 31 possible regression equations was superior in each of the applications of the MWV method at different seed numbers, and Table 15 presents the average (Fisher Z transformed) weight validity coefficient, associated test of significance and the number of iterations for which the weight validity was highest for the best and the next best equations. The sample-generated least square estimates of regression weights of each selected variable and the validity coefficient of the best equation on the total sample were as follows:

$$\hat{FGPA} = 17.5916 + .1176\ (EL) + .1559\ (PH) + .2003\ (CH) + .0748\ (BI),$$
$$R = .6675 \text{ and } R^2 = .4456.$$

Traditional methods  Table 16 presents the selected variables, the regression coefficients of these variables and their significances. The following regression equation, R and $R^2$ were the same by the three traditional methods.

$$\hat{FGPA} = 20.6446 + .1229\ (EL) + .1617\ (PH) + .2280\ (CH),$$
$$R = .6678 \text{ and } R^2 = .4460.$$

Table 14. A frequency of superiority for the 31 possible regression equations run at different seed numbers (SN) and at fixed significance level (α) for large sample size utilizing the MWV technique

| Variables in the equations | Run 1[a] | Run 2[b] | Total |
|---|---|---|---|
| 5 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 |
| 4 5 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 |
| 3 5 | 0 | 0 | 0 |
| 3 4 | 1 | 0 | 0 |
| 3 4 5 | 4 | 1 | 5 |
| 2 | 0 | 0 | 0 |
| 2 5 | 0 | 0 | 0 |
| 2 4 | 0 | 0 | 0 |
| 2 4 5 | 0 | 0 | 0 |
| 2 3 | 0 | 0 | 0 |
| 2 3 5 | 0 | 0 | 0 |
| 2 3 4 | 0 | 0 | 0 |
| 2 3 4 5 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 |
| 1 5 | 0 | 0 | 0 |
| 1 4 | 2 | 0 | 2 |
| 1 4 5 | 1 | 1 | 2 |
| 1 3 | 0 | 0 | 0 |
| 1 3 5 | 8 | 0 | 8 |
| 1 3 4 | 116 | 29 | 145 |
| 1 3 4 5 | 149[c] | 47[c] | 196[c] |
| 1 2 | 0 | 0 | 0 |
| 1 2 5 | 0 | 0 | 0 |
| 1 2 4 | 3 | 1 | 4 |
| 1 2 4 5 | 3 | 1 | 4 |
| 1 2 3 | 0 | 0 | 0 |
| 1 2 3 5 | 3 | 0 | 3 |
| 1 2 3 4 | 39 | 15 | 54 |
| 1 2 3 4 5 | 24 | 9 | 33 |

[a]SN = 10652319 and α = .05.

[b]SN = 22176865 and α = .05.

[c]Selected as the best predictors to be in the prediction equation.

Table 15. The average weight validity coefficient, probability and frequency of superiority of the best and the next best equations for large sample size utilizing the MWV technique

| | Equation | Variables | Average weight validity coefficient ($\bar{\bar{R}}_C$) | Probability | Frequency of superiority |
|---|---|---|---|---|---|
| Run 1 | best | 1,3,4 and 5 | .6318 | .0001 | 149 |
| | next best | 1,3 and 4 | .6337 | .0001 | 116 |
| Run 2 | best | 1,3,4 and 5 | .6327 | .0001 | 47 |
| | next best | 1,3 and 4 | .6335 | .0001 | 29 |

Table 16. Analysis of traditional multiple regression selection methods using the scores of five predictors and the criterion under the condition of large sample size

| Source of variation | Degrees of freedom | Mean square | F value | Probability |
|---|---|---|---|---|
| Regression | 3 | 2130.532 | 51.26 | .0001 |
| Error | 196 | 41.565 | | |
| Total | 199 | | | |

| | B value | F value | Probability |
|---|---|---|---|
| Intercept | 20.6446 | | |
| EL | .1229 | 21.03 | .0001 |
| PH | .1617 | 17.03 | .0001 |
| CH | .2280 | 29.70 | .0001 |

The variables in the equation selected by the traditional methods were the variables in the next best equation selected by MWV. The total number of times highest, of the next best equation over two runs, was 145, while it was 196 for the best equation. This indicated the superiority of the MWV over the traditional methods even when the sample size was relatively large.

## Results when Sex was Utilized as a Moderator Variable

### Maximizing weight validity method

Table 17 presents the number of times each of the 31 possible regression equations was superior in each of the applications of the MWV for males, females and the total group. Table 18 presents the mean weight validity, associated test of significance and the number of iterations for which the weight validity was highest for the best equation of each group. The sample-generated least square estimates of the regression weights for each selected variable and the validity coefficient of the best equation for each group were as follows:

A. Males

$$\widehat{FGPA} = 29.4566 + .0769\ (EL) + .1387\ (PH) + .2042\ (CH),$$

$R = .5118$ and $R^2 = .262$

B. Females

$$\widehat{FGPA} = 7.5459 + .0783\ (EL) + .0457\ (MA) + .1217\ (PH) + .2327\ (CH) + .1566\ (BI),$$

$R = .8171$ and $R^2 = .6677$.

Table 17. A frequency of superiority for the 31 possible regression equations run at different seed numbers (SN) and at fixed significance level (α) for males, females and total group utilizing the MWV technique

| Variables | Male | | | Female | | | Total group run (T)[c] |
|---|---|---|---|---|---|---|---|
| | Run 1[a] | Run 2[b] | Total | Run 1 | Run 2 | Total | |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 4 | 6 | 1 | 7 | 0 | 0 | 0 | 2 |
| 3 4 5 | 2 | 0 | 2 | 1 | 0 | 1 | 4 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 4 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 3 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 3 4 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| 2 3 4 5 | 0 | 0 | 0 | 2 | 0 | 2 | 1 |
| 1 | 0 | 0 | 0 | 0 | 3 | 3 | 0 |
| 1 | 5 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 4 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| 1 4 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 3 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 3 4 | 17 | 7 | 24[d] | 0 | 0 | 0 | 102 |
| 1 3 4 5 | 0 | 1 | 1 | 0 | 0 | 0 | 133[d] |
| 1 2 | 0 | 0 | 0 | 0 | 3 | 3 | 0 |
| 1 2 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 2 4 | 2 | 0 | 2 | 0 | 0 | 0 | 0 |
| 1 2 4 5 | 0 | 0 | 0 | 2 | 3 | 5 | 0 |
| 1 2 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 2 3 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 2 3 4 | 4 | 1 | 5 | 2 | 1 | 3 | 65 |
| 1 2 3 4 5 | 2 | 0 | 2 | 9 | 10 | 19[d] | 46 |

[a]Run 1 at SN = 22176865 and α = .05.

[b]Run 2 at SN = 10652319 and α = .05.

[c]Run T at SN = 22176765 and α = .05.

[d]Selected as the best predictors to be in the prediction equation.

Table 18. Average weight validity coefficient, probability and frequency of superiority of the best equation for males, females and the total group utilizing the MWV technique

| | Variables in the best equation | Mean weight validity ($\bar{R}_C$) | Probability | Frequency of superiority |
|---|---|---|---|---|
| Male | 1,3 and 4 | .4628 | .0001 | 24 |
| Female | 1,2,3,4 and 5 | .7835 | .0001 | 19 |
| Total | 1,3,4 and 5 | .6219 | .0001 | 133 |

C. Total group

$$\hat{FGPA} = 20.4594 + .0889\ (EL) + .1629\ (PH) + .2356\ (CH) + .0536\ (BI),$$

$R = .6437$ and $R^2 = .414$.

Traditional methods

Tables 19, 20 and 21 present the selected variables, the regression coefficients of these variables and their probabilities for the subsets of males, females as well as the total group respectively. The regression equation, R and $R^2$ of each group were as follows:

A. Males

$$\hat{FGPA} = 29.4453 + .0773\ (EL) + .1385\ (PH) + .2039\ (CH),$$

$R = .512$ and $R^2 = .2620$.

B. Females

$$\hat{FGPA} = 7.5828 + .0782\ (EL) + .0458\ (MA) + .1216\ (PH) + .2326\ (CH) + .1561\ (BI).$$

$R = .817$ and $R^2 = .6679$.

C. Total group

$$\hat{FGPA} = 22.2442 + .0958\ (EL) + .1671\ (PH) + .2551\ (CH),$$

$R = .642$ and $R^2 = .4123$.

These results indicated that males, females were two different populations, i.e. sex was a moderator variable. The variables selected by the traditional and MWV methods were the same for the subset of males and females but those selected for males differed from those selected for females. Females were more predictable ($R^2 = .67$) than males ($R^2 = .26$). The performance of the prediction equation for males, irrespective

Table 19. Analysis of traditional multiple regression methods using the scores of five predictor variables and the criterion for males

| Source of variation | Degrees of freedom | Mean square | F value | Probability |
|---|---|---|---|---|
| Regression | 3 | 1246.645 | 25.58 | .0001 |
| Error | 216 | 48.730 | | |
| Total | 219 | | | |

| | B value | F value | Probability |
|---|---|---|---|
| Intercept | 29.4453 | | |
| EL | .07725 | 7.63 | .0062 |
| PH | .13851 | 13.04 | .0004 |
| CH | .20390 | 22.27 | .0001 |

Table 20. Analysis of traditional multiple regression selection methods using the scores of five predictor variables and the criterion for females

| Source of variation | Degrees of freedom | Mean square | F value | Probability |
|---|---|---|---|---|
| Regression | 4 | 1858.507 | 55.16 | .0001 |
| Error | 119 | 26.442 | | |
| Total | 123 | | | |

| | B values | F value | Probability |
|---|---|---|---|
| Intercept | 7.5829 | | |
| EL | .0782 | 6.49 | .0121 |
| MA | .0458 | 6.80 | .0103 |
| PH | .1216 | 6.53 | .0119 |
| CH | .2326 | 22.53 | .0001 |
| CI | .1561 | 6.81 | .0102 |

Table 21. Analysis of traditional multiple regression methods using the scores of five predictor variables, and the criterion for the total group

| Source of variation | Degrees of freedom | Mean square | F value | Probability |
|---|---|---|---|---|
| Regression | 3 | 3397.6986 | 79.51 | .0001 |
| Error | 340 | 42.7324 | | |
| Total | 343 | | | |

| | B value | F value | Probability |
|---|---|---|---|
| Intercept | 22.2441 | | |
| EL | .0958 | 20.70 | .0001 |
| PH | .1671 | 31.10 | .0001 |
| CH | .2551 | 60.46 | .0001 |

of the selection method, was not as effective when compared with the prediction of females or compared with the average amount of explained variance of FGPA or college success as reflected by student grades on different subjects.

The variables selected by MWV to be in the prediction equation were different than those selected by the traditional methods when one considered the total group. Three variables were selected by the traditional methods but four variables were selected by MWV. Three predictor variables were common in both equations. The prediction equation of the four variables was superior to that of the three variables. This superiority could not be explained by the proportion of the explained variance but through the frequency of superiority in subsequent samples. Because of the varying quantity and quality of the selected variables in the two equations, differing regression weights resulted. The regression weights of the variables in the superior equation were more valid (generalizable).

Comparing the performance of the traditional methods and the MWV method within groups and among groups, which were classified according to ratio (small, intermediate and large) and sex (males, females and total group) revealed the following:

1. The traditional and MWV methods selected the same variables for small ratio and none of these methods appeared to be superior, which led to one of the following conclusions.

   A. None of the methods was to be preferred over the rest and using any of these methods did not make any difference.

B. None of these methods were appropriate and another method was required.

2. The MWV method was superior to the traditional methods in the case of the intermediate ratio. The variables (3 and 4) selected by the traditional methods were the same variables selected by all methods in the case of the small ratio, however, three variables (1, 3 and 4) were selected by the MWV technique.

3. In the case of large ratio, variable 1 was added to the set of predictor variables (3 and 4) selected by the traditional methods for the intermediate ratio. Variable 5 was added to the predictor variables (1, 3 and 4) by the MWV technique for intermediate ratio. The MWV technique was superior to the traditional methods for intermediate and large ratio cases.

4. The variables (1, 3 and 4) selected by the traditional methods for large ratio were the same variables selected by these methods for the total group. The variables (1, 3, 4 and 5) selected by the MWV method for large ratio were the same variables selected by this method for the total group, which meant that a ratio of 40 subjects/variable or more provided the same results utilizing the same method, and that increasing the ratio did not make any difference in the quality of prediction utilizing the same technique.

5. When males and females were treated separately, the variable, selected by the traditional and MWV method within the same sex group were the same, but the variables selected between the sex

groups were different. All methods revealed that females were more predictable than males and none of the regression methods was superior. The variables selected to be in the prediction equation of males were a subset of the variables selected to be in the prediction equation of females, which meant that a combination of ratio and homogeneity determined whether a method was superior or not, i.e. the superiority of a given method was situational specific.

The MWV method was recommended as a superior method from the results of one previous comparison study. The data of that study possessed the characteristics of intermediate ratio, heterogeneity and multicollinearity. Combining the results of that comparison study with the results of the current study brings into question the superiority of the MWV technique. The results of the current study in case of an intermediate ratio were different from the results for small ratio regarding group homogeneity. Sex was a moderator variable in this study and when the both subgroups were treated as a single group a loss of information occurred and the superiority of the MWV method became more questionable. In the previous comparison study, the data had two potential moderator variables, sex and race, and they may have been completely different if the groups of sex and race had been treated separately, because this would have led a classification of all of the subgroups within the small ratio level (11 black males, 24 black females, 18 white males and 30 white females).

Table 22 presents the intercorrelations of the predictor variables and the criterion for an intermediate sample size. The intercorrelations

Table 22. Intercorrelations of the five predictor variables and the criterion for intermediate sample size

| | MA | PH | CH | BI | FGPA |
|---|---|---|---|---|---|
| EL | .153 | .287 | .235 | .395 | .292 |
| MA | | .561 | .383 | .340 | .291 |
| PH | | | .495 | .305 | .460 |
| CH | | | | .408 | .594 |
| BI | | | | | .326 |

of the prediction variables had no multicollinearity, the maximum zero-order correlation was .561, (correlation between PH and MA), and this was classified as of moderate size. The variables PH and CH were selected as the best predictor, by the traditional methods. These two variable had the highest zero-order correlations with the criterion. The third predictor variable (EL), which was added to the previous two predictors by the MWV method, was not the variable which had the third highest zero-order correlation with the criterion, but it did have second and third lowest intercorrelations with the predictors variables. The variable (MA) had the lowest zero-order correlation with the EL and with the criterion but was not selected because of its high correlation with (PH), which had been previously selected. The results of the previous comparison study were not as complex because the intercorrelations from that studied variables were categorized into two distinct sets. However, the current study had no multicollinearity and also more intercorrelations. This introduced a new situation of comparison which revealed the superiority of the MWV technique over the traditional techniques under certain conditions. The problem of multicollinearity in regression analysis led to the following statement in the words of Nie et al. (1975):

> When extreme multicollinearity exists there is no acceptable way to perform regression analysis using the given set of independent variables.

### Results of Mathematical and Empirical Estimation of the Shrinkage in the Validity Coefficient

The estimates of the population multiple correlations which were calculated empirically, by the MWV technique designated as $\bar{R}_C$, and mathematically, by the three shrinkage formulas (Wherry, Lord-Nicholson and Stein-Darlington) designated as $\bar{R}_W$, $\bar{R}_{L-W}$ and $\bar{R}_{S-W}$ respectively, for each of the six groups which were classified according to ratio and sex were presented in Table 23. The entries of this table were the squared quantities.

Table 24 revealed the amount of shrinkage, that could be expected when the weights developed in a sample were applied to subsequent samples from the same population, which were calculated mathematically and empirically for all groups. The results indicated that the amount of shrinkage was underestimated differentially by all the formulas over all the groups. Applying the criterion of Schmitt et al. (1976), which said that an underestimation of greater than .03 is sizeable, the Wherry formula underestimates the amount of shrinkage for all groups. The amount of shrinkage estimated by the other two formulas for each group

Table 23. The squared validity coefficient ($R_T^2$), the empirical estimation of the squared weight validity coefficient by MWV technique ($\bar{R}_C^2$) and the mathematical estimation of the squared weight validity coefficient by the three shrinkage formulas (Wherry, Lord-Nicholson and Stein-Darlington) designated as $\bar{R}_W^2$, $\bar{R}_{L-N}^2$ and $\bar{R}_{S-D}^2$ respectively

| Classification | | $R_T^2$ | $\bar{R}_W^2$ | $\bar{R}_{L-N}^2$ | $\bar{R}_{S-D}^2$ | $\bar{R}_C^2$ |
|---|---|---|---|---|---|---|
| | small ratio | .4456 | .4341 | .3873 | .3848 | .3437 |
| Ratio levels | intermediate ratio | .3890 | .3828 | .3577 | .3570 | .3339 |
| | large ratio | .4460 | .4302 | .4263 | .4260 | .3997 |
| | total group | .4123 | .4089 | .4002 | .4001 | .3868 |
| Group category | males | .2620 | .2552 | .2381 | .2379 | .2144 |
| | females | .6679 | .6567 | .6371 | .6361 | .6140 |

Table 24. The amount of shrinkage of $R_T^2$ calculated by MWV technique and the three mathematical formulas for each ratio level and for each group category

| Comparison | Levels of ratio | | | Group category | | |
|---|---|---|---|---|---|---|
| | Small | Intermediate | Large | Total | Males | Females |
| $R_T^2 - \bar{R}_C^2$ | .1019 | .0551 | .0463 | .0255 | .0476 | .0539 |
| $R_T^2 - \bar{R}_W^2$ | .0119 | .0062 | .0158 | .0034 | .0068 | .0112 |
| $R_T^2 - \bar{R}_{L-N}^2$ | .0583 | .0313 | .0197 | .0121 | .0239 | .0308 |
| $R_T^2 - \bar{R}_{S-D}^2$ | .0608 | .0320 | .0200 | .0122 | .0241 | .0318 |

$R_T$ : validity coefficient

$\bar{R}_C$ : weight validity calculated empirically utilizing the MWV technique

$\bar{R}_W$ : weight validity calculated mathematically by Wherry formula

$\bar{R}_{L-N}$ : weight validity calculated mathematically by Lord-Nicholson formula

$\bar{R}_{S-D}$ : weight validity calculated mathematically by Stein-Darlington formula was approximately equal, but it can be said that the Stein-Darlington formula provides the best estimates. The amount of shrinkage was under-estimated by each of the three mathematical formulas in the case of the small ratio, which revealed that none of the shrinkage formulas can be an alternative to the MWV technique, as an empirical method of estimation, for small ratio. However, the Stein-Darlington formula may be recommended for the other ratio levels and group categories.

It was indicated earlier that the MWV was not superior to the traditional methods in the case of the small ratio sample, and none of the mathematical shrinkage formulas can be used as an alternative of the MWV. This led to the dilemma that none of the four methods was appropriate for small ratio sample. However, the superiority of the MWV technique cannot be generalized to all situations but may be superior under certain conditions. Drehmer and Morris (1981) described a new method with an accompanying computer program which they recommended in the small ratio sample situation.

## Results of the Validity of the Prediction Equation Over Time

Table 25 reveals the results of analyzing the data set of the academic year 1978-1979 (set A) and the data set of the academic year 1979-1980 (set B) utilizing the stepwise regression procedure. The selected variables, standard regression weight of each variable, the probabilities and the amount of explained variance of FGPA by the selected variables for males, females as well as the total group are displayed in the same table. The results indicated that:

Table 25. Variables, standard regression weights, probabilities and the amount of explained variance of the FGPA by the selected variables for males, females and total group for the data set of the academic year 1978-1979 (set A) and the data set of the academic year 1979-1980 (set B) utilizing the stepwise regression technique

| | Set A | | | | Set B | | | |
|---|---|---|---|---|---|---|---|---|
| Comparison | Variables | BETA | F | $R^2$ | Variables | B value | F | $R^2$ |
| | PH | .40439 | 40.05** | | CH | .3041 | 22.269** | |
| | EL | .25030 | 16.92* | | PH | .2284 | 13.039** | |
| Males | MA | .15978 | 6.20** | | EL | .1671 | 7.632** | |
| | | | | .292 | | | | .262 |
| | CH | .37621 | 16.095** | | CH | .3469 | 22.535** | |
| Females | AL | .31582 | 11.460** | | PH | .1935 | 6.531* | |
| | EL | .25632 | 8.341** | | BI | .1784 | 6.813* | |
| | | | | | MA | .1927 | 6.797* | |
| | | | | | EL | .1481 | 6.495* | |
| | | | | .447 | | | | .668 |
| Total | PH | .33202 | 31.446** | | CH | .3748 | 60.460** | |
| | EL | .25911 | 25.291** | | PH | .2652 | 31.105** | |
| | CH | .19847 | 10.926** | | EL | .1989 | 20.698** | |
| | | | | .306 | | | | .412 |

*$p < .05$.

**$p < .01$.

1. Females of set A and set B were more predictable than males. This confirmed the previous results which indicated that sex was a moderator variable.
2. The performance of the prediction equation for males in both sets was not practically significant. An explained variance of less than 35% in academic achievement situation was not considered practical. The selected variables to be in the prediction equation for males of set A were different from those of set B, which meant that the prediction equation of males from set A was not valid for males in set B. The existence of two common predictor variables (CH, PH) and the approximate equality of the explained variance were not sufficient indices of the validity of the prediction equation over both sets.
3. The predictor variables for females of set A were different, qualitatively and quantitatively, from those of set B. The amount of explained variance was also different. These two indices were enough to insure the invalidity of the prediction equation of females from set A over the females from set B.
4. When males and females were treated as a single group in both set A and set B, the same variables (PH, CH and EL) were selected for each set. The square of the correlation coefficient between the predicted FGPA and the actual FGPA of set A (.306) differed from that of set B (.412). The regression weights of the three predictor variables were not consistent over the two

sets. The calculated prediction equation for the total group had better validity than that for males and females treated separately because the same variables were used in the equations of both sets. However, other indices indicated the invalidity of the selected prediction equation. In addition to the inconsistency of the regression weights and the $R^2$ values, the accuracy of prediction for the two equations differed. The Mean Absolute Error (MAE) of prediction by the equation selected for set A was 5.65, while the MAE of prediction by the equation selected for set B was 4.97. This revealed that the equation of set B had better accuracy. Another index of the effectiveness of a prediction equation was the proportion of students whose predicted scores fall within a certain range utilizing the prediction error. Figures 1 and 2 revealed the scatterplots of the standardized predicted FGPA against the standardized residuals. Figure 1 revealed that 73% (201/275) of the students of set A were within one standard deviation of residuals, and Figure 2 revealed that 72% (246/344) of the students of set B were within one standard deviation of residuals. The two proportions are equals practically. The overpredicted proportion of students of set A and set B were 15% and 16%, respectively. The underpredicted proportion of students of set A and set B were both 12%. The last measures confirmed the validity of the prediction equation, calculated for one of the two sets, over the other set. The overall indices of the validity of the

Figure 1. Scatterplots of the standardized predicted FGPA (abscissa) against the standardized residuals (ordinate) for the data set of the academic year 1978-1979

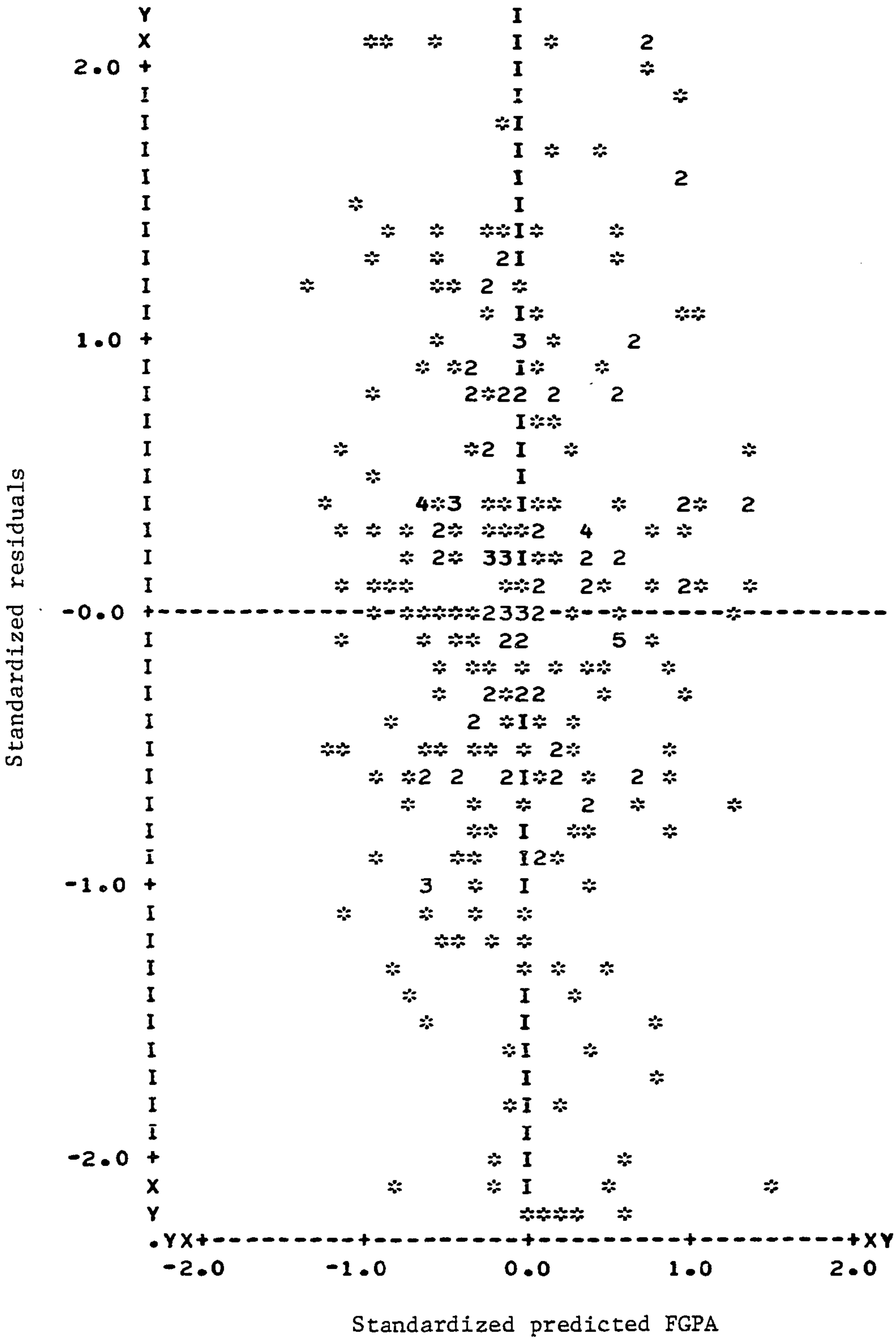
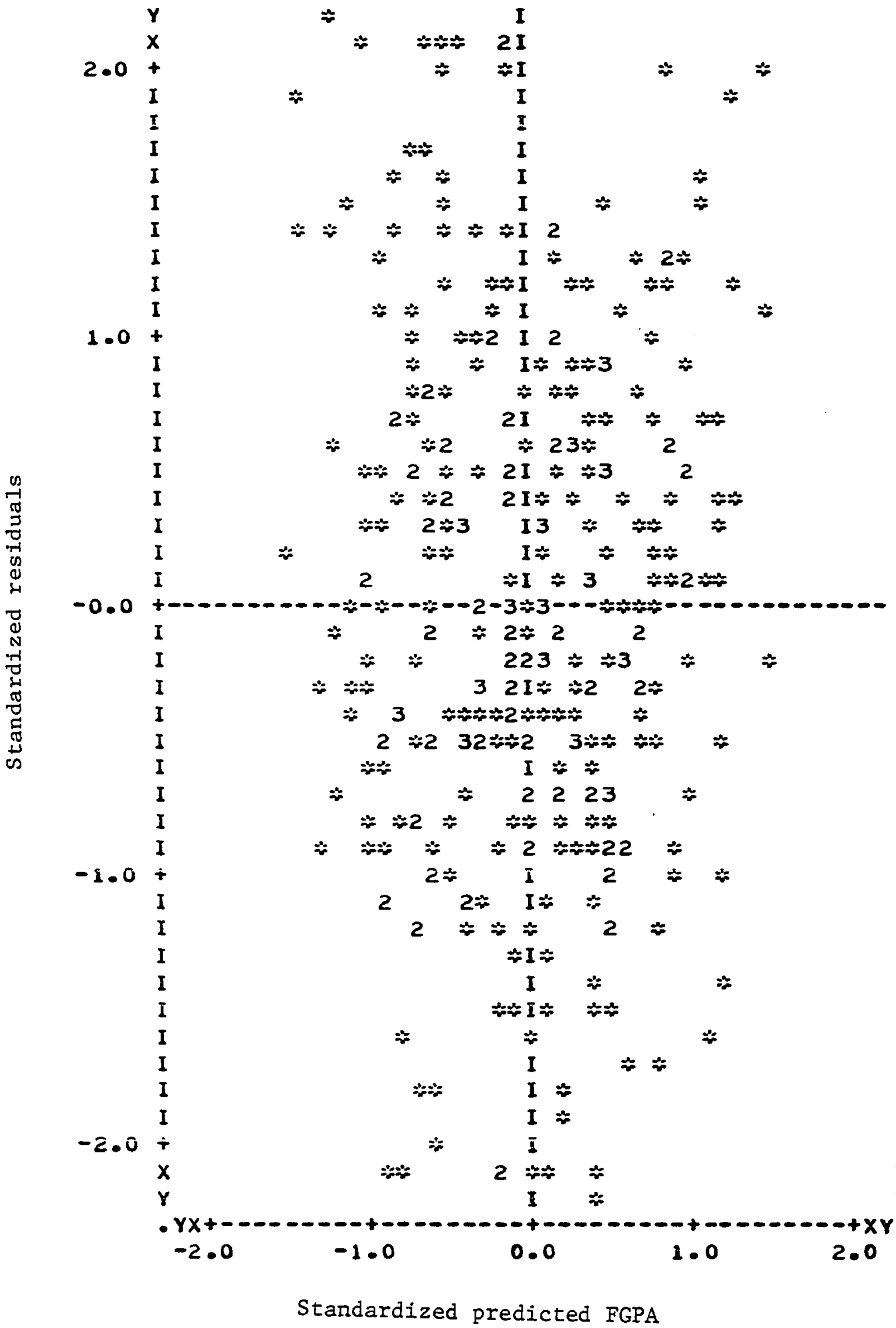

2.0
1.0
-0.0
-1.0
-2.0
-2.0
-1.0
0.0
1.0
2.0
Standardized residuals
Standardized predicted FGPA

Figure 2. Scatterplots of the standardized predicted FGPA (abscissa) against the standardized residuals (ordinate) for the data set of the academic year 1979-1980



Standardized residuals
2.0
1.0
-0.0
-1.0
-2.0
-2.0
-1.0
0.0
1.0
2.0
Standardized predicted FGPA

prediction equation for the total in each set indicated that the validity of a prediction equation over time is a matter of subjective judgment. The prediction equation of the total group was relatively valid and the regression weights calculated from set B were recommended for subsequent selection, assuming the absence of other alternatives.

The invalidity of the prediction equations for males, females, and total group could be due to one or more of the following reasons.

1. Interference of academic achievement with social characteristics. The inconsistency of the interference effect, through the transition period of high school and college could cause the unstability of predicting student achievement.
2. The process of evaluating freshman student achievement in the college of science at Yarmouk University may be unsystematic. The measurement tools utilized, mainly instructor-made tests, could be of weak reliability and validity leading to the inference that FGPA did not reflect the student achievement as it was assumed in this study, especially if the university evaluation process was dependent on the instructors' grading policies.
3. The tests of GSECE were not standardized. The selection of a student by a specific institution depended on his/her grades on these tests. Thus, the students recognized the achieved grades as extremely important. This fact may create other factors such as: anxiety, mental fatigue, tutoring, coaching, etc.,

which made the reflection of the student's achievement by high school grades questionable. The prediction of student achievement could be improved by improving the tests of GSECE qualitatively and quantitatively and/or by using other appropriate tests.

These factors in addition to other factors such as changes in freshman curriculum and changes in the distribution of academic ability among entering students of set A and set B, which were assumed to be constant, could be large enough to cause old prediction equations to be inaccurate.

### Results for the Quality of Prediction of Factor Scores, the Data-Level Variables and the High School GPA

The previous results indicated that males and females belong to two different populations. Consistent with the results, the data of males, females and the total group were factor analyzed separately. Table 26 presents the eigenvalues (EV) and the proportion of variance (PV) for males, females and the total group which were calculated from the unaltered correlation matrix. Two factors met the Kaiser's criterion within the three populations. Table 27 presents the eigenvalues and the percent of the common variance accounted for by the unrotated factors which were extracted by the principal component method (SPSS-PA2). The two factors extracted were rotated orthogonally by the varimax method. The loadings of each variable on the two rotated factors (F1 and F2) are presented in Table 28. The variables MA, PH, CH and BI possessed high loadings on F1, while the variables AL and AP had high loadings on F2, and variable IE had a moderate loading on the same factor. This

Table 26. The eigenvalue (EV) and proportion of variance (PV) of each factor for males, females and the total group, calculated from the unaltered correlation matrix

| | Males | | Females | | Total | |
|---|---|---|---|---|---|---|
| Factor | EV | PV | EV | PV | EV | PV |
| 1 | 2.967 | 37.1 | 3.648 | 45.6 | 3.233 | 40.4 |
| 2 | 1.280 | 16.1 | 1.054 | 13.2 | 1.157 | 14.5 |
| 3 | .988 | 12.4 | .813 | 10.2 | .904 | 11.3 |
| 4 | .744 | 9.3 | .797 | 10.0 | .724 | 9.1 |
| 5 | .623 | 7.8 | .526 | 6.6 | .624 | 7.8 |
| 6 | .554 | 6.9 | .478 | 6.0 | .543 | 6.8 |
| 7 | .442 | 5.5 | .354 | 4.4 | .418 | 5.2 |
| 8 | .402 | 5.0 | .331 | 4.1 | 3.98 | 5.0 |

Table 27. The eigenvalues (EV) and the percent of common variance (PV) accounted for by the unrotated factors extracted by the principal component method (SPSS-PA2)

| | Males | | Females | | Total | |
|---|---|---|---|---|---|---|
| Factor | EV | PV | EV | PV | EV | PV |
| 1 | 3.170 | 87.0 | 2.867 | 76.4 | 2.683 | 81.9 |
| 2 | .475 | 13.0 | .737 | 23.6 | .593 | 18.1 |

Table 28. The loadings of the eight variables on each of the two rotated factors (F1 and F2) for males, females and the total group

| | Males | | Females | | Total | |
|---|---|---|---|---|---|---|
| Variables | F1 | F2 | F1 | F2 | F1 | F2 |
| IE | .12849 | .42084 | .27763 | .32331 | .17988 | .42351 |
| AL | .33547 | .63994 | .52975 | .54607 | .41197 | .64165 |
| EL | .24303 | .32938 | .40046 | .05373 | .29872 | .28478 |
| MA | .64321 | .20212 | .69187 | .26591 | .64985 | .20315 |
| PH | .61176 | .09712 | .78587 | .19727 | .69126 | .13132 |
| CH | .63685 | .13753 | .70569 | .28579 | .68808 | .18390 |
| BI | .48255 | .40923 | .63313 | .26055 | .51230 | .39126 |
| AP | -.00002 | .75978 | .07104 | .68035 | .04350 | .65526 |

subjective classification could be explained by the existence of two factors. F1 could be an appropriate scale of scientific ability and F2 could be an appropriate scale of literate ability, while both factors could be an appropriate scale of academic ability or general academic ability because the loadings did not possess the criterion of simple structure. Each of the variables AL and BI also had a moderate loading on the factor which was not its main factor. The variable EL had moderate loadings on both factors for the total group, and its contribution on the two factors were a function of sex. The variables AL and EL could be classified as literate subjects and as scientific tools at the same time. Both languages were used inconsistently in the college of science at Yarmouk University as scientific tools and both were independent subjects of the same weight in the student high school. The need of a student or the required skills for learning any language as a goal were different from those using the language as a tool. Therefore, it is difficult to evaluate the role of each variable (AL and EL) in predicting the success of freshman in the college of science at Yarmouk University. The variables IE and AP had moderate or high loadings on F2 and small loadings on F1, thus one could offer that the two variables had approximately pure loadings on F2, a factor which was relatively not related to the criterion. The contribution of each factor was reported in Table 27. The existence of two factors, one which was more related to the criterion than the other one highlighted the existence of the national goals or objectives for education. The inference of these national objectives made the factor variables, as predictors, more appropriate than the scores of the original variables. In the words of Morris and Guertin (1977):

> . . . , factor scores are more appealing than the numerous data-level variables because of their parsimony; yet they are still based on an empirically derived conceptual system.

The traditional index of selecting the variables which belong to each factor was a factor loading of .30 or more, but due to the characteristics of the data used in this study, which are presented in Table 28, made this criterion not practical. Investigating the entries of Table 28 it was found that the variables AL, AP, MA, PH, CH and BI have the highest loadings within each sex group as well as with the total group. The first two (AL and AP) belong to factor F2 and the last four belong to factor F2. The minimum loading was .48. Using this criterion, led to a relatively simple structure for all groups, but a loading below .48 led to differential selection of variables per factor among groups and to shared variables between the two factors within groups. Hence, the criterion for a variable to load on a factor was set at .48 or more.

Two factor scores were calculated, for each student in every group, by using the factor score coefficients of the variables which belonged to each factor. The factor score coefficients are presented in Table 29. The standardized factor scores of factor F1 and factor F2 (FACZ1 and FACZ2 respectively) were used to predict the FGPA utilizing the stepwise regression procedure. The percentage of the explained variance ($R^2$ %) of FGPA from the two factor variables was compared with the percentage of the explained variance ($R_S^2$ %) of FGPA from the variables which were selected by the stepwise regression procedure and with the percentage of

Table 29. Factor score coefficients of the eight variables on each factor for males, females and the total group using the data set of the academic year 1979-1980

| Variables | Males | | Females | | Total | |
|---|---|---|---|---|---|---|
| | F1 | F2 | F1 | F2 | F1 | F2 |
| IE | -.01442 | .12086 | .00260 | .11280 | -.01843 | .15242 |
| AL | .06745 | .33073 | .05140 | .35814 | .03695 | .42514 |
| EL | .04814 | .06714 | .08974 | -.05034 | .04733 | .06854 |
| MA | .32363 | -.01359 | .20862 | .02705 | .27074 | -.03426 |
| PH | .28038 | -.07108 | .39014 | -.12443 | .33479 | -.11447 |
| CH | .31373 | -.04114 | .23434 | -.00004 | .33066 | -.07793 |
| BI | .17076 | .10041 | .19348 | .00631 | .14165 | .12217 |
| AP | -.19529 | .52805 | -.17715 | .52446 | -.14691 | .40709 |

the explained variance ($r^2$ %) of FGPA from high school GPA. This was presented in Table 30. The entries indicated that $R_S^2$ % was greater than $R^2$ % and much greater than $r^2$ % for each group, which meant that the HGPA in this situation was a weak predictor and the factor variables or the variables which were selected by the stepwise procedure should be used instead of the high school GPA. The $R^2$ % was not the optimal value because the rule used for calculating the loading on the factors was chosen to provide some practical advantages, while $R_S^2$ % was the optimal value. Therefore, the optimal performance of the factor variables and the original variables was compared. For comparative purposes, the data set of females was used. The optimal weights of factor variables were obtained by more than one rule or criterion. The traditional rule of a loading of .30 or more (rule 1) was used, but it was conditional, because the rotation of factors was positive manifold rotation but not a simple structure. It was assumed that any variable which met this criterion on both factors shared the construct with the factor of its higher loading. According to that criterion, the variables EL, MA, PH, CH and BI formed one factor, and the variables IE, AL and AP formed the other factor. The standardized factor scores were calculated and analyzed by the stepwise regression procedure. The same computations were repeated by using the factor score coefficients of all the variables for every factor (rule 2). The results of these analyses are presented in Tables 31 and 32. Table 31 presents the intercorrelations of the factor variables (FACZ1 and FACZ2) and the criterion for rule 1 (found above the diagonal of the matrix) and for rule 2 (found below the

Table 30. The percentages of explained variance $R^2$ %, $R_S^2$ % and $r^2$ % of FGPA by the factor variables, variables selected by the stepwise procedure and by the single predictor (high school GPA) respectively, for males, females and the total group

| Group | $R^2$ % | $R_S^2$ % | $r^2$ % |
|---|---|---|---|
| Males | 20 | 27 | 12 |
| Females | 64 | 68 | 59 |
| Total | 38 | 42 | 28 |

Table 31. Intercorrelations of the criterion FGPAZ and the factor variables of rule 1 (above diagonal) and rule 2 (below diagonal)

| | FGPAZ | FACZ1 | FACZ2 |
|---|---|---|---|
| FGPAZ | | .81 | .49 |
| FACZ1 | .78 | | .50 |
| FACZ2 | .41 | .24 | |

Table 32. The factor variables, their standard regression weights and the proportion of the explained variance using a loading of .03 or more to be significant (rule 1) and using a loading of any value to be significant (rule 2)

| | Variables | BETA | F | $R^2$ |
|---|---|---|---|---|
| Rule 1 | FACZ1 | .74360 | 147.6*** | .648 |
| | FACZ2 | .12322 | 4.1* | .660 |
| Rule 2 | FACZ1 | .72122 | 172.4*** | .650 |
| | FACZ2 | .23345 | 18.1* | .656 |

$^*p < .05$.

$^{***}p < .001$.

diagonal of the same matrix). Table 32 presents the factor variables, the standard regression weights and F-values for rule 1 and rule 2. The proportions of the explained criterion variance from the two factor variables, for each rule, are presented in the same table. The main contribution from applying both rules was attributed to the scores of the first factor (FACZ1) and an additional contribution was gained from the scores of factor two (FACZ2) at a .05 significance level.

The quality of predicting the FGPA by the factor variables using rule 1 and rule 2 (C1 and C2 respectively) was compared with the quality of prediction by the variables which were selected utilizing the stepwise regression procedure (C3) and with that of using the single predictor, high school GPA (C4). The percentage of the explained variance $R^2$ %, the mean absolute error (MAE), the proportion of students whose predicted scores were within one standard deviation of residuals (P), and the amount of shrinkage in $R^2$, estimated by the Stein-Darlington mathematical shrinkage formula, for each comparison are presented in Table 33. The percentage of explained variance ($R^2$ %) for factor scores (C1 and C2) and data-level variables (C3) were very similar in magnitude, but the mean absolute error (MAE) was lower for data-level variables. The proportion of students within one standard deviation of residuals (P) was lower for data-level variable. Thus, the proportion of the over predicted or underpredicted students was less for the factor scores. The shrinkage of $R^2$ was higher for data level variables which means that the regression weights of factor variables had more stability, in subsequent samples, than those of data-level variables.

Table 33. The percentage of explained variance $R^2$%, the mean absolute error (MAE), the proportion of students whose predicted scores were within one standard deviation of residuals (P) and the estimated shrinkage in $R^2$, for each of the four approaches of manipulating the data set of females (C1, C2, C3 and C4)

| Comparison | $R^2$% | MAE | P | Shrinkage of $R^2$ |
|---|---|---|---|---|
| C1 | 66.0 | 4.506 | .73 | .014 |
| C2 | 65.6 | 4.515 | .73 | .014 |
| C3 | 66.8 | 3.646 | .69 | .032 |
| C4 | 59.4 | 4.167 | .74 | .010 |

C1: The data were factor scores of the two factors assuming that a loading of .30 or more was practically significant.

C2: The data were factor scores of the two factors assuming that any loading value was practically significant.

C3: The data were the scores of the original variables analyzed by stepwise procedure.

C4: The data were the scores of the variable, High School Grade Point Average (HGPA).

The proportion of explained variance for factor scores was higher than for the single variable scores, but the MAE, P and shrinkage of $R^2$ of the factor scores and the single variable scores were approximately similar in magnitude.

Consideration of the results of applying these four indices revealed that the factor scores led the data-level variables by two to one and they leaded the single variable by one to nothing, while the data-level variables and the single variable were equal -- two to two. The overall results indicated that the quality of prediction can be improved by using factor scores. A subsequent advantage of using factor scores is that most or all of the original variables were included with their appropriate contributions to the criterion within one or more of the factors. The retention of the original variables also provided a practical advantage because the original variables helped to keep equal academic pressure upon the student involved and thereby help to fill the country's educational needs as well as helping to meet its national goals.

# CHAPTER V. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

## Summary

This study investigated the performance of four multiple regression techniques (stepwise, forward, backward and maximizing weight validity) in predicting the success of freshman students, as measured by Freshman Grade Point Average (FGPA), in the college of science at Yarmouk University, Irbid-Jordan, from the grades of five academic subjects selected from eight available subjects from the General Secondary Education Certificate Examination (GSECE). It was assumed that the Freshman Grade Point Average (FGPA) reflected the student college achievement and that the high school grades reflected the students' achievement in high school. The shrinkage in $R^2$ which was obtained empirically by the maximizing weight validity technique was compared with the corresponding values calculated from the three mathematical shrinkage formulas (Wherry, Lord-Nicholson and Stein-Darlington). The original eight variables were numerically reduced by three techniques (judgment, factor analysis and regression). A comparison of the quality of prediction from the reduced variables by the three techniques was one of the objectives of this study.

The data were collected from the records of 344 freshman (220 males and 124 females) who were available in the Admission and Registration Department of Yarmouk University. Sex was a moderator variable in this study. Therefore, the data of males and females were treated separately and the whole data set was also treated as a single group, assuming that the loss of information due to the latter treatment was negligible. The number of elements per variable was controlled in the comparison of the performance of the regression methods as well as in the estimation of shrinkage in $R^2$. These comparisons were done for three different ratios, small, intermediate and large ratios (10, 20 and 40 students per variable respectively) assuming that the gap between these ratios was large enough to be accommodated in this classification system.

The original independent variables of this study were: Islamic Education (IE), Arabic Language (AL), English Language (EL), Mathematics (MA), Physics (PH), Chemistry (CH), Biology (BI) and Arabic society and Palestine problem (AP).

This study was conducted to answer the following questions:

1. Using the collected data of the variables, all of which relevant to the criterion:
   a. Do the four regression methods result in the same variables being contained with the developed prediction equations for a fixed ratio?
   b. Do the four regression methods provide the same quality of prediction at a fixed ratio, using the frequency of superiority as an index?
   c. Is there an effect of ratio on $R^2$ and on the selected predictors, quantitatively and qualitatively, for a fixed method?
2. Can any one of the three mathematical shrinkage formulas be an alternative to the maximizing weight validity technique as an empirical approach of estimating the weight validity coefficient in subsequent samples?

3. Is there an effect of sex on the prediction equation and the quality of prediction utilizing the same method at a fixed ratio?
4. Is the prediction equation, which was selected by the same method for a fixed sex and ratio, stable over a two year period?
5. Is there any difference in the quality of the prediction of the reduced models utilizing the three techniques (judgment, factor analysis and regression), using $R^2$, Mean Absolute Error (MAE), proportion of elements whose predicted scores were within one standard deviation (p) and the shrinkage of $R^2$ as indices of the quality of prediction?

The elements for the different ratios were selected by a simple random procedure from the frame of males or females and by a proportionate stratified random procedure from the total frame which was purified of the foreign elements before selection.

The answers to the previous questions were obtained by applying the following systematic analysis:

1. The intercorrelations of all the variables and the general linear model were used in the exploratory analysis in which the variables which were related to the criterion were selected.
2. The best prediction equation utilizing each of the four methods was calculated for each of the three ratios.
3. The best prediction equation utilizing each of the four methods was calculated for males and females ignoring the effect of ratio size.

4. The shrinkage in the validity coefficient was calculated by estimating the weight validity coefficient empirically, utilizing the MWV technique, and mathematically, utilizing three shrinkage formulas.
5. The indices of the quality of prediction of the three reduced models were calculated. For this purpose, the data from the total group were factor analyzed using the principal component method (PA2), and Kaiser's criterion and varimax rotation for the subset of males and females as well as for the total group. The factor scores were calculated and the prediction equations were developed for the factor scores, the original variable scores and the composite scores.

The prediction equations over two successive academic years were calculated by the stepwise procedure to verify the validity of the prediction equation over time using a cross-sectional technique.

The regression, correlation and factor analysis subprograms of the statistical packages, Statistical Analysis System (SAS) and Statistical Package for Social Sciences (SPSS) were used. The maximizing weight validity technique had its own Fortran program designed by Morris (1977) for a Control Data Corporation (CDD) Fortran Compiler. This program was modified to fit the International Business Machine (IBM) and to minimize the cost of analysis. The weight validity, the validity coefficient of the reduced model and the prediction equation were the main outputs of this program.

An exploratory analysis led to the selection of five variables (EL, MA, PH, CH and BI) which were related to the criterion. These variables were used in comparing the performance of the stepwise, forward, backward and maximizing weight validity techniques for three different ratios. A regression equation was developed for each ratio. The frequency of superiority and the validity coefficient were calculated. The results indicated the equality of the performance of the MWV technique and the traditional techniques for samples of small ratios. The results of the analysis handled in intermediate and large sample ratios indicated the superiority of the MWV prediction equation.

The effect of sex on the performance of the four methods was investigated. The findings for each method indicated that females were consistently more predictable than males. The variables selected by the traditional and MWV methods were the same for each group of males and females, but they were different between these two groups for the same regression procedure. The MWV method was superior to the traditional methods when males and females were treated as a single group. The comparison results which were related to sex and ratio were presented in a summary table (Table 34).

One point of interest dealt with the empirical and mathematical estimation of shrinkage in $R^2$. The results revealed that the amount of shrinkage was underestimated differentially by the three mathematical shrinkage formulas for all comparisons. The highest underestimation was obtained from the Wherry formula. The best estimation for large and intermediate sample ratios was obtained from the Stein-Darlington formula.

Table 34. A summary table of the selected variables by each method for each comparison (ratio or group), frequency of superiority and the validity coefficient of each model

| Comparison | Regression method | Variables | Validity coefficient | Frequency of superiority |
|---|---|---|---|---|
| Small ratio | Traditional | PH and CH | .6675 | 107 |
| | MWV | PH and CH | .6675 | 107 |
| Intermediate ratio | Traditional | PH and CH | .6240 | 158 |
| | MWV | EL,PH and CH | .6350 | 213 |
| Large ratio | Traditional | EL,PH and CH | .6678 | 145 |
| | MWV | EL,MA,PH,CH and BI | .6675 | 196 |
| Total group | Traditional | EL, PH and CH | .6420 | 102 |
| | MWV | EL,PH,CH and BI | .6437 | 133 |
| Males | Traditional | EL,PH and CH | .5118 | 24 |
| | MWV | EL,PH and CH | .5120 | 24 |
| Females | Traditional | EL,MA,PH,CH and BI | .8170 | 19 |
| | MWV | EL,MA,PH,CH and BI | .8171 | 19 |

According to Schmitt's criterion, the shrinkage was underestimated by all three formulas in case of small sample ratio.

A stepwise procedure was used to test the validity of the prediction equation over time using the eight variables and a cross-sectional technique. The results indicated that the calculated regression equation and $R^2$ for males and females over a two year period were different. These two indices were more than enough to insure the invalidity of the prediction equation of males and females from the first year data set applied to the second year data set, while the prediction equation for the total group revealed stronger validity, but some information was lost when males and females were treated as a single group. The results of the two sets confirmed that females were more predictable than males and the performance of the prediction equation for males was not of practical significance.

The last goal in this study was to compare the quality of prediction of the reduced model for the original variables utilizing the stepwise procedure, the factor variable model and the judgmental single composite model for the subgroups of males and females as well as the total group. The variables which contributed to each factor were selected using a threshold loading of .48, which allowed every variable to contribute to only one factor. The results indicated that the reduced model utilizing the stepwise procedure performed better than the other two models for all groups, when using $R^2$% as an index of performance, but this index was not enough to determine completely the quality of prediction, and the rule of selecting the variables contributing to each factor was not optimal.

Therefore, the data set of females was used and two reduced models of factor variables were calculated, utilizing two different rules for selecting the variables which contributed to each factor. One rule was the conditional traditional rule of a .30 threshold loading such that each variable contributed to only one factor. The second rule was the rule specified by the SPSS factor analysis subprogram which considered the contribution of each variable to each factor irrespective of the loading values. Using the first rule, the variables EL, MA, PH, CH and BI formed the first factor. The variable IE, AL and AP formed the second factor. The quality of prediction was judged by four indices ($R^2$%, MAE, p and shrinkage of $R^2$). The results indicated that the performance of the factor variables under the two rules was equal. The numerical results of the four indices in Table 29 indicated that factor scores had better quality of prediction than the other two reduced models. Using the factor variables, the original variables were retained, which could be considered as a qualitative practical advantage of the factor variables, because the original variables were planned to keep equal academic pressure upon the student involved and thereby help to fill the country's educational needs and meet its national goals.

## Conclusions and Recommendations

It should be remembered that the findings in this study were based on a practical or life data set. These findings could be situation specific. It should also be kept in mind that in attempting to predict academic achievement one was not dealing with fixed laws but in probabilities. In addition, there were many factors which affected academic achievement that were not examined by this study.

Bearing these reservations in mind, the following conclusions and suggestions were drawn from the findings of this study:

1. The computation of the best prediction equation by the MWV technique becomes relatively expensive, when it is compared with the cost of the traditional methods, especially when the number of predictors is greater than five. The cost of analysis could be a minor factor in crucial situations. Therefore, the cost of analysis is not the main criterion in utilizing the MWV technique.
2. The MWV technique was not superior to the traditional methods for small sample size. Therefore, the generalization of the MWV technique over all conditions is questionable.
3. The shrinkage of $R^2$ was underestimated by the three mathematical shrinkage formulas, in case of small sample size. The Stein-Darlington formula provided the closest estimation to the empirical approach. Therefore, the known three shrinkage formulas cannot be alternatives to the empirical estimation for sample size of small ratio and the Stein-Darlington formula is recommended for other ratios.
4. One of the disadvantages of the MWV technique was that the samples which were obtained by a random splitting of the original sample were not independent. This procedure may capitalize the sampling error, especially in case of small sample size. A

comparison study of the MWV technique and Drehmer and Morris technique may provide practical and theoretical information about the two techniques. One of the two techniques could be more economical and possibly superior.

5. The quality of prediction of the MWV technique was investigated in one study by Morris utilizing multicollinear data. It was known that when extreme multicollinearity existed there was no acceptable way to perform regression analysis. The question remains: Is the MWV technique the remedy of the problems in regression analysis caused by multicollinearity? The question will require more than just a yes or no answer. This question was not investigated in this study.
6. The MWV technique added a new index for judging the quality of prediction of a regression equation. This index was the frequency of superiority of that equation in subsequent samples.
7. Ignoring some factors of the mathematical shrinkage formula, a technique which was used in some statistical packages and suggested by some authors, could lead to an indirect negative effect upon developing new precise procedures which are necessary for quantitative evaluation.
8. The differences of the findings of this study attributed to sex could be due to the interaction of social factors with academic achievement rather than to physiological factors. The question remains: To what extent do the high school grades and the freshman grade point averages reflect the achievement of the

students in the population of this study? The whole situation should be evaluated through the educational needs and the national goals of the Jordanian society.

9. Factor scores improved the quality of prediction for subsequent samples, but the stages of the analysis may cost more than the cost of analyzing the original scores by any of the regression techniques. The factor variables were more appropriate in determining which variables that measure the same ability and at the same time are relevant to the ability measured by the criterion.

10. The composite scores which were used for the selection of the students in the population of this study provided a nonpractical quality of prediction, especially for the subset of males. The prediction equations obtained by applying regression methods on the original variables or on the factor variables could be two temporary alternatives to the use of a single composite variable, but these do not pose a permanent or completely satisfactory solution.

BIBLIOGRAPHY

American Psychological Association. Standards for educational and psychological tests. Washington, D.C.: American Psychological Association, 1974.

Asher, H. B. Causal Modeling. London: Sage Publications, Inc., 1976.

Billeh, V. Y., Salah, M. and Taki, A. The prediction of academic achievement at the University of Jordan from the scholastic background through the stepwise multiple regression method. Dirasat, 1974, 1, 87-117.

Blum, M. and Naylor, J. C. Industrial Psychology: Its theoretical and social foundations. New York: Harper & Row, 1968.

Borg, W. R. and Gall, M. D. Educational Research. New York: Longman, Inc., 1979.

Bowers, J. and Loeb, J. Changes in predictor weights in GPA regression equation. Measurement and Evaluation of Guidance, 1972, 5, 438-442.

Burket, G. R. A study of reduced rank models for multiple prediction. Psychometric Monographs, 1964, 12, 1-65.

Cattell, R. B. (Ed). Handbook of multivariate experimental psychology. Chicago: Rand-McNally, 1966.

Daniel, C. and Wood, F. S. Fitting equations to data. New York: John Wiley & Sons, Inc., 1980.

Darlington, R. B. Multiple regression in psychological research and practice. Psychological Bulletin, 1968, 69, 161-182.

Daws, R. B. and Corrigan, B. Linear models in decision making. Psychological Bulletin, 1974, 81, 95-106.

Draper, N. R. and Smith, H. Applied regression analysis. New York: John Wiley & Sons, Inc., 1966.

Drehmer, D. E. and Morris, G. W. Cross-Validation with small samples: An algorithm for computing Gollob's estimator. Educational and Psychological Measurement, 1981, 41, 195-200.

Durflinger, G. W. The prediction of college success. College University Journal, 1943, 19, 68-78.

Gurtin, W. H. and Bailey, J. P. Introduction to modern factor analysis. Ann Arbor, Michigan: Edwards, 1970.

Hills, J., Klock, J., and Bush, M. The use of academic prediction equations with subsequent classes. American Educational Research Journal, 1965, 2, 203-206.

Jackson, R. A. Prediction of the academic success of college freshman. Journal of Educational Psychology, 1955, 46, 296-301.

Kerlinger, F. G. and Pedhazur, E. J. Multiple regression in behavioral research. New York: Holt, Rinehart, and Winston, 1973.

Lord, F. M. Efficiency of prediction when a regression equation from one sample is used in a new sample. Princeton, N.J.: Educational Testing Service, 1950.

Lord, F. M. and Novick, M. R. Statistical theories of mental tests. Reading, Mass.: Addison-Wesley, 1968.

Marriott, F. H. The interpretation of multiple observations. London: Academic Press, 1975.

McDonald, R. T. and Gawkoski, R. S. Predictive value of SAT scores and high school achievement for success in a college honors program. Educational and Psychological Measurement, 1979, 79, 411-420.

Morris, J. D. Selecting the best regression equation by maximizing double cross validation. Behavior Research Methods and Instrumentation, 1976, 8, 389.

Morris, J. D. Selecting the best regression equation by maximizing weight validity. Journal of Marketing Research, 1977, 14, 410-412.

Morris, J. D. and Guertin, W. H. The superiority of factor scores as predictors. Journal of Experimental Education, 1977, 45, 41-44.

Morris, J. D., Morgan, F. B., and Maynor, W. On selecting the best set of regression predictors. Journal of Experimental Education, 1980, 48, 100-103.

Mosier, C. I. Problems and designs of cross-validation. Educational and Psychological Measurement, 1951, 11, 5-11.

Nicholson, G. E. Prediction in future samples. In I. Olkin et al. (Eds.), Contributions to probability and statistics. Stanford, Calif.: Stanford University Press, 1960.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., and Bent, D. SPSS: Statistical Package for the Social Sciences (2nd ed.). New York: McGraw-Hill, Inc., 1975.

Nunnally, J. C. Psychometric theory. New York: McGraw-Hill Inc., 1978.

Perrin, D. and Whitney, D. Methods for smoothing expectancy tables applied in the prediction of success in college. Journal of Educational Measurement, 1976, 13, 223-231.

Sawyer, R. and Maxey, J. The validity of college grade prediction equations over time. Journal of Educational Measurement, 1979, 16, 279-284.

Scannell, D. P. Prediction of college success from elementary and secondary school performance. Journal of Educational Psychology, 1960, 51, 130-134.

Schmitt, N., Coyle, B. W., and Rauschenberger, J. A Monte Carlo evaluation of three formula estimates of cross-validated multiple correlation. Psychological Bulletin, 1976, 84, 751-758.

Stein, C. Multiple regression. In I. Olkin et al. (Eds.), Contributions to probability and statistics. Stanford, Calif.: Stanford University Press, 1960.

Tatsuoka, M. M. Multivariate analysis in educational research. In F. N. Kerlinger (Ed.), Review of Research in Education. Itasca, Ill.: Peacock, 1973, 273-319.

Wayne, H. H. and William, F. B. Evaluating the study habits and attitudes of high school students. Journal of Educational Psychology, 1968, 59, 404-409.

Wherry, R. J. A new formula for predicting the shrinkage of the coefficient of multiple correlation. Annals of Mathematical Statistics, 1931, 2, 440-457.

Webb, S. C. The prediction of achievement for first year dental students. Educational and Psychological Measurement, 1957, 16, 543-548.

## ACKNOWLEDGMENT

The author wishes to express his sincere gratitude and appreciation to the following:

To Dr. Anton J. Netusil, major professor, and Dr. Robert F. Strahan for their time, effort, encouragement and assistance in pursuit of this study.

To Dr. Mary E. Huba, Dr. Frederick G. Brown and Dr. Rex A. Thomas for their valuable counseling, advice and for serving on the committee.

To Mr. Bud Meador for his assistance in modifying the computer program.

To the staff of the Admission and Registration Department at Yarmouk University, Irbid-Jordan, for their help in data collection.

To my wife, Fatmeh, and to my children, Audeh, Camilia, Suleiman, Tamara, Mohammad and Asma for their patience and understanding.

To my parents and my brothers for their support to my family.

# VITA

Ahmad S. Audeh was born on December 15, 1948 and raised in Sakib, a rural town in Jordan. He received an elementary education from Sakib, a preparatory education from Alkita, a secondary education from Irbid High School graduating in 1967.

In October 1967, Ahmad entered the University of Jordan. The degree of Bachelor of Science in physics was awarded to him in July 1971. Ahmad was a teacher of physics in Jerash senior high schools from August 1971 to August 1979. While he was working as a teacher, he entered the College of Education at the University of Jordan in October 1973. A Diploma of Education was awarded to him in August 1975. Ten months later he entered the same college and received a Master's degree in Educational Psychology in August 1978.

From September 1979 to August 1980, Ahmad was a research assistant in the Education Department of Yarmouk University, Irbid-Jordan.

In August 1980, Ahmad entered the Research and Evaluation section of the Professional Studies Department at Iowa State University in the United States of America and worked toward the Doctor of Philosophy in Research and Evaluation which was awarded to him on July 24, 1982.

On a hierarchal scale from one to ten, Ahmad is in the bottom of this hierarchy looking toward the top. The top itself is not the goal from his point of view. He believes that the top is far removed but he believes at the same time that it can be reached.

# APPENDIX: A COMPUTER PROGRAM OF MAXIMIZING WEIGHT VALIDITY MULTIPLE REGRESSION TECHNIQUE

```
C      IMPLICIT REAL*8(A-H,O-Z)
C PROGRAM:  MULTIPLE REGRESSION VARIABLE SELECTION MAXIMIZING
C  DOUBLE CROSS-VALIDATION REPLICATION CORRELATION - 628
C  RECORDS.
C PROGRAMMER:  JOHN D. MORRIS, GEORGIA SOUTHERN COLLEGE, 1975.
C MORE COMPLETE DESCRIPTION IN:
C MORRIS, JOHN D.  SELECTING THE BEST REGRESSION EQUATION BY
C     MAXIMIZING DOUBLE CROSS-VALIDATION CORRELATION.  BEHAVIOR
C     RESEARCH METHODS AND INSTRUMENTATION, 1976, 8, 389.
C AND,
C MORRIS, JOHN D.  STREP:  SELECTING THE BEST REGRESSION
C     EQUATION BY MAXIMIZING WEIGHT VALIDITY.  JOURNAL OF
C     MARKETING RESEARCH, 1977, 14, 410-412.
C DATA REQUIRED:
C 1) TITLE CARD - ANY CHARACTERS IN ANY COLUMNS - TIT(20).
C 2) VARIABLE FORMAT FOR READING SCORE VECTORS(DEP VAR LAST).
C 3) CONTROL CARD -
C     COL 1-5 NUMBER OF SUBJECTS INPUT - N.
C     COL 6-10 NUMBER OF INDEPENDENT VARIABLES (MAX=10)-NI.
C     COL 11-15 EITHER NUMBER OF ITERATIONS OR PROBABILITY
C      DESIRED FOR TEST TO STOP ITERATIONS (PUNCH DECIMAL, F5.4)
C      (ASSUMED .05 IF BLANK).
C     COL 16-20 1 IF ITERATION TO PROCEED UNTIL "BEST" EQUATION
C      IDENTIFIED, OTHERWISE THE 'BEST' GROUP OF EQUATIONS WILL
C      BE IDENTIFIED - ITERC.
C     COL 21-25 1 IF THIS IS THE LAST JOB - IJOB.
C     COL 26-35 UP TO NI - 1 FORCED VARIABLE INDICES (LEFT
C      JUSTIFIED) - IFOR.
C 4) EIGHT DIGIT RANDOM NUMBER FOR SEED.
C 5) SUBJECTS' SCORE VECTORS ACCORDING TO # 2.
C*SUBPROGRAMS RGRESS, IVERSE, MATOUT, VECOUT, PRBF, AND
C RANDOM REQUIRED.
C***SCRATCH UNIT 1 REQUIRED FOR SUBJECT STORAGE.
C*INCLUDE 'PROGRAM' STMT ONLY IF CDC FORTRAN COMPILER IS USED,
C THE INPUT FILE NAME IS INFILE.
      DIMENSION X(06),R(06,06,2),XB(06,2),SD(06,2),
     +BR(06),BN(06),RT(06,06),A(06,06),C(06,06),RI(06,06),
     +TIT(20),FND2(2),IVAR(06),XBT(06),SDT(06),S(06),RA(06,06),
     +SUM(0031,3),XBA(06),SDA(06),FORM(20),
     +REP(0031,2),IFOR(9)
      INTEGER*2 MC(0031,06)
      NDEM = 06
   90 READ(5,1) (TIT(I),I=1,20),(FORM(I),I=1,20),N,NI,PROB,
     +ITERC,IJOB,(IFOR(I), I = 1,9)
      REWIND 1
    1 FORMAT(20A4/20A4/2I5,F5.4,2I5,9I1)
C CALCULATE NUMBER OF FORCED VARIABLES.
      NFO = 1
   92 IF (IFOR(NFO) .EQ. 0) GO TO 93
```

```
      NFO = NFO + 1
      IF (NFO .LT. NI) GO TO 92
   93 NFO = NFO - 1
C DECIDE IF PROB IS NUMBER OF ITERATIONS OR STOPPING
C PROBABILITY.
      IPROB = 0
      IF (PROB .LE. 0.0) PROB = .05
      IF (PROB .GE. 1.) IPROB = PROB
      WRITE(6,2) (TIT(I), I = 1,20),(FORM(I), I = 1,20),N,NI
     +,PROB,ITERC
      IF (NFO .GT. 0) WRITE(6,97) (IFOR(I), I = 1,NFO)
   97 FORMAT(19H FORCED VARIABLES =,9I3)
    2 FORMAT(1H1,36X,
     +41H***MULTIPLE REGRESSION VARIABLE SELECTION,
     +14H MAXIMIZING***/34X,
     +38H***DOUBLE CROSS-VALIDATION REPLICATION,
     +15H CORRELATION***/40X,
     +36H***JOHN D. MORRIS - GEORGIA SOUTHERN,
     +17H COLLEGE, 1975***/1H0,20A4/10H FORMAT = ,20A4/
     +11H SUBJECTS =,
     +I5/24H INDEPENDENT VARIABLES =,I5/
     +28H PROBABILITY LEVEL ENTERED =,F8.3/8H ITERC =,I2)
      FN = N
      I = N/2
      IF (I .GT. NI) GO TO 76
      WRITE(6,77)
   77 FORMAT(39H**TOO FEW SUBS/VAR FOR SPLITS - LINEAR ,
     +19HDEPENDENCIES ARISE.)
      STOP
   76 FND = I
      ND = NI + 1
      NIT = 0
      SL = RANDOM(1)
      NP = 2**NI - 1
      IF (NFO .GT. 0) NP = 2**(NI - NFO)
      DO 3 J = 1,ND
      XBA(J) = 0.
      DO 3 K = J,ND
    3 RA(J,K) = 0.
      DO 62 I = 1,NP
      REP(I,1) = 0.
   62 REP(I,2) = 0.
C READ IN SUBJECTS AND ACCUMULATE.
      DO 60 M = 1,N
      READ(5,FORM) (X(I), I = 1,ND)
      WRITE(1)   X
      DO 60 J = 1,ND
      XBA(J) = XBA(J) + X(J)
      DO 60 K = J,ND
   60 RA(J,K) = RA(J,K) + X(J)*X(K)
```

```
      DO 61 I = 1,ND
      XBA(I) = XBA(I)/FN
   61 SDA(I) =  SQRT(RA(I,I)/FN - XBA(I)*XBA(I))
      DO 33 I = 1,ND
      DO 33 K = I,ND
      RA(I,K) = (RA(I,K)/FN-XBA(I)*XBA(K))/(SDA(I)*SDA(K))
   33 RA(K,I) = RA(I,K)
      WRITE(6,43)
   43 FORMAT(//13H MEAN VECTOR:)
      CALL VECOUT(XBA,ND,NDEM)
      WRITE(6,46)
   46 FORMAT(//27H STANDARD DEVIATION VECTOR:)
      CALL VECOUT(SDA,ND,NDEM)
      WRITE(6,47)
   47 FORMAT(//20H CORRELATION MATRIX:)
      CALL MATOUT(RA,IVAR,ND,ND,NDEM,NDEM)
C CALCULATE BINARY EQUIVALENTS AND VARIABLE INDICES.
      IS = NI
      NF = 2**NI - 1
      K = 1
      DO 22 IP = 1,NF
      IF ((IP - 2**(NI - IS)) .GT. 0) IS = IS - 1
      M = 0
      J = IP
      DO 27 I = IS,NI
      JP = 2**(NI - I)
      IB = J/JP
      IF (IB .NE. 1) GO TO 27
      M = M + 1
      MC(K,M) = I
      J = J - JP
      IF (J .EQ. 0) GO TO 99
   27 CONTINUE
   99 MC(K,NDEM) = M
      IF (NFO .EQ. 0) GO TO 94
      I = 1
   96 DO 95 J = 1,M
      IF (MC(K,J) .EQ. IFOR(I)) GO TO 98
   95 CONTINUE
      GO TO 22
   98 I = I + 1
      IF (I .LE. NFO) GO TO 96
   94 K = K + 1
   22 CONTINUE
C BEGIN AN ITERATION.
      JP = 0
   63 DO 7 I = 1,2
      DO 7 J = 1,ND
      XB(J,I) = 0.
      DO 7 K = J,ND
```

```
    7 R(J,K,I) = 0.
C READ IN SUBJECTS AND RANDOMLY SPLIT.
      REWIND 1
      FND2(1) = 0.
      FND2(2) = 0.
      DO 8 M = 1,N
      READ(1)      X
      IF (FND2(1) .LT. FND) GO TO 106
      I = 2
      GO TO 12
  106 IF (FND2(2) .LT. FND) GO TO 11
      I = 1
      GO TO 12
   11 I = 2*RANDOM(0) + 1
   12 FND2(I) = FND2(I) + 1.
      DO 8 J = 1,ND
      XB(J,I) = XB(J,I) + X(J)
      DO 8 K = J,ND
    8 R(J,K,I) = R(J,K,I) + X(J)*X(K)
      DO 10 L = 1,2
      DO 9 I = 1,ND
      XB(I,L) = XB(I,L)/FND2(L)
    9 SD(I,L) =  SQRT(R(I,I,L)/FND2(L) - XB(I,L)*XB(I,L))
      DO 10 I = 1,ND
      DO 10 K = I,ND
      R(I,K,L) = (R(I,K,L)/FND2(L)-XB(I,L)*XB(K,L))/
     +    (SD(I,L)*SD(K,L))
   10 R(K,I,L) = R(I,K,L)
C CALCULATE REGRESSION EQUATIONS AND WEIGHT VALIDITIES.
      FP = XB(ND,1)*FND2(1) + XB(ND,2)*FND2(2)
      SLP = (SD(ND,1)*SD(ND,1)+XB(ND,1)*XB(ND,1))*FND2(1)+
     +(SD(ND,2)*SD(ND,2)+XB(ND,2)*XB(ND,2))*FND2(2)
      SL = -2.
      DO 105 IP = 1,NP
      DO 14 I = 1,3
   14 SUM(IP,I) = 0.
      DO 13 L = 1,2
      IF (L .EQ. 1) II = 2
      IF (L .EQ. 2) II = 1
      DO 15 I = 1,ND
      XBT(I) = XB(I,L)
      SDT(I) = SD(I,L)
      DO 15 J = 1,ND
   15 RT(I,J) = R(I,J,L)
      M = MC(IP,NDEM)
      DO 69 I = 1,M
   69 IVAR(I) = MC(IP,I)
      CALL RGRESS(RT,RI,A,C,SDT,XBT,S,NDEM,ND,ND,M,IVAR,
     +BR,BN,YI,R2,IB)
      IF (IB .NE. 1) GO TO 26
```

```
      JP = JP + 1
      IF (JP .LT. 10) GO TO 63
      WRITE(6,77)
      STOP
   26 IB = M-1
      F = 0.
      R2 = 0.
      IF (IB .EQ. 0) GO TO 100
      DO 21 I = 1,IB
      IV = IVAR(I)
      F = F + BR(I)*R(ND,IV,II)*SD(IV,II)*SD(ND,II)
      YI = YI + BR(I)*XB(IV,II)
      R2 = R2 + BR(I)*BR(I)*SD(IV,II)*SD(IV,II)
      IP1 = I + 1
      DO 21 J = IP1,M
      JV = IVAR(J)
   21 R2 = R2 + 2.*BR(I)*BR(J)*R(IV,JV,II)*SD(IV,II)*SD(JV,II)
  100 IV = IVAR(M)
      R2 = R2 + BR(M) * BR(M)*SD(IV,II)*SD(IV,II)
      F = F + BR(M)*R(ND,IV,II)*SD(IV,II)*SD(ND,II)
      YI = YI + BR(M)*XB(IV,II)
      SUM(IP,1) = YI*FND2(II) + SUM(IP,1)
      SUM(IP,2) = (R2 + YI*YI)*FND2(II) + SUM(IP,2)
   13 SUM(IP,3) = (F + YI*XB(ND,II))*FND2(II) +SUM(IP,3)
      YI = FN*SUM(IP,2)-SUM(IP,1)*SUM(IP,1)
      IF (YI .NE. 0.) GO TO 128
      YI = 0.
      GO TO 129
  128 YI=(FN*SUM(IP,3)-FP*SUM(IP,1))/ SQRT(YI*(FN*SLP-FP*FP))
  129 REP(IP,1) = REP(IP,1) + .5*ALOG((1. + YI)/(1. - YI))
      IF (YI .LE. SL) GO TO 105
      SL = YI
      K = IP
  105 CONTINUE
      REP(K,2) = REP(K,2) + 1.
      NIT = NIT + 1
      IF (PROB .LT. 1.) GO TO 64
      IF (NIT - IPROB) 63,83,83
   64 IF (NIT .LT. 5) GO TO 63
C CHECK FOR STOP OF ITERATIONS.
C  SELECT HIGH AND LOW FREQUENCY AND CHECK FOR DIFF.
      R2 = REP(1,2)
      F = R2
      M = 1
      DO 65 IP = 2,NP
      IF (REP(IP,2) .LE. R2) GO TO 87
      R2 = REP(IP,2)
      M = IP
      GO TO 65
   87 IF (REP(IP,2) .GE. F) GO TO 65
```

```
      F = REP(IP,2)
   65 CONTINUE
      IP = F
      L = R2
      IF (STEST(L,IP) .GE. PROB) GO TO 63
      IF (ITERC .EQ. 0) GO TO 89
C SELECT 2CND BEST EQUATION.
      I = L
      DO 66 J = 1,NP
      IF (J .EQ. M) GO TO 66
      IF ((L - REP(J,2)) .LT. I) I = L - REP(J,2)
   66 CONTINUE
C CALCULATE BINOMIAL SIGN TEST THAT BEST EQUATION IS
C  SIGNIFICANTLY BETTER THAN SECOND BEST.
      I = L - I
      SL = STEST(L,I)
      IF (SL .GE. PROB) GO TO 63
      IB = L - 1
      GO TO 83
C SELECT BEST EQUATION INFERIOR TO MTH.
   89 IP = L - IP
      DO 80 I = 1,NP
      IF (I .EQ. M) GO TO 80
      IB = REP(I,2)
      IF (STEST(L,IB) .GE. PROB) GO TO 80
      IB = L - IB
      IF (IB .LT. IP) IP = IB
   80 CONTINUE
      IB = L - IP
C CHECK SMALLEST EQUATION SUPERIOR TO IB.
      IP = L
      DO 101 I = 1,NP
      J = REP(I,2)
      IF ((J .GT. IB) .AND. (J .LT. IP)) IP = J
  101 CONTINUE
      IF (STEST(IP,IB) .GE. PROB) GO TO 63
   83 WRITE(6,35)
   35 FORMAT(10H1VARIABLES,11X,18HMEAN REPLICATION R,
     +25H        PROB   TIMES HIGHEST)
      PROB = FN - 2.
      R2 = NIT
      DO 68 IP = 1,NP
      YI =  EXP(2.*REP(IP,1)/R2)
      YI = (YI - 1.)/(1. + YI)
      F = YI*YI*(PROB/(1. - YI*YI))
      F = PRBF(1.,PROB,F)
      II = MC(IP,NDEM)
   68 WRITE(6,37) YI,F,REP(IP,2),(MC(IP,I), I = 1,II)
   37 FORMAT(1H ,31X,F7.4,3X,F7.4,10X,F5.0,T1,10I2)
      IF (IPROB .GE. 1) GO TO 91
```

```
      WRITE(6,84)
   84 FORMAT(18H1BEST EQUATION(S):)
      DO 85 I = 1,NP
      IF (REP(I,2) .LE. IB) GO TO 85
      II = MC(I,NDEM)
      DO 86 J = 1,II
   86 IVAR(J) = MC(I,J)
      WRITE(6,51) (IVAR(J), J = 1,II)
   51 FORMAT(////12H VARIABLES =,10I3)
      CALL RGRESS(RA,RI,A,C,SDA,XBA,S,NDEM,ND,ND,II,IVAR,
     +BR,BN,YI,SL,J)
      IF (J .NE. 1) GO TO 49
      WRITE(6,16)
   16 FORMAT(36H0***SINGULAR R - DISREGARD ANALYSIS.)
      GO TO 85
   49 SL =  SQRT(SL)
      WRITE(6,50) SL
   50 FORMAT(23H MULTIPLE CORRELATION =,F7.4//11H RAW SCORE ,
     +19HREGRESSION WEIGHTS:)
      CALL VECOUT(BR,II,NDEM)
      WRITE(6,52) YI
   52 FORMAT(11H CONSTANT =,F12.4//24H STANDARDIZED REGRESSION,
     +9H WEIGHTS:)
      CALL VECOUT(BN,II,NDEM)
      WRITE(6,88)
   88 FORMAT(1H ,24(5H*****))
   85 CONTINUE
   91 IF (IJOB .NE. 1) GO TO 90
      STOP
      END
      SUBROUTINE RGRESS(X,XI,B,XC,SD,XB,C,ND,NV,IC,NP,IPR,BR,
     +BN,YI,R2,
     +ISING)
C     IMPLICIT REAL*8(A-H,O-Z)
C THIS SUBROUTINE CALCULATES A REGRESSION EQUATION FROM AN
C  INPUT
C CORRELATION MATRIX FOR AN INDEXED CRITERION FOR ANY SUBSET OF
C THE INPUT VARIABLES.
C PROGRAMMER:  JOHN D. MORRIS - UNIVERSITY OF FLORIDA, 1974.
C    X - INPUT CORRELATION MATRIX CONTAINING CRITERION.
C    XI - RETURNED INVERSE OF PREDICTOR CORRELATION MATRIX.
C    B - STORAGE MATRIX.
C    XC - STORAGE MATRIX.
C    SD - INPUT STANDARD DEVIATIONS.
C    XB - INPUT MEANS.
C    C - STORAGE VECTOR.
C    ND - DIMENSION FOR ALL ARRAYS IN CALLING PROGRAM.
C    NV - TOTAL NUMBERS OF VARIABLES IN X INCLUDING CRITERION.
C    IC - INDEX OF THE CRITERION VARIABLE IN X.
C    NP - NUMBER OF PREDICTOR VARIABLES.
```

```
C     IPR - PREDICTOR INDICES IN NUMERICAL ORDER.
C     BR - VECTOR OUTPUT HOLDING RAW SCORE REGRESSION WEIGHTS.
C     BN - VECTOR OUTPUT HOLDING NORMALIZED REGRESSION WEIGHTS.
C     YI - CRITERION INTERCEPT.
C     R2 - MULTIPLE CORRELATION COEFFICIENT SQUARED.
C     ISING - SINGNAL FOR A SINGULAR CORRELATION MATRIX.
C*SUBROUTINES REQUIRED: IVERSE.
      DIMENSION X(ND,ND),XI(ND,ND),B(ND,ND),C(ND),BR(ND),
     +BN(ND),SD(ND)
     +,XC(ND,ND),XB(ND),IPR(ND)
C BRANCH IF MORE THAN ONE PREDICTOR.
      ISING = 0
      IF (NP .GT. 1) GO TO 18
      I = IPR(1)
      BN(1) = X(IC,I)
      R2 = BN(1)*BN(1)
      BR(1) = SD(IC)*BN(1)/SD(I)
      YI = XB(IC) - BR(1)*XB(I)
      RETURN
C STORE CRITERION AND COLLAPSE TO PREDICTORS.
   18 DO 1 I = 1,NV
    1 C(I) = X(IC,I)
      IP1 = 0
      DO 2 I = 1,NV
      IJ = 0
      DO 14 J = 1,NP
      IF (I .EQ. IPR(J)) IJ = 1
   14 CONTINUE
      IF (IJ .NE. 1) GO TO 2
      IP1 = IP1 + 1
      DO 12 J = 1,NV
   12 XC(IP1,J) = X(I,J)
    2 CONTINUE
      IP1 = 0
      DO 3 I = 1,NV
      IJ = 0
      DO 15 J = 1,NP
      IF (I .EQ. IPR(J)) IJ = 1
   15 CONTINUE
      IF (IJ .NE. 1) GO TO 3
      IP1 = IP1 + 1
      DO 13 J = 1,NP
   13 XC(J,IP1) = XC(J,I)
    3 CONTINUE
C COMPUTE IVERSE OF PREDICTOR R MATRIX - BRANCH IF SINGULAR.
      CALL IVERSE(XC,XI,B,NP,ND,R2,ISING)
      IF (ISING .EQ. 1) GO TO 4
C CALCULATE WEIGHTS, MULTIPLE RSQ, AND CRITERION INTERCEPT.
      YI = 0.
      R2 = 0.
```

```
      IP1 = 0
      DO 6 I = 1,NV
      IJ = 0
      DO 16 J = 1,NP
      IF (I .EQ. IPR(J)) IJ = 1
   16 CONTINUE
      IF (IJ .NE. 1) GO TO 6
      IP1 = IP1 + 1
      BN(IP1) = 0.
      IL1 = 0
      DO 5 J = 1,NV
      IJ = 0
      DO 17 K = 1,NP
      IF (J .EQ. IPR(K)) IJ = 1
   17 CONTINUE
      IF (IJ .NE. 1) GO TO 5
      IL1 =IL1 + 1
      BN(IP1) = XI(IP1,IL1)*C(J) + BN(IP1)
    5 CONTINUE
      BR(IP1) = BN(IP1)*SD(IC)/SD(I)
      R2 = BN(IP1)*C(I) + R2
      YI = XB(I)*BR(IP1) + YI
    6 CONTINUE
      YI = XB(IC) - YI
    4 RETURN
      END
      SUBROUTINE IVERSE(A,AI,B,NV,ND,D,ISING)
C PROGRAMMER:  JOHN D. MORRIS - UNIVERSITY OF FLORIDA, 1974.
C     IMPLICIT REAL*8(A-H,O-Z)
C THIS SUBROUTINE CALCULATES THE INVERSE OF A SQUARE INPUT
C MATRIX BY THE METHOD AS OUTLINED BY TATSUOKA WITH ROW
C INTERCHANGE TO PUT THE LARGEST ELEMENT ON THE DIAGONAL.
C (MULTIVARIATE ANALYSIS - 1971, 253-60).  A CHECK IS MADE
C ON THE INVERSE AND IF SINGULAR ISING IS SET TO 1.
C ARGUMENTS ARE:
C     A - SQUARE INPUT MATRIX, NOT DESTROYED IN PROCESS.
C     AI - STORAGE MATRIX, OUTPUT AS A-INVERSE.
C     B  - STORAGE MATRIX.
C     NV - ORDER OF A
C     ND - DIMENSION FOR A, B & C IN CALLING PROGRAM.
C     D - RETURNED AS THE DETERMINANT OF A.
C     ISING - SINGNAL FOR SINGULAR MATRIX.
      DIMENSION A(ND,ND),AI(ND,ND),B(ND,ND)
      ISING = 0
      D = 1.
C INITIALIZE STORAGE MATRICES.
      DO 6 I = 1,NV
      DO 6 J = 1,NV
      B(I,J) = A(I,J)
      AI(I,J) = 0.
```

```
    6 IF (I .EQ. J) AI(I,J) = 1.
      DO 1 IL = 1,NV
      IF (ISING .EQ. 1) GO TO 1
C FIND LARGEST PIVOT ELEMENT IN COLUMN.
      KMAX = IL
      K = IL + 1
      TEMP1 = B(IL,IL)
      IF (IL .EQ. NV) GO TO 3
    5 IF ( ABS(TEMP1) .LT.  ABS(B(K,IL))) GO TO 2
      K = K + 1
      IF (K .LE. NV) GO TO 5
      GO TO 3
    2 TEMP1 = B(K,IL)
      KMAX = K
      K = K + 1
      IF (K .LE. NV) GO TO 5
    3 D = D*B(KMAX,IL)
      IF (D .EQ. 0.) ISING = 1
      IF (ISING .EQ. 1) GO TO 1
      IF (IL .EQ. NV) GO TO 7
C ENTERCHANGE ROWS.
      IF (KMAX .EQ. IL) GO TO 7
      D = -D
      DO 8 J = 1,NV
      TEMP1 = B(IL,J)
      TEMP2 = B(KMAX,J)
      B(KMAX,J) = TEMP1
      B(IL,J) = TEMP2
      TEMP1 = AI(IL,J)
      TEMP2 = AI(KMAX,J)
      AI(IL,J) = TEMP2
    8 AI(KMAX,J) = TEMP1
C DIVIDE PIVOT ROW BY PIVOT.
    7 TEMP1 = B(IL,IL)
      DO 9 I = 1,NV
      B(IL,I) = B(IL,I)/TEMP1
    9 AI(IL,I) = AI(IL,I)/TEMP1
      IF (IL .EQ. NV) GO TO 1
C SUBTRACT MULTIPLES OF THE PIVOT ROW.
      K = IL + 1
      DO 10 I = K,NV
      TEMP1 = B(I,IL)
      DO 10 IA = 1,NV
      B(I,IA) = B(I,IA) - TEMP1*B(IL,IA)
   10 AI(I,IA) = AI(I,IA) - TEMP1*AI(IL,IA)
    1 CONTINUE
      IF (ISING .EQ. 1) GO TO 13
C ACCOMPLISH BACKWARD SOLUTION.
      NVL1 = NV - 1
      DO 11 I = 1,NVL1
```

```
      IA = NV - I
      IAP1 = IA + 1
      DO 11 J = IAP1,NV
      TEMP1 = B(IA,J)
      DO 11 K = 1,NV
      B(IA,K) = B(IA,K) - TEMP1*B(J,K)
   11 AI(IA,K) = AI(IA,K) - TEMP1*AI(J,K)
C CHECK INVERSE.
      DO 14 I = 1,NV
      DO 14 J = 1,NV
      B(I,J) = 0.
      DO 15 IA = 1,NV
   15 B(I,J) = A(I,IA)*AI(IA,J) + B(I,J)
      IF ((I .EQ. J).AND.( ABS(B(I,J)-1.).GT..01)) ISING = 1
   14 IF ((I .NE. J).AND.( ABS(B(I,J)).GT..01)) ISING =1
   13 RETURN
      END
      SUBROUTINE MATOUT(A,N,NR,NC,NDR,NDC)
C     IMPLICIT REAL*8(A-H,O-Z)
C THIS SUBROUTINE OUTPUTS A RECTANGULAR MATRIX.
C PROGRAMMER:  JOHN D. MORRIS - UNIVERSITY OF FLORIDA, 1974.
C    A - MATRIX TO BE OUTPUT.
C    N - STORAGE VECTOR DIMENSIONED NDC.
C    NR - NUMBER OF ROWS IN A.
C    NC - NUMBER OF COLUMNS IN A.
C    NDR - DIMENSION FOR ROWS IN THE CALLING PROGRAM.
C    NDC - DIMENSION FOR COLUMNS IN THE CALLING PROGRAM.
      DIMENSION A(NDR,NDC),N(NDC)
      DO 1 I = 1,NC
    1 N(I) = I
      ICUM = 0
      LI = 0
    6 ICUM = ICUM + 10
      IP = LI + 1
      LI = ICUM
      IF (ICUM .GE. NC) LI = NC
      WRITE(6,2) (N(I), I = IP,LI)
    2 FORMAT(1H ,5X,10I10)
      DO 4 I = 1,NR
    4 WRITE(6,5) I,(A(I,J), J = IP,LI)
    5 FORMAT(1H ,I5,10F10.4)
      IF (LI .LT. NC) GO TO 6
      RETURN
      END
      SUBROUTINE VECOUT(X,NE,ND)
C     IMPLICIT REAL*8(A-H,O-Z)
C THIS SUBROUTINE OUTPUTS A VECTOR.
C PROGRAMMER:  JOHN D. MORRIS - UNIVERSITY OF FLORIDA, 1974.
C    X - VECTOR TO BE OUTPUT.
C    NE - NUMBER OF ELEMENTS IN THE VECTOR.
```

```
C     ND - DIMENSION FOR X IN THE CALLING PROGRAM.
      DIMENSION X(ND)
      WRITE(6,2) (I,X(I), I = 1,NE)
    2 FORMAT(6(I6,2H) ,F10.4))
      RETURN
      END
      FUNCTION PRBF(DA,DB,FR)
C     IMPLICIT REAL*8(A-H,O-Z)
C FROM D. J. VELDMAN LIBRARY.
      PRBF = 1.
      IF (DA*DB*FR .EQ. 0.) RETURN
      IF (FR .LT. 1.) GO TO 5
      A = DA
      B = DB
      F = FR
      GO TO 10
    5 A = DB
      B =DA
      F = 1./FR
   10 AA = 2./(9.*A)
      BB = 2./(9.*B)
      Z =  ABS(((1. - BB)*F**.333333 - 1. + AA)/ SQRT(BB*F
     +**.666667 + AA))
      IF (B .LT. 4.) Z = Z*(1. + .08*Z**4./B**3.)
      PRBF = .5/(1. + Z*(.196854 + Z*(.115194 + Z*
     +(.000344 + Z*.019527))))**4.
      IF (FR .LT. 1.) PRBF = 1. - PRBF
      RETURN
      END
      FUNCTION RANDOM(K)
C     IMPLICIT REAL*8(A-H,O-Z)
C THIS FUNCTION CREATES A UNIFORMLY DISTRIBUTED VARIATE-
C FROM COOLEY AND LOHNES LIBRARY.
      IF (K) 1,2,1
    1 READ(5,3) K1,K2,K3,K4
      DK=100*(100*(100*K1+K2)+K3)+K4
    3 FORMAT(4I2)
      WRITE(6,4) K1,K2,K3,K4
    4 FORMAT(15H RANDOM SEED = ,4I2)
C   2 M1 = 11*K4
C     M2 = 11*K3
C     M3 = 11*K2 + K4
C     M4 = 11*K1 + K3
C     J = M1/100
C     K4 = M1 - 100*J
C     M2 = M2 + J
C     J = M2/100
C     K3 = M2 - 100*J
C     M3 = M3 + J
C     J = M3/100
```

```
C       K2 = M3 - 100*J
C       M4 = M4 + J
C       J = M4/100
C       K1 = M4 - 100*J
C       X1 = K1
C       X2 = K2
C       RANDOM = X1*1.E-2 + X2*1.E-4
      2 RANDOM=GGUBFS(DK)
        RETURN
        END
        FUNCTION STEST(NP,NM)
C       IMPLICIT REAL*8(A-H,O-Z)
C THIS PROGRAM WILL CALCULATE THE PROBABILITY THAT PLUSES OCCUR
C THE INPUT NUMBER OF TIMES (NP) OR MORE UNDER THE ASSUMPTION
C THAT PLUSES AND MINUSES (NM) ARE EQUIPROBABLE USING THE
C BINOMIAL EXPANSION IF NP + NM .LE. 20, AND A NORMAL
C    APPROXIMA-
C TION OTHERWISE.
C THIS IS TO BE USED WITH A CORRELATED SAMPLE SIGN TEST.
C*FUNCTION PRBF REQUIRED.
        N = NP + NM
        IF (N .GT. 20) GO TO 1
        Y = .5**N
        FACTN = 1.
        IF (N .LE. 1) GO TO 5
        DO 2 I = 2,N
      2 FACTN = FACTN*I
      5 STEST = 0.
        DO 8 J = NP,N
        R = 1.
        IF (J .LE. 1) GO TO 6
        DO 3 I = 2,J
      3 R = R*I
      6 X = FACTN/R
        R = 1.
        K = N - J
        IF (K .LE. 1) GO TO 7
        DO 4 I = 2,K
      4 R = R*I
      7 X = X/R
        X = X*Y
      8 STEST = STEST + X
        RETURN
      1 R = N
        STEST = IABS(NP -NM)
        STEST = (STEST - 1.)*(STEST - 1.)/R
        STEST = PRBF(1.,1000.,STEST)
        RETURN
        END
```