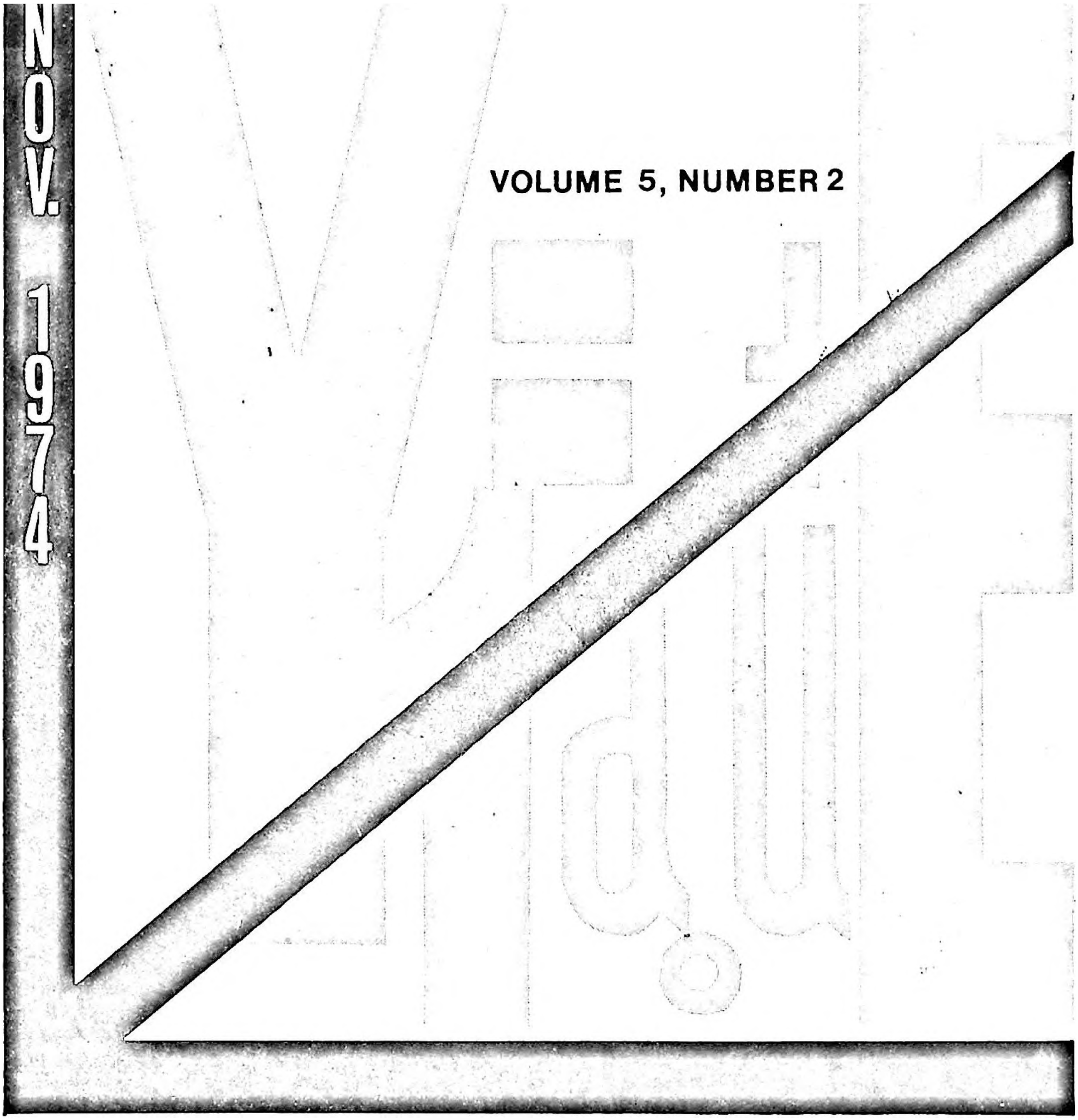


VOLUME 5, NUMBER 2

MULTIPLE LINEAR REGRESSION VIEWPOINTS

A publication of the Special Interest Group
on Multiple Linear Regression
of The American Educational Research Association.

IF YOU ARE SUBMITTING A RESEARCH ATRICLE OTHER THAN NOTES OR COMMENTS, I WOULD LIKE TO SUGGEST THAT YOU USE THE FOLLOWING FORMAT, AS MUCH AS POSSIBLE:

TITLE
AUTHOR AND AFFILIATION
SINGLE-SPACED INDENTED ABSTRACT (THE REST OF THE MANUSCRIPT SHOULD BE DOUBLE-SPACED)
INTRODUCTION (PURPOSE—SHORT REVIEW OF LITERATURE, ETC.)
METHOD
RESULTS
DISCUSSION (CONCLUSION)
REFERENCES

ALL MANUSCRIPTS SHOULD BE SENT TO THE EDITOR AT THE ABOVE ADDRESS. (ALL MANUSCRIPTS SHOULD BE CAMERA-READY COPY)

TABLE OF CONTENTS

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 5, No. 2, 1974

CAUSALITY AND PREDICTION: SIMILAR BUT NOT SYNONYMOUS

by

Joseph F. Maola, Assistant Professor
School of Education
Duquesne University
Pittsburgh, Pennsylvania

A recent issue of Viewpoints (June, 1974) published an article that began with the following sentence: "A universal objective of scientists involved with explanations of behavior or phenomenon is to demonstrate their knowledge of what is causing it" (Byrne). Upon reading this sentence the initial responses of typical behavioral researchers may probably have been the following: (1) What a true statement! and (2) The article is going to show me how to explain causality! Only upon completion of the paper, however, does the reader realize that the author of the paper, while extremely knowledgable and versed in the power and strength of multiple regression analysis, had fallen victim to that antagonist of all social scientific reporters; semantics.

Byrne (1974) reports that causality exists upon three levels, however, of the three levels stated: (1) the real cause, (2) the hypothesized cause, and (3) the statistical cause, only the first can truly be considered a <u>cause.</u> What is really causing a phenomena can never be answered through guessing or inferential statistics; likewise, "The Use of Multiple Regression Equations to Demonstrate Causality" is an erroneous title which may appropriately be retitled through substitution of the term "Prediction" rather than causality. The paper by Byrne (1974) seems to imply that if $R^2 = 1.00$, then since 100 percent of the predictor variance is accounted for there is causality.

Prediction and causality are not synonymous for given $R^2 = 1.00$ answering the question of which variables "cause" the phenomena is still not answered.

Byrne (1974) presents an excellent rationale for reporting R^2 to enable readers of research to determine whether significant differences are important differences. The error was the statement implies that reporting of R^2 demonstrates causality at the hypothesized level if 100 percent of the variance is accounted.

The major value of the Byrne (1974) paper is his presentation of a rationale for reporting R^2. There are two criticisms however: (1) The author has deliberately or accidentally misinterpreted causality as synonymous with prediction, and (2) The limitations of R^2 were ostensively absent. The calculation of R^2 is dependent upon the order of the variables in the equations, which maximizes the effect of the first variables, and that the variables are independent of each other. In most cases R^2 provides an overestimate of the predictor.

The purpose for most papers is given at the beginning, however, the purpose for this paper was not previously stated in order that the author would present a number of points which were necessary for interpreting Byrne (1974). This present paper should provide the reader with: (1) an understanding of the semantic difference of prediction and causality, (2) the value of Byrne (1974) for presenting a rationale for reporting R^2, and (3) a brief encounter with the limitations of R^2. A literary review of Byrne (1974) would be highly favorable as would a critical review. The misunderstandings, which may have developed from reading Byrne (1974) needed to be clarified and this was the only purpose for this paper.

REFERENCES: Byrne, J. F. The use of regression equations to demonstrate causality. Multiple Linear Regression Viewpoints, 1974, 5, 11-22.

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 5, No. 2, 1974

REGRESSION SOLUTIONS TO THE A X B X S DESIGN

John D. Williams
University of North Dakota

Summary - An alternative regression solution to the repeated measures design (A X B X S) is given, contrasting to a solution given earlier by Pyle (1974). The solution given here can be completed without altering the criterion measures to find the A and B effects. Also, the present solution can be translated quite easily into experimental design terminology.

Pyle (1974) has demonstrated one form of a regression solution to the A X B X S design. Recently, others have given regression solutions to the repeated measures design, including Fanning and Newman (1974) and Pohlman and McShane (1974).

Using Pyle's solution and data as a point of departure, several comments can be made. First, assuredly, his solution is correct. In that his solution requires redefining the criterion scores depending upon the particular main effect being considered, the solution is to some degree inefficient. While Fanning and Newman and Pohlman and McShane are, like Pyle's solution, in a Ward and Jennings (1973) vein, the present solution is meant to be helpful to those researchers who feel a need (1) to end up with results that can be framed in experimental design terminology and (2) is empirically easy to solve. Because of the second requirement, and because the A X B X S design is proportional, an intuitively pleasing solution can be given without resorting to the algebraic manipulations such as are found in Pyle's solution, but following the Ward and Jennings (1973) and Bottenberg and Ward (1963) tradition.

Pyle's data follow:

	B_1		B_2	
	A_1	A_2	A_1	A_2
S_1	1	3	4	4
S_2	2	2	5	4
S_3	3	1	6	6

Also, several variables can be formed which can be useful for arriving at a solution:

Y = the criterion score;

X_1 = 1 if from Subject 1, 0 otherwise;

X_2 = 1 if from Subject 2, 0 otherwise;

X_3 = 1 if from treatment A, 0 otherwise;

X_4 = 1 if from treatment B, 0 otherwise;

$X_5 = X_3 \cdot X_4$;

$X_6 = X_1 \cdot X_3$;

$X_7 = X_2 \cdot X_3$;

$X_8 = X_1 \cdot X_4$;

$X_9 = X_2 \cdot X_4$;

$X_{10} = X_1 \cdot X_3 \cdot X_4$; and

$X_{11} = X_2 \cdot X_3 \cdot X_4$.

Table 1 includes the previously given information.

TABLE 1

Regression Formulation of the A X B X S Design

	P_1	P_2	A_1	B_1	A_1B_1	A_1P_1	A_1P_2	B_1P_1	B_1P_2	$A_1B_1P_1$	$A_1B_1P_2$
Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
1	1	0	1	1	1	1	0	1	0	1	0
3	1	0	0	1	0	0	0	1	0	0	0
4	1	0	1	0	0	1	0	0	0	0	0
4	1	0	0	0	0	0	0	0	0	0	0
2	0	1	1	1	1	0	1	0	1	0	1
2	0	1	0	1	0	0	0	0	1	0	0
5	0	1	1	0	0	0	1	0	0	0	0
4	0	1	0	0	0	0	0	0	0	0	0
3	0	0	1	1	1	0	0	0	0	0	0
1	0	0	0	1	0	0	0	0	0	0	0
6	0	0	1	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0

For the moment, a model can be considered,

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 + b_9X_9 + b_{10}X_{10} + b_{11}X_{11} + e_1, \quad (1)$$

where $b_0 - b_{11}$ are regression coefficients,

$X_1 - X_{11}$ are as previously defined, and

e_1 = the error in prediction with this model.

If equation 1 is inspected, it can be seen that X_1 and X_2 represent the subject variables, X_3 represents the A variable, X_4 represents the B variable, X_5 represents the AB interaction, X_6 and X_7 represent the AS interaction, X_8 and X_9 represent the BS interaction and X_{10} and X_{11} represent the ABS interaction. Thus, the sources of variation involved with this model are A, B, AB, AS, BS, and ABS.

One difficulty ensues with the direct utilization of equation 1; because <u>all</u> the variables are represented in the design, there can be no error variance (i.e., $e_1 = 0$ for every Y score). However, if equation 1 can abstractly serve as a fullest model, (which it certainly is: $R^2 = 1$) a very direct approach can be used to find the ABS effect, which can be found as follows:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 + b_9X_9 + e_2, \quad (2)$$

which implicitly has the restriction $b_{10} = b_{11} = 0$. The ABS variation is given as $1 - R_2^2$ where R_2^2 results from the use of equation 2. The sum of squares due to ABS can be found as $(1 - R_2^2)(SS_T)$ or also as $(1 - R_2^2)(s_y^2)(n - 1)$. Many regression programs report the sum of squares for deviation from regression directly. SS_{ABS} can be found to be 2.167, with $1 - R^2 = .06583$.

In that the A X B X S design is a three way design, one simple approach is to conceive the design variously as a one way design three different times (with dimensions A, B and C), as a two way design three different times (with dimensions A and B, A and S and B and S); and finally, as a three way design (to find the ABS interaction). Table 2 includes a representation of the solution for the sources of variation.

TABLE 2

Solution to the A X B X S Design in Terms of R^2

Source of Variation	df	SS
A	1	$SS_T(R^2_A)$
B	1	$SS_T(R^2_B)$
S	2	$SS_T(R^2_S)$
AB	1	$SS_T(R^2_{A,B,AB} - R^2_{A,B})$
AS	2	$SS_T(R^2_{A,S,AS} - R^2_{A,S})$
BS	2	$SS_T(R^2_{B,S,BS} - R^2_{B,S})$
ABS	2	$SS_T(1 - R^2_{A,B,S,AB,AS,BS})$
Total	11	SS_T

It should be noted that several alternative methods could be used to calculate the sources of variation in Table 2. For example, SS_A is shown as being calculated directly using the linear model $Y = b_0 + b_1X_3 + e_3$ (3), leading to $SS_T(R^2_A)$ for the SS_A term. However, SS_A is also equal to $SS_T(R^2_{A,B} - R^2_B)$, and to $SS_T(R^2_{A,S} - R^2_S)$ and to $SS_T(R^2_{A,B,AB} - R^2_{A,AB})$, among others. There would also exist alternative ways to calculate SS_B and SS_S. While the several computations will yield the same numerical result, they represent a somewhat different conception of the main effect. For example, if SS_A is calculated from $SS_T(R^2_{A,B} - R^2_B)$, the question being answered is: what is the effect of A independent of B? When the cell frequencies are equal, the computational results of these approaches are identical.

Table 3 contains the necessary linear models for solving the A X B X S design, using the formulations from Table 2.

TABLE 3

Linear Models for Solving the A X B X S Design

Effect	Linear Model
A	$Y = b_0 + b_1X_3 + e_3$ (3)
B	$Y = b_0 + b_1X_4 + e_4$ (4)
S	$Y = b_0 + b_1X_1 + b_2X_2 + e_5$ (5)
A,B,AB	$Y = b_0 + b_1X_3 + b_2X_4 + b_3X_5 + e_6$ (6)
A,B	$Y = b_0 + b_1X_3 + b_2X_4 + e_7$ (7)
A,S,AS	$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_6 + b_5X_7 + e_8$ (8)
A,S	$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + e_9$ (9)
B,S,BS	$Y = b_0 + b_1X_1 + b_2X_2 + b_4X_4 + b_5X_8 + b_6X_9 + e_{10}$ (10)
B,S	$Y = b_0 + b_1X_1 + b_2X_2 + b_4X_4 + e_{11}$ (11)
A,B,S,AB,AS,BS	$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 + b_9X_9 + e_2$ (2)

While the main effects are resultant directly from the use of the corresponding linear models, the solutions for the interactions require the use of the difference (either in R^2 units or sum of squares) between two linear models. For example, SS_{AB} utilizes equations 6 and 7; similarly, SS_{AS} uses equations 8 and 9 and SS_{BS} uses equations 10 and 11. The solution for the data given earlier is shown in Table 4.

TABLE 4

Summary Table for A X B X S Design

Source of Variation	df	SS	MS	F
A	1	.083	.083	.077
B	1	24.083	24.083	22.237
S	2	2.167		
AB	1	.083	.083	.077
AS	2	2.167	1.083	
BS	2	2.167	1.083	
ABS	2	2.167	1.083	
Total	11	32.917		

where

$F_A = MS_A/MS_{AS}$,
$F_B = MS_B/MS_{BS}$, and
$F_{AB} = MS_{AB}/MS_{ABS}$.

Discussion

The use of regression for the solution to the repeated measures design has been previously described in a number of recent publications, each with their unique direction and differing in the models (but not the numerical solution) from the present solution. In turn, these solutions give identical results to the A X B X S design described by Lindquist (1953).

In that the repeated measures design for equal cell frequencies has been documented with a computer solution (Hartley, 1960) and the A X B X S design necessarily involves equal cell frequencies, one might reasonably ask, why spend time and effort necessary for a regression solution to this design. This thought surely will occur to the researcher who has even a modest sample size (say N = 30).

The rationale for using a regression solution would seem to lie in the intuitive value gained from working through the linear models. Hopefully, this experience will allow a better understanding of the research problem. A practical suggestion for researchers who wish a solution to the A X B X S design, is that they solve a small problem such as the one given here so that they can get a good "feel" for the problem; when analyzing a real life problem, the Hartley program is more efficient. If the researcher has an understanding of the meaning of the sources of variation, the actual computer solution used is not as an important a detail as the understanding of the solution.

REFERENCES

Bottenberg, R. A., and Ward, J. H. Applied multiple linear regression. Lackland Air Force Base, Texas: Personnel Research Laboratory PRL-TDR-63-6, 1963.

Fanning, F. W., and Newman, I. The development and demonstration of multiple regression modules for operant conditioning questions. Multiple Linear Regression Viewpoints, 1974, 4, No. 4, 65-87.

Hartley, A. O. Analysis of variance. In Ralston, A., and Wilf, H. S. (Eds.), Mathematical Methods for Digital Computers. New York: John Wiley & Sons, 1960.

Lindquist, E. F. Design and analysis of experiments in psychology and education. Boston: Houghton Mifflin Company, 1953.

Pohlman, J. T., and McShane, M. G. Applying the general linear model to repeated measures problems. Presented at the American Educational Research Association Convention, Chicago, April 17, 1974.

Pyle, T. W. Classical analysis of variance of completely within-subject factorial designs using regression techniques. Multiple Linear Regression Viewpoints, 1974, 5, No. 1, 23-32.

Ward, J. H., and Jennings, E. Introduction to linear models. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 5, No. 2, 1974

POSSIBLE RELATIONSHIPS BETWEEN PREDICTOR AND CRITERION VARIABLES

Robert G. St.Pierre
Boston College
Chestnut Hill, Massachusetts 02167

Introduction

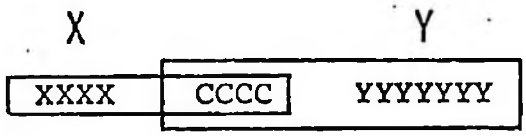
The possible relationships between a set of predictor variables and a criterion variable can be defined by thinking in terms of the amount of overlap between variables. Garrett (1946) and McNemar (1969) discuss the interpretation of the correlation coefficient between two variables when each variable can be perceived as the summation of a number of equally potent, equally likely, independent elements, being either present or absent. The correlation between two such variables is a function of the number of elements common to the variables and is given by:

$$r_{xy} = \frac{n_{xy}}{\sqrt{n_x + n_{xy}}\ \sqrt{n_y + n_{xy}}}$$

where n_x is the number of elements unique to variable X, n_y is the number of elements unique to variable Y, and n_{xy} is the number of elements common to both variables.

Figure 1 shows that if X is comprised of eight elements and Y is comprised of eleven elements with four elements being common to both X and Y, the correlation between the two variables is equal to .426.

FIGURE 1



$$r_{xy} = \frac{4}{\sqrt{4+4}\quad\sqrt{4+7}} = \frac{4}{\sqrt{8}\quad\sqrt{11}} = .426$$

This interpretation of the correlation between two variables has been extended to the multiple correlation and regression framework and a computer program called VARGEN was written (St.Pierre, 1974) as an educational tool to assist students in understanding regression, correlation, and the statistical properties of data. VARGEN will create a set of data consisting of a number of variables for each of a number of subjects in the following manner. For each subject, one hundred random numbers are drawn from a rectangular distribution which ranges from 0.0 to 1.0, and are placed in the cells of a ten by ten matrix which defines the universe of possible elements. The student can specify the relative size and location of each variable (predictor or criterion) within the universe. The value of each variable is determined by summing the random numbers contained in the cells of the universe that are covered by each variable. If two or more variables overlap, that is, have one or more cells in common, the random numbers contained in the overlapping cells are added to both variables. Once each variable has been calculated for the first subject, a new set of random numbers is generated and the value of each variable calculated for the second subject. This process is continued until data is generated for as many subjects as desired. A regression analysis is then performed on the created data set (the student defines which variables are predictors and which is the criterion), and a visual display showing which cells are covered by each variable is output.

In this way the student can pose and answer questions about a set of data such as: How much of the criterion variance in this data set is unique, and how much is held in common with predictor variance? How are the predictors related - to what extend do they share common variance?

Possible Relationships

This paper will define the possible relationships between a set of predictors and a criterion, and present a modified VARGEN visual display and regression analysis appropriate to that relationship (based on a sample size of 1000). The effect of each type of relationship on the square multiple correlation (R^2) will be the primary focus of the paper. Three major cases will be presented.

Case 1 - Situations where the predictors and the criterion share no common variance. While the predictors do not overlap the criterion, they may or may not share common variance with each other. In this case $R^2 = 0$ as no criterion variance can be explained by the predictors (See Figure 2).

<u>Case 2</u> - Situations where the predictors and criterion share common variance, however, the criterion retains some unique variance. While the predictors must overlap the criterion to some degree, they may or may not share common variance with each other. In this case $0 < R^2 < 1$ since some portion of the criterion variance can be explained by the predictors (See Figure 3).

<u>Case 3</u> - Situations where the criterion has no unique variance. This case can be divided into three parts as follows:

Case 3a - The predictors account for all and only the criterion variance, and do not share common variance with each other. In other words, the predictors completely overlap the criterion, do not overlap each other, and do not contain any variance not contained in the criterion. This is the only case in which $R^2 = 1$ (See Figure 4).

Case 3b - The predictors account for all and only the criterion variance, but share common variance with each other. That is, the predictors completely overlap the criterion, do not contain any variance not contained in the criterion, but overlap each other to some extent and thus do not account for unique portions of the criterion variance. In this case, $0 < R^2 < 1$ (See Figure 5).

Case 3c - The predictors account for all the criterion variance and also some non-criterion variance. That is, the predictors more than overlap the criterion. They may or may not share common variance with each other. In this case, $0 < R^2 < 1$ (See Figure 6).

Conclusion

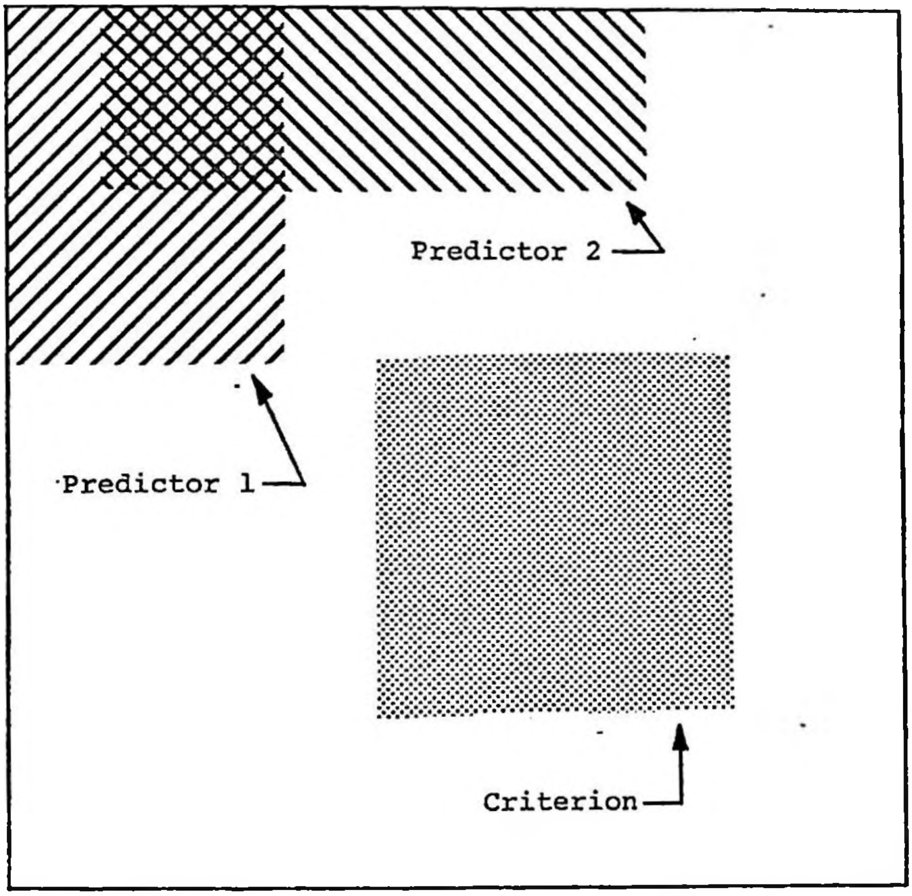
Situations can be expanded to include any number of predictors without altering the statements which have been made concerning R^2. While most students intuitively understand the results arrived at in Cases 1, 2 and 3a, the results of Cases 3b and 3c are somewhat more difficult to understand. It is important to realize that the only situation where $R^2=1$ occurs when the following three conditions are fulfilled:

1) the predictors account for all the criterion variance, that is, the predictors completely overlap the criterion;
2) the predictors account for only the criterion variance, that is, the predictors do not contain variance not belonging to the criterion;
3) the predictors are uncorrelated, that is, no predictor shares common variance with any other predictor.

While it may seem that R^2 should equal 1 for Cases 3a, 3b, and 3c, we can see that only Case 3a meets the three stated conditions. Case 3b violates condition 3, and Case 3c violates condition 2 (See Appendix A for a proof that $R^2 = 1$ only under the stated conditions). That is, whenever the predictors contain any error variance, or are correlated, it is impossible for R^2 to be equal to 1, regardless of the extent to which the predictors are correlated with the criterion.

Onv way to investigate the possible relationships between predictor and criterion variables is to use VARGEN, which, as described earlier, allows the student to create his own data sets. He is free to vary the amount of overlap between variables and thus acquire a feel for what is meant by common and unique variance and by overlap between variables, as well as to see the effect of varying relationships between predictors and criterion on a regression analysis.

FIGURE 2

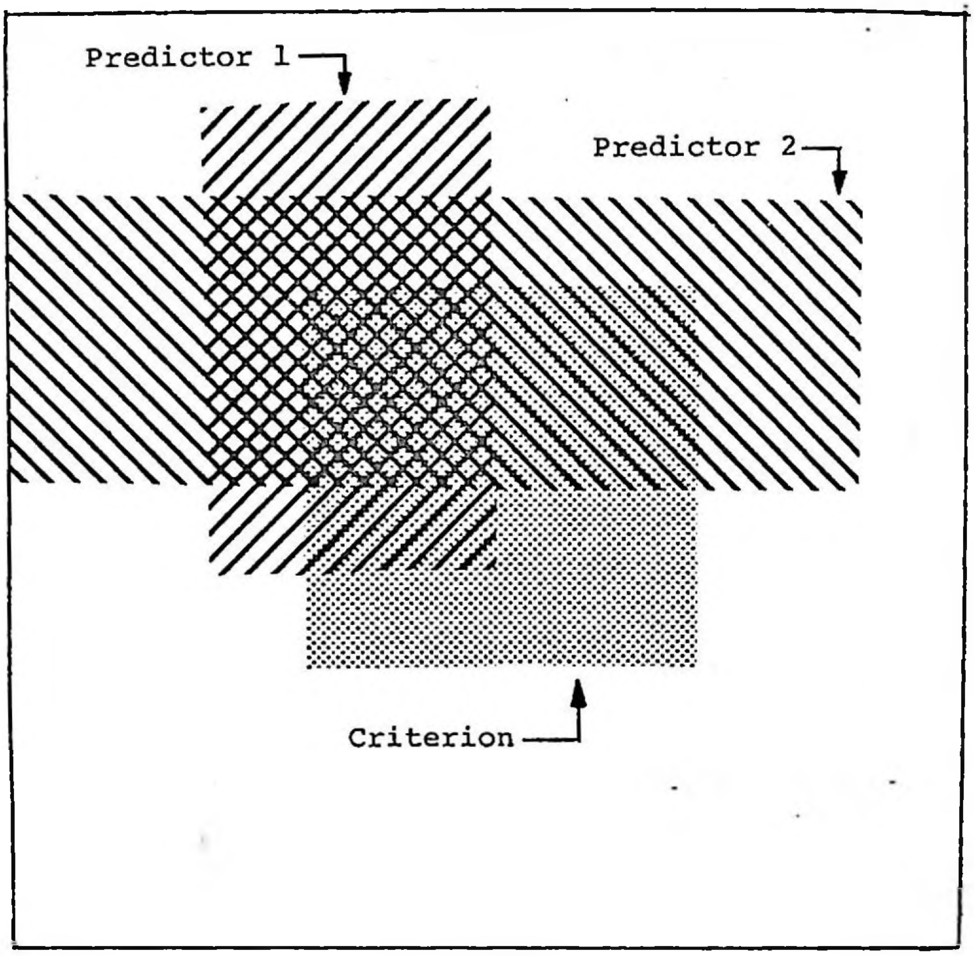


Multiple R Square = 0.002
Multiple R = 0.046
F (Anova on R) = 1.063
N.D.F.1 = 2.000
N.D.F.2 = 997.000

Predictor	B Weights	Beta	Beta Sq.	R(Criterion)	Beta*R
1	-0.013	-0.011	0.000	-0.025	0.000
2	-0.051	-0.041	0.002	-0.045	0.002

Intercept Constant = 8.397

FIGURE 3

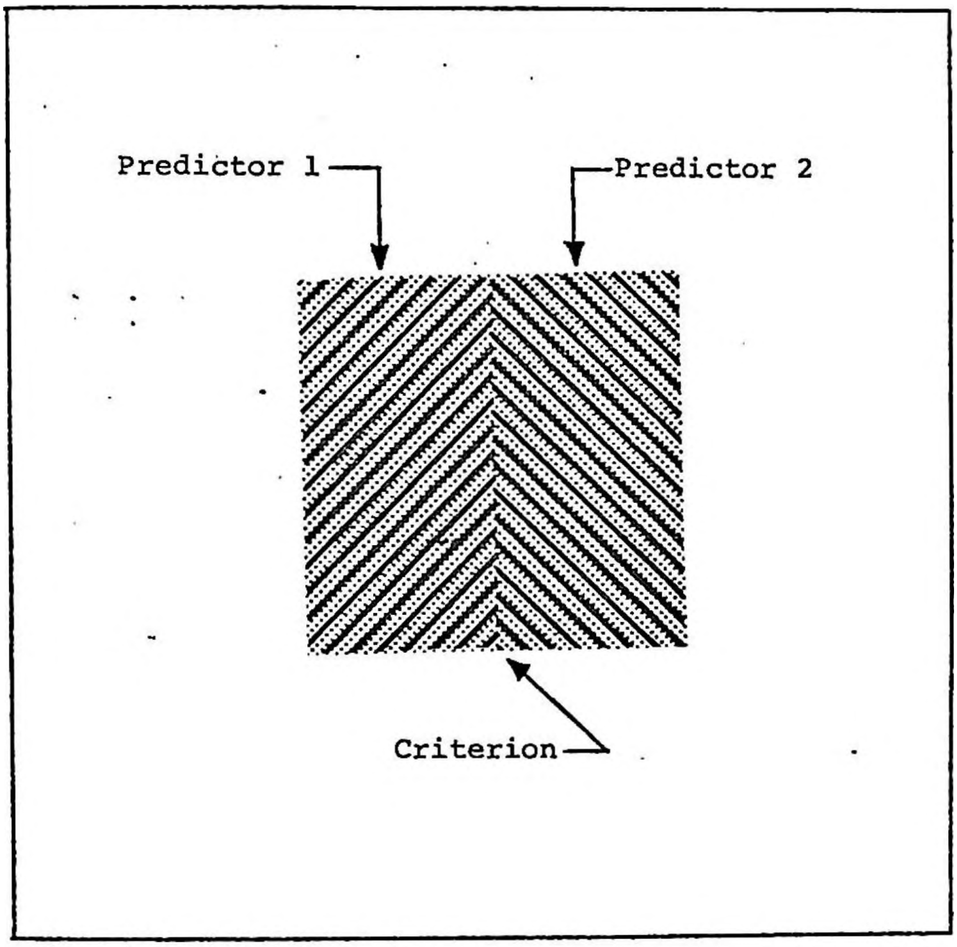


Multiple R Square	=	0.206
Multiple R	=	0.454
F (Anova on R)	=	129.497
N.D.F.1	=	2.000
N.D.F.2	=	997.000

Criterion	B Weights	Beta	Beta Sq.	R(Criterion)	Beta*R
1	0.250	0.230	0.053	0.365	0.084
2	0.241	0.302	0.091	0.405	0.122

Intercept Constant = 2.894

FIGURE 4

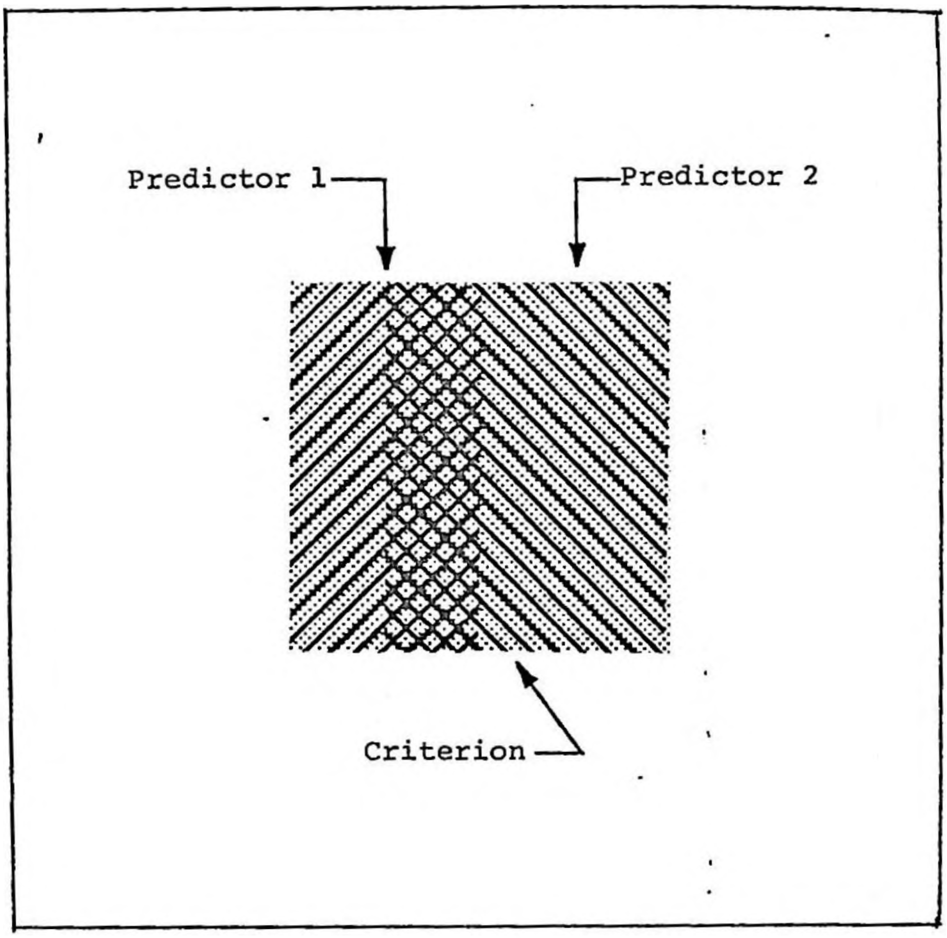


Multiple R Square = 1.000
Multiple R = 1.000
F (Anova on R) = *****
N.D.F.1 = 2.000
N.D.F.2 = 997.000

Predictor	B Weights	Beta	Beta Sq.	R(Criterion)	Beta*R
1	1.006	0.713	0.508	0.710	0.506
2	1.006	0.705	0.497	0.702	0.495

Intercept Constant = -0.048

FIGURE 5

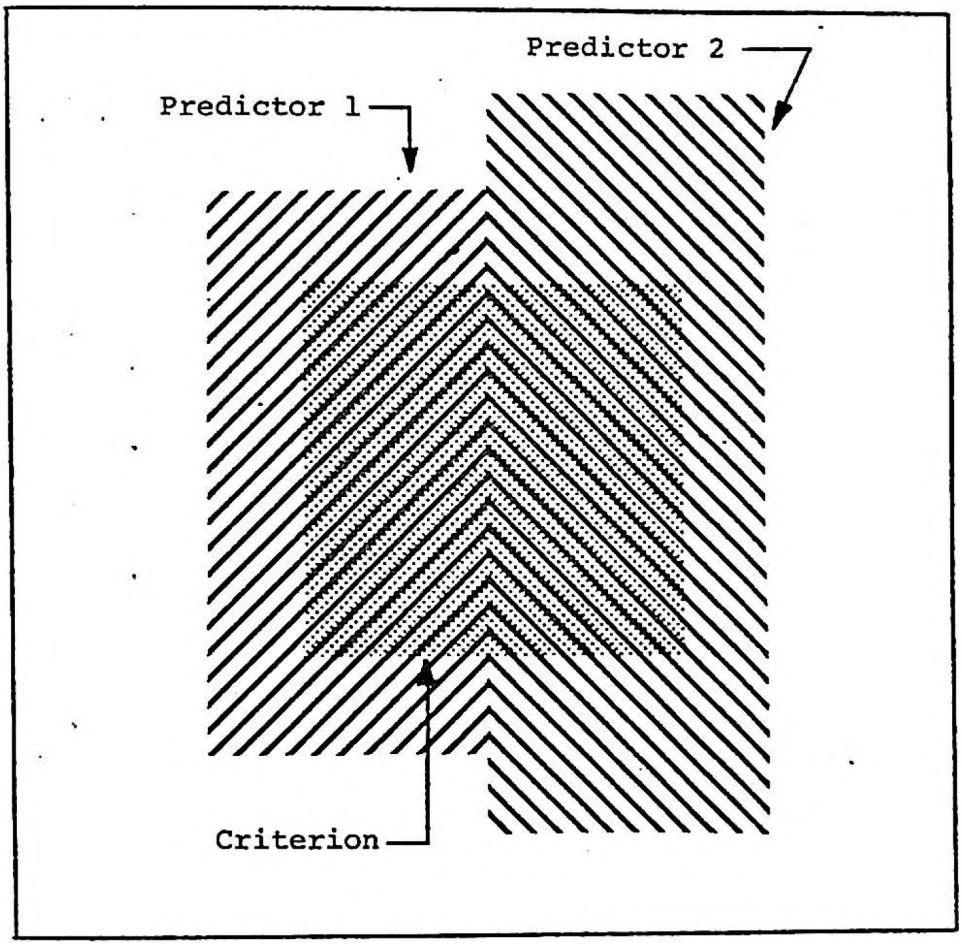


Multiple R Square = 0.903
Multiple R = 0.950
F (Anova on R) =4614.898
N.D.F.1 = 2.000
N.D.F.2 = 997.000

Predictor	B Weights	Beta	Beta Sq.	R(Criterion)	Beta*R
1	0.609	0.432	0.186	0.710	0.306
2	0.803	0.690	0.476	0.864	0.596

Intercept Constant = 0.745

FIGURE 6



Multiple R Square = 0.415
Multiple R = 0.645
R (Anova on R) = 354.212
N.D.F.1 = 2.000
N.D.F.2 = 997.000

Predictor	B Weights	Beta	Beta Sq.	R(Criterion)	Beta*R
1	0.464	0.488	0.239	0.473	0.231
2	0.371	0.438	0.192	0.421	0.185

Intercept Constant = -0.625

REFERENCES

Garrett, H. *Statistics in psychology and education*. New York: Longmans, Green and Company, 1946.

McNemar, Q. *Psychological statistics*. New York: Wiley, 1969.

St.Pierre, R. G., VARGEN: A multiple regression teaching program. *Multiple Linear Regression Viewpoints*, 1974, 4, 21-29.

APPENDIX A

Given the following common elements correlation formula:

$$r_{xy} = \frac{n_{xy}}{\sqrt{n_x n_y}}$$

where n_{xy} = the number of elements common to x and y

n_x = the number of elements in x

n_y = the number of elements in y

The three conditions stated earlier imply that the following relationships must hold for $R^2 = 1$. That is, $R^2 = 1$ if and only if:

$$n_{12} = 0 \tag{1}$$

$$n_1 + n_2 = n_c \tag{2}$$

$$n_{1c} = n_1 \tag{3}$$

$$n_{2c} = n_2 \tag{4}$$

Note that in the above conditions the subscript "c" refers to the criterion, "1" to the first predictor, and "2" to the second predictor.

Form the R matrix

$$R = \left[\begin{array}{cc|c} 1 & \frac{n_{12}}{\sqrt{n_1 n_2}} & \frac{n_{1c}}{\sqrt{n_1 n_c}} \\ \frac{n_{12}}{\sqrt{n_1 n_2}} & 1 & \frac{n_{2c}}{\sqrt{n_2 n_c}} \\ \hline \frac{n_{1c}}{\sqrt{n_1 n_c}} & \frac{n_{2c}}{\sqrt{n_2 n_c}} & 1 \end{array}\right] = \left[\begin{array}{c|c} R_{11} & R_{12} \\ \hline R_{21} & R_{22} \end{array}\right] \tag{5}$$

The multiple regression B weights can be obtained by: $B = R_{11}^{-1} R_{12}$ (6)

To solve for R_{11}^{-1} we need to obtain the determinant -

$$\det(R_{11}) = (1)(1) - \frac{n_{12}}{\sqrt{n_1 n_2}} \frac{n_{12}}{\sqrt{n_1 n_2}} = 1 - \frac{n_{12}^2}{n_1 n_2} \tag{7}$$

Solving the inverse by the cofactor method we obtain,

$$R_{11}^{-1} = \begin{bmatrix} \dfrac{1}{1 - \dfrac{n_{12}^2}{n_1 n_2}} & \dfrac{\dfrac{-n_{12}}{\sqrt{n_1 n_2}}}{1 - \dfrac{n_{12}^2}{n_1 n_2}} \\ \dfrac{\dfrac{-n_{12}}{\sqrt{n_1 n_2}}}{1 - \dfrac{n_{12}^2}{n_1 n_2}} & \dfrac{1}{1 - \dfrac{n_{12}^2}{n_1 n_2}} \end{bmatrix} \tag{8}$$

Simplifying (8) and substituting into (6) we obtain:

$$B = \begin{bmatrix} \dfrac{n_1 n_2}{n_1 n_2 - n_{12}^2} & \dfrac{-n_{12}\sqrt{n_1 n_2}}{n_1 n_2 - n_{12}^2} \\ \dfrac{-n_{12}\sqrt{n_1 n_2}}{n_1 n_2 - n_{12}^2} & \dfrac{n_1 n_2}{n_1 n_2 - n_{12}^2} \end{bmatrix} \begin{bmatrix} \dfrac{n_{1c}}{\sqrt{n_1 n_c}} \\ \dfrac{n_{2c}}{\sqrt{n_2 n_c}} \end{bmatrix} \tag{9}$$

$$= \begin{bmatrix} \dfrac{\sqrt{n_1}}{n_1 n_2 - n_{12}^2} \left[\dfrac{n_{1c} n_2 - n_{12} n_{2c}}{\sqrt{n_c}} \right] \\ \dfrac{\sqrt{n_2}}{n_1 n_2 - n_{12}^2} \left[\dfrac{n_{2c} n_1 - n_{12} n_{1c}}{\sqrt{n_c}} \right] \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{10}$$

The formula for R^2 is:

$$R^2 = B'R_{12} = b_1r_{1c}+b_2r_{2c} \tag{11}$$

Substituting (9) into (11) we obtain:

$$R^2 = \left[\frac{\sqrt{n_1}}{n_1n_2-n_{12}^2} \left[\frac{n_{1c}n_2-n_{12}n_{2c}}{\sqrt{n_c}}\right], \; \frac{\sqrt{n_2}}{n_1n_2-n_{12}^2} \left[\frac{n_{2c}n_1-n_{12}n_{1c}}{\sqrt{n_c}}\right] \right] \begin{bmatrix} \dfrac{n_{1c}}{\sqrt{n_1n_c}} \\ \dfrac{n_{2c}}{\sqrt{n_2n_c}} \end{bmatrix} \tag{12}$$

Which can be simplified to:

$$R^2 = \frac{n_{1c}(n_{1c}n_2-n_{2c}n_{12}) + n_{2c}(n_{2c}n_1-n_{1c}n_{12})}{n_c(n_1n_2-n_{12}^2)} \tag{13}$$

We now show that $R^2 = 1$ if (1), (2), (3), and (4) hold. Since $n_{12}=0$ from (1):

$$R^2 = \frac{n_{1c}(n_{1c}n_2) + n_{2c}(n_{2c}n_1)}{n_1n_2n_c} = \frac{n_{1c}^2}{n_1n_c} + \frac{n_{2c}^2}{n_2n_c} \tag{14}$$

let $n_1=n_{1c}$ and $n_2=n_{2c}$ from (3) and (4)

$$R^2 = \frac{n_1^2}{n_1n_c} + \frac{n_2^2}{n_2n_c} = \frac{n_1+n_2}{n_c} \tag{15}$$

Given $n_1+n_2=n_c$ from (2), we have

$$R^2 = \frac{n_c}{n_c} = 1 \tag{16}$$

Thus when all four conditions hold $R^2 = 1$.

We now show that if any one of the four conditions does not hold $R^2<1$

<u>Violate (1), i.e., let $n_{12}>0$,</u> then using (13), (3) and (4)

$$R^2 = \frac{n_1(n_1n_2-n_2n_{12}) + n_2(n_2n_1-n_1n_{12})}{n_c(n_1n_2-n_{12}^2)} \tag{17}$$

$$= \frac{n_1n_2(n_1+n_2-2n_{12})}{n_c(n_1n_2-n_{12}^2)} \tag{18}$$

Using (2) we have

$$R^2 = \frac{n_1n_2(n_c-2n_{12})}{n_c(n_1n_2-n_{12}^2)} = \frac{n_1n_2n_c-2n_1n_2n_{12}}{n_1n_2n_c-n_cn_{12}^2} \tag{19}$$

If $2n_1n_2n_{12} < n_cn_{12}^2$ then $R^2>1$ which is impossible.

If $2n_1n_2n_{12} > n_cn_{12}^2$ then $R^2<1$ which was to be proved.

If $2n_1n_2n_{12} = n_cn_{12}^2$ then $R^2=1$. We will show this is impossible.

$$2n_1n_2n_{12}=n_cn_{12}^2 \tag{20}$$

$$2n_1n_2=n_cn_{12} \quad \text{dividing by } n_{12} \tag{21}$$

$$2n_1n_2 = (n_1+n_2)\ n_{12} \quad \text{from (2)} \tag{22}$$

$$2n_1n_2 = n_1n_{12}+n_2n_{12} \tag{23}$$

$$n_1n_2+n_1n_2 = n_1n_{12}+n_2n_{12} \tag{24}$$

This is only true if $n_1=n_2=n_{12}=k$ (25)

which would imply $r_{12} = \frac{k}{\sqrt{k \cdot k}} = 1$. If this is true, the determinant of R_{11} in (7) is zero and so R_{11}^{-1} cannot be found. Thus, if (25) holds, R^2 cannot be obtained. Therefore, $2n_1n_2n_{12} > n_cn_{12}^2$ and

$$\frac{n_1n_2n_c-2n_1n_2n_{12}}{n_1n_2n_c-n_cn_{12}^2} = R^2 < 1$$

Violate (2), i.e.: let $n_1+n_2 \neq n_c$. The proof can be followed from (13), (14) and (15). Picking up at (15), we have

$$R^2 = \frac{n_1+n_2}{n_c} \tag{15}$$

If $n_1+n_2 > n_c$ then $R^2>1$ which is impossible.

If $n_1+n_2 = n_c$ then $R^2=1$. However, this contradicts our assumption.

If $n_1+n_2 < n_c$ then $R^2<1$ which was to be proved.

Violate (3), i.e.: let $n_1 \neq n_{1c}$. From (13) we have

$$R^2 = \frac{n_{1c}(n_{1c}n_2-n_{2c}n_{12}) + n_{2c}(n_{2c}n_1-n_{1c}n_{12})}{n_c(n_1n_2-n_{12}^2)} \tag{13}$$

Since $n_{12}=0$

$$R^2 = \frac{n_{1c}^2n_2 + n_{2c}^2n_1}{n_1n_2n_c} \tag{27}$$

using (4)

$$R^2 = \frac{n_{1c}^2 + n_1n_2}{n_1n_c} \tag{28}$$

If $n_{1c}^2 + n_1n_2>n_1n_c$ then $R^2 > 1$ which is impossible.

If $n_{1c}^2 + n_1n_2<n_1n_c$ then $R^2 < 1$ which was to be proved.

If $n_{1c}^2 + n_1n_2=n_1n_c$ then $R^2 = 1$. We will show this is impossible.

$$n_{1c}^2 + n_1 n_2 = n_1 n_c \tag{29}$$

Using (2)

$$n_{1c}^2 + n_1 n_2 = n_1 (n_1 + n_2)$$

$$n_{1c} = n_1$$

But this contradicts our. assumption, thus $R^2 < 1$

Violate (4), i.e.: let $n_2 \neq n_{2c}$. This proof follows directly from the proof when (3) was violated.

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 5, No. 2, 1974

REGRESSION CHI-SQUARE: TESTING FOR A LINEAR TREND IN PROPORTIONS IN A 2 X C CONTINGENCY TABLE

Samuel R. Houston
University of Northern Colorado
and
James T. Bolding, Jr.
University of Arkansas

ABSTRACT

The usual chi-square test for a 2 X C contingency table can fail to produce statistical significance when, in fact, a significant linear trend in proportions is present in the data. The ordinary chi-square test lacks power in the case when the variables can be considered ordered classifications. A regression chi-square test is described and illustrated with hypothetical data in which the usual chi-square test produced non-significant results even though a significant linear trend was present in the data.

Consider the following hypothetical data. A survey was conducted to gather information from voters (registered and non-registered) concerning support for a local school bond. Each voter was asked to check one of five categories ranging from strongly oppose to strongly favor and each response was used as a measure of support for the school bond. The school district believed that a more appropriate campaign could be conducted if systematic differences were detected on the initial survey. A statistican was consulted *ex post facto* (as usual!) and it was suggested that the data be arranged in a 2 X 5 contingency table and the usual chi-square test be completed.

Chi-Square Analysis of the Data

Table 1 is a summary table of responses from the survey.

TABLE 1

120 VOTERS CLASSIFIED ACCORDING TO REGISTRATION STATUS AND DEGREE OF SUPPORT FOR LOCAL SCHOOL BOND

Voter Registration Status	Attitude Toward Local School Bond					Total
	Strongly Oppose (SO)	Oppose (O)	Neutral (N)	Favor (F)	Strongly Favor (SF)	
Registered	7	15	25	9	4	60
Non-registered	3	11	25	13	8	60
Total	10	26	50	22	12	120

of p_i on X_i should be a good test criterion . On the null hypothesis each p_i is distributed about the same mean, estimated by $\bar{p}$, with variance $\overline{pq}/n_i$. The regression coefficient b is calculated as

$$b = [\sum a_i X_i - (\sum a_i)(\sum n_i X_i)/N] \;/\; [\sum n_i X_i^2 - (\sum n_i X_i)^2/N]$$

Its standard error, S.E. (b), is given by

$$\text{S.E.}(b) = [\bar{p}\cdot\bar{q}/\{\sum n_i X_i^2 - (\sum n_i X_i)^2/N\}]^{1/2}$$

For the data in Table 3, $b = [168 - (60)(360)/120]/\;[1216 - (360)^2/120]$
$= -0.088235$

and S.E.(b) = 0.042868. The normal deviate for testing the null hypothesis $\beta = 0$, is $Z = b/\text{S.E.}(b) = -2.06$. In this example the regression test gave Z = -2.06 (significant at 0.05 level) while the t-test gave t = -2.08 (118 d.f.). The difference in results arises because the 2 approaches use different large-sample approximations to the exact distribution of Z and t.

When the regression on p_i on X_i is used as a test criterion, it is of interest to examine whether the regression is linear. Armitage (1) has shown that this can be done by first computing chi-square $= \sum n_i (p_i - \bar{p})^2/\bar{p}\cdot\bar{q}$ $= \{\sum a_i p_i - A^2/N\} / \bar{p}\cdot\bar{q}$. This chi-square, with (C - 1) d.f., measures the total variation among the C values of p_i. The chi-square for linear regression, with 1 d.f., is found by squaring Z, since the square of a normal deviate has a chi-square distribution with 1 d.f. The difference,

$\chi^2_{(C-1)} - \chi^2_{(1)}$, is a chi-square with (C - 2) d.f. for testing the deviations of the p_i from their linear regression on the X_i. For the school bond example, the total chi-square is 4.72 with 4 d.f., while Z^2 = 4.24 with 1 d.f. Thus the chi-square for the deviations is 0.48 with 3 d.f., which is in agreement with the hypothesis of linearity.

REFERENCES

1. P. Armitage. Biometrics, 11:375(1955).
2. W. G. Cochran. Biometrics, 10:417(1954).
3. G. W. Snedecor and W. G. Cochran. Statistical Methods, 6th ed. .The Iowa State University Press, Ames, Iowa (1967).
4. F. Yates. Biometrika, 35:176(1948).

Oppose, Oppose, Neutral, Favor, and Strongly Favor classes. Having assigned the scores we may think of the school survey data as consisting of two independent samples of 60 and 60 voters, respectively. In Table 3 the data are presented in coded units and frequencies.

TABLE 3

SCHOOL BOND DATA DISPLAYED BY CODED SCORES

Attitude to School Bond	Voter Registration Status	
	Registered	Non-registered
X_i	f_i	f_i
1	7	3
2	15	11
3	25	25
4	9	13
5	4	8
Total	60	60
Mean ($\overline{X}$)	2.8	3.2

On the X scale the average attitude toward the school bond is 2.8 for the registered voters and 3.2 for the non-registered voters. The difference, $\overline{D}$, is -0.4, with a standard error $\pm$ 0.193 (118 d.f.), computed in the usual way. The value of t = -0.4/0.193 = -2.08 is significant at 0.05 level. Contrary to the initial chi-square test, this test reveals a significantly greater negative attitude score for registered voters than for non-registered voters.

Snedecor and Cochran (3) state that the assignment of coded scores is appropriate when (i) the phenomenon in question is one that could be measured on a continuous scale, and (ii) the ordered classification can be regarded as a kind of grouping of this continuous scale by a cruder scale that is the best we can do in the present state of knowledge. The net result, of course, is that we have been provided with a more powerful method of data analysis when contrasted with the initial chi-square test.

Regression Chi-Square Analysis of the Data

Another procedure, a regression chi-square test, has been suggested by Yates (4) and it has been shown to yield results consistent (not identical) with the t-test described above. For this test the null hypothesis is that there is no relation between p_i and X_i. The regression coefficient

The calculation of the usual chi-square value produces a value of 4.28, which is not significant for 4 degrees of freedom (d. f.). On the basis of this test the researcher might conclude at this point that there is no evidence to suggest that voter registration status is related to attitude of support for local school bond. Yet, a closer look at the data suggests that a linear trend in proportions might be present. A modification of Table 1 to include information which is suggestive of a linear trend appears in Table 2.

TABLE 2

TESTING A LINEAR REGRESSION ON p_i ON SUPPORT SCORE REGRESSION

Voter Registration Status	Attitude Toward Local School Bond					Total
	SO	O	N	F	SF	
Registered (a_i)	7	15	25	9	4	60
Non-registered	3	11	25	13	8	60
Total (n_i)	10	26	50	22	12	120 (N)
Registered Voter Rate ($p_i = a_i/n_i$)	0.70	0.56	0.50	0.45	0.33	0.5 ($\bar{p}$)
Coded Score (X_i)	1	2	3	4	5	

The proportions in the voter registration rate seem to be decreasing with the ordered categories which reflect attitude toward the local school bond issue. (See the row of Registered Voter Rates in Table 2.) The rates start at 0.70 (Strongly Oppose) and decrease sequentially to a low of 0.33 (Strongly Favor). This relationship is not detected in the usual chi-square test as the chi-square test is not sensitive to ordered categories which may be present in the classification variables.

<u>Two Sample t-test Analysis of the Data</u>

One alternative to the usual chi-square test, a t-test for two independent samples, has been suggested by Cochran (2), Armitage (1) and Yates (4). This approach involves attaching a coded score to each class so that an ordered scale is created. To illustrate from the school bond example, let us assign a 1, 2, 3, 4, 5, respectively to Strongly

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 5, No. 2, 1974

FOUR-WAY DISPROPORTIONATE HIERARCHICAL MODELS

John D. Williams
University of North Dakota

Summary - A solution for the four-way disproportionate fixed effects hierarchical model is given. Because there are 15 effects to be ordered, and the ordering will have a major impact on each different effect, the n-way hierarchical model should be used only when there is a strong _a priori_ ordering suggesting itself to the researcher.

Disproportionate n-way designs are troublesome* but necessary for the applied researcher. Most of the solutions for the disproportionate case are explained for the two-way model (Jennings, 1967; Overall and Spiegel, 1969; Williams, 1972). While allusions are made to solutions for higher dimensional designs, actual solutions are almost non-available. The most commonly accepted solution, called the fitting constants solution, wherein each main effect is measured independent of the other main effect, has the property that the sum of squares from each source of variation typically will not add to the sum of squares total when disproportionate cells occur.

For the applied researcher who has learned his statistical methodology in traditional experimental design courses where only proportionate cell frequencies were encountered, the non-additive quality of the fitting constants solution is bewildering. Because the hierarchical solution (Cohen, 1968) _is_ additive, some researchers may be attracted to it because of that quality. Correctly used, the hierarchical model requires the researcher to order the effects as to their research interest. For the four-way model, this is no mere task; there are 15 effects to be so-ordered: 4 main effects, 6 two-way interactions, 4 three-way interactions, and a four-way interaction.

*Without the prodding questions of Isadore Newman at the 1974 American Educational Research Association Convention, it is doubtful that I would have ever guessed any one would be _interested_ in a solution to the higher-dimensional hierarchical model.

The hierarchical solution would proceed in the following fashion: the most important effect is measured directly; the second most important effect is adjusted for the first effect; the third most important effect is adjusted for the first two effects; this process is continued until the fifteenth effect is adjusted for the first fourteen effects.

While a description of the model is useful, an <u>actual</u> set of data can be more instructive regarding the process for finding a solution. The data for a four-way disproportionate design is given in Table 1.

TABLE 1

Four-Way Disproportionate Data

		A_1		A_2		A_3	
		D_1	D_2	D_1	D_2	D_1	D_2
B_1	C_1	8 6	7 4	10 3	8 2	5 8 3	3 1
	C_2	5 3	8 2	11 7	12 8	5 2	7
B_2	C_1	1	12 4 2	9 9 7	9 7	5	4
	C_2	2	2 4 8	2 2 1	4 0	3 2 8 7	1 1 3 2
B_3	C_1	3 3 0 1	7	15 2 7	5 4	7 2 8	2 2 6
	C_2	2 8	2 1 3 2	1 1 3	2 8 3 2	1 1 1 1	2 4
B_4	C_1	13 12 10 6	5 3 11 0	17 12 4 1	15 13 12	9 6 4	17 12
	C_2	2 8 4 1	5 2 2 7	10 8 6	17	10 10 8	20

The coding for a solution to the four-way model is as follows:

Y = the criterion score,

X_1 = 1 if a member of A_1; 0 otherwise,

X_2 = 1 if a member of A_2; 0 otherwise,

X_3 = 1 if a member of B_1; 0 otherwise,

X_4 = 1 if a member of B_2; 0 otherwise,

X_5 = 1 if a member of B_3; 0 otherwise,

X_6 = 1 if a member of C_1; 0 otherwise,

X_7 = 1 if a member of D_1; 0 otherwise,

$$\left.\begin{array}{l} X_8 = X_1 \cdot X_3, \\ X_9 = X_1 \cdot X_4, \\ X_{10} = X_1 \cdot X_5, \\ X_{11} = X_2 \cdot X_3, \\ X_{12} = X_2 \cdot X_4, \\ X_{13} = X_2 \cdot X_5, \end{array}\right\} \text{AB interaction}$$

$$\left.\begin{array}{l} X_{14} = X_1 \cdot X_6, \\ X_{15} = X_2 \cdot X_6, \end{array}\right\} \text{AC interaction}$$

$$\left.\begin{array}{l} X_{16} = X_1 \cdot X_7, \\ X_{17} = X_2 \cdot X_7, \end{array}\right\} \text{AD interaction}$$

$$\left.\begin{array}{l} X_{18} = X_3 \cdot X_6, \\ X_{19} = X_4 \cdot X_6, \\ X_{20} = X_5 \cdot X_6, \end{array}\right\} \text{BC interaction}$$

$$\left.\begin{array}{l} X_{21} = X_3 \cdot X_7, \\ X_{22} = X_4 \cdot X_7, \\ X_{23} = X_5 \cdot X_7, \end{array}\right\} \text{BD interaction}$$

$$\left.\begin{array}{l} X_{24} = X_6 \cdot X_7, \end{array}\right\} \text{CD interaction}$$

$$\left.\begin{aligned}
X_{25} &= X_1 \cdot X_3 \cdot X_6,\\
X_{26} &= X_1 \cdot X_4 \cdot X_6,\\
X_{27} &= X_1 \cdot X_5 \cdot X_6,\\
X_{28} &= X_2 \cdot X_3 \cdot X_6,\\
X_{29} &= X_2 \cdot X_4 \cdot X_6,\\
X_{30} &= X_2 \cdot X_5 \cdot X_6,
\end{aligned}\right\}\quad \text{ABC interaction}$$

$$\left.\begin{aligned}
X_{31} &= X_1 \cdot X_3 \cdot X_7,\\
X_{32} &= X_1 \cdot X_4 \cdot X_7,\\
X_{33} &= X_1 \cdot X_5 \cdot X_7,\\
X_{34} &= X_2 \cdot X_3 \cdot X_7,\\
X_{35} &= X_2 \cdot X_4 \cdot X_7,\\
X_{36} &= X_2 \cdot X_5 \cdot X_7,
\end{aligned}\right\}\quad \text{ABD interaction}$$

$$\left.\begin{aligned}
X_{37} &= X_1 \cdot X_6 \cdot X_7,\\
X_{38} &= X_2 \cdot X_6 \cdot X_7,
\end{aligned}\right\}\quad \text{ACD interaction}$$

$$\left.\begin{aligned}
X_{39} &= X_3 \cdot X_6 \cdot X_7,\\
X_{40} &= X_4 \cdot X_6 \cdot X_7,\\
X_{41} &= X_5 \cdot X_6 \cdot X_7,
\end{aligned}\right\}\quad \text{BCD interaction}$$

$$\left.\begin{aligned}
X_{42} &= X_1 \cdot X_3 \cdot X_6 \cdot X_7,\\
X_{43} &= X_1 \cdot X_4 \cdot X_6 \cdot X_7,\\
X_{44} &= X_1 \cdot X_5 \cdot X_6 \cdot X_7,\\
X_{45} &= X_2 \cdot X_3 \cdot X_6 \cdot X_7,\\
X_{46} &= X_2 \cdot X_4 \cdot X_6 \cdot X_7,\\
X_{47} &= X_2 \cdot X_5 \cdot X_6 \cdot X_7.
\end{aligned}\right\}\quad \text{ABCD interaction}$$

Once the coding is performed, the question of ordering the effects remains. As indicated, the ordering of the several effects is no small problem; 15! different arrangements could be considered. In that many problems are such that the main effects (A,B,C,D) are usually of prime importance, the two-way interactions are of next importance (AB, AC, AD, BC, BD, CD), the three-way interactions are of next importance (ABC, ABD, ACD, BCD) and the four-way interaction is least in importance (ABCD), the problem can be reduced somewhat. While other orderings could be considered, the ordering given here is A, B, C, D, AB, AC, AD, BC, BD, CD, ABC, ABD, ACD, BCD ABCD.

Using the just mentioned ordering, the following linear models can be developed:

$Y = b_0 + b_1X_1 + b_2X_2 + e_1$, (1)

$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + e_2$, (2)

$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + e_3$, (3)

$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + e_4$, (4)

$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 +$
$b_9X_9 + b_{10}X_{10} + b_{11}X_{11} + b_{12}X_{12} + b_{13}X_{13} + e_5$, (5)

$Y = b_0 + b_1X_1 + \ldots + b_{14}X_{14} + b_{15}X_{15} + e_6$, (6)

$Y = b_0 + b_1X_1 + \ldots + b_{16}X_{16} + b_{17}X_{17} + e_7$, (7)

$Y = b_0 + b_1X_1 + \ldots + b_{18}X_{18} + b_{19}X_{19} + b_{20}X_{20} + e_8$, (8)

$Y = b_0 + b_1X_1 + \ldots + b_{21}X_{21} + b_{22}X_{22} + b_{23}X_{23} + e_9$, (9)

$Y = b_0 + b_1X_1 + \ldots + b_{24}X_{24} + e_{10}$, (10)

$Y = b_0 + b_1X_1 + \ldots + b_{25}X_{25} + b_{26}X_{26} + b_{27}X_{27} + b_{28}X_{28} + b_{29}X_{29} +$
$b_{30}X_{30} + e_{11}$, (11)

$Y = b_0 + b_1X_1 + \ldots + b_{31}X_{31} + b_{32}X_{32} + b_{33}X_{33} + b_{34}X_{34} + b_{35}X_{35} +$
$b_{36}X_{36} + e_{12}$, (12)

$Y = b_0 + b_1X_1 + \ldots + b_{37}X_{37} + b_{38}X_{38} + e_{13}$, (13)

$Y = b_0 + b_1X_1 + \ldots + b_{39}X_{39} + b_{40}X_{40} + b_{41}X_{41} + e_{14}$, (14)

$Y = b_0 + b_1X_1 + \ldots + b_{42}X_{42} + b_{43}X_{43} + b_{44}X_{44} + b_{45}X_{45} + b_{46}X_{46} +$
$b_{47}X_{47} + e_{15}$, (15)

where

$b_0 - b_{47}$ are regression coefficients (likely to be different from equation to equation),

$X_1 - X_{47}$ are as previously defined, and

$e_1 - e_{15}$ are the errors in prediction with each respective model.

Any given effect is found as the difference between the model including that effect and the previous model. Computer programs that include an analysis of variance for the regression allow a direct calculation of the sum of squares. Otherwise, $SS_T = (n-1)s^2$ where s^2 is the variance of the Y scores. Then, for the k^{th} effect considered,

$$SS_k = SS_T(R^2_k - R^2_{k-1}). \qquad (16)$$

The hierarchical solution for the data in Table 1 is given in Table 2.

TABLE 2

Hierarchical Solution for the Four-Way Data

Source of Variation	df	SS	MS	F
A	2	102.167	51.084	5.201
B	3	520.703	173.568	17.671
C	1	67.335	67.335	6.856
D	1	10.216	10.216	1.040
AB	6	137.709	22.952	2.337
AC	2	15.296	7.648	.779
AD	2	12.236	6.118	.623
BC	3	55.870	18.623	1.896
BD	3	48.640	16.213	1.651
CD	1	20.749	20.749	2.112
ABC	6	161.724	26.954	2.744
ABD	6	248.427	41.404	4.215
ACD	2	8.856	4.428	.451
BCD	3	37.032	12.344	1.257
ABCD	6	60.750	10.125	1.031
Within	75	736.688	9.822	
Total	122	2244.398		

Reducing the Number of Acceptable Solutions

While there are 15! different arrangements of ordering the effects for a four-way design, not all of those orderings will in fact yield a meaningful solution. For example, the AB interaction cannot be meaningfully interpreted unless the AB variables have been preceeded by both the A and B effects in the ordering. This is not to say that some interaction effects cannot take precedence over some main effects; rather, for any n^{th} order interaction effect to be included, it must first be preceded by the main effects and lower level interaction effects directly involved in a n-way layout of the data. As an example of a permissable ordering of the effects, the following ordering shows several interaction effects being measured prior to at least one main effect: A, B, AB, C, AC, BC, ABC, D, AD, BD, ABD, CD, ACD, BCD, ABCD.

The just given ordering demonstrates some of the flexibility of the present hierarchical model: the AB effect can be measured any time after the A and B main effects have been included, and can occur prior to the C and D main effects in the hierarchical ordering. Similarly, the AB, AC, BC and ABC interactions can precede the D main effects in the ordering.

A Note on Coding Systems

The coding system described earlier is a "dummy" or binary coding system. Some researchers may prefer "effect" coding, wherein the "left out" group is coded -1 for the effect in which it has been omitted. While both coding systems should give the same results, computer accuracy (or inaccuracy) can play a part in using as many as 47 predictor variables, as is necessary for a 3 X 4 X 2 X 2 design. In particular, the results reported in Table 2 were achieved using effect coding; the results using the binary coding system were very similar, never deviating more than .182 in sum of squares for any of the effects given in Table 2; all main effects and five of the interaction effects showed differences of less than .01 in sum of squares. Because the effect coding provided the least distortion in the final model in reproducing the expected values for the 48 cells, the results from the effect coding were used.

Discussion

Having provided a solution to the n-way hierarchical model (specifically, the four-way model), the usefulness of such a solution must still be examined. Perhaps it's major attractiveness to the traditional researcher is that the hierarchical model is an additive model and thus bears a superficial relationship to the more commonly encountered proportionate designs. This attractiveness must be countered with what could be a major drawback: if the ordering is changed, all the effects except the ABCD effect can be affected by the change so that the corresponding sum of squares will be different.

On the other hand, if either a fitting constants solution or unadjusted main effects solution (Williams, 1972) are employed, both methods yield unique (but usually non-additive) solutions. Further, these latter two solutions will yield identical results for all interaction effects; the main effects will of course differ. On the other hand, the interaction effects for the hierarchical model will usually differ for most if not all of the interaction effects from the fitting constants and unadjusted main effect solutions.

Thus, the concern for the additive nature of the hierarchical model should not be oversold: the use for the n-way hierarchical model would seem to be for the very specific situation wherein a researcher has a strong *a priori* ordering of the main effects and interaction terms.

REFERENCES

Cohen, J. Multiple regression as a general data-analytic system. Psychological Bulletin, 1968, 70, 426-443.

Jennings, E. Fixed effects analysis of variance by regression analysis. Multivariate Behavioral Research, 1967, 2, 95-108.

Overall, J. E., and Spiegel, D. K. Concerning least squares analysis of experimental data. Psychological Bulletin, 1969, 72, 311-322.

Williams, J. D. Two way fixed effects analysis of variance with disproportionate cell frequencies. Multivariate Behavioral Research, 1972, 7, 67-83.

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 5, No. 2, 1974

ISSUES IN TEACHING MULTIPLE LINEAR REGRESSION FOR BEHAVIORAL RESEARCH

Samuel R. Houston
University of Northern Colorado
and
E. LaMonte Ohlson
The Ohio State University

ABSTRACT

Several questions, issues and approaches to teaching a basic course in Multiple Linear Regression (MLR) are raised in this article. These include the following: 1)MLR as a generalized procedure; 2) Use of matrix algebra in MLR; 3) Redundant models in MLR; 4) Orthogonal coding; and 5) Data analysis and MLR. Suggestions and recommendations are made for the issues raised.

Issue 1--MLR as a Generalized Procedure

The behavioral researcher should be easily convinced after some exposition that the methods of MLR when contrasted with ANOVA are superior or the only appropriate method of attack in the following situations: 1) when the predictor variable is continuous, that is the experimental treatments with varying degrees of the same variable; 2) when cell frequencies in a factorial design are unequal and disproportionate; 3) when the predictor variables are both continuous and categorical, as in the analysis of covariance; 4) when studying trends in data, etc. As a result, there might be the temptation to discard classical ANOVA procedures in favor of the more generalized MLR techniques. In many situations the writers would support such a position, especially in the case in which the objective is to analyze data for possible relationships to be verified later with tight experimental studies. This role for the behavioral scientist has been supported by John Tukey (7).

However, in the more restricted experimental study, the researcher might be investigating a set of very specific questions which lend themselves to a classical ANOVA design. Suppose, for example, the answers to the research questions can be obtained from a classical 2-factor ANOVA design. Should the researcher rely on MLR in this case? We strongly hesitate to make such a recommendation in favor of the generalized procedure. Would it not be a sad discovery to learn that one's elementary school son came home from school only to announce that since he had discovered that addition is the most generalized of arithmetic operations (i.e., subtraction, division and multiplication can be conceptualized and expressed as special cases of addition) he would rely on methods of addition in all future encounters with problems involving multiplication! So, too, it might be argued that classical ANOVA procedures are more sensitive and efficient in some research situations than are generalized MLR techniques. We recommend that the behavioral scientist develop expertise in both approaches and, armed with knowledge of their interrelationships, allow the research questions to dictate which approach should be undertaken.

Issue 2--Use of Matrix Algebra in MLR

A study of the MLR textbooks of Kelly et al. (4) and Ward and Jennings (8) reveals that the authors introduce the students to the concept of vectors and avoid the use of matrix algebra. While it might be argued that an introductory course in linear models can be taught successfully sans matrix algebra, the mechanics of matrix algebra, which should be considered an essential language for a behavioral scientist, are not difficult. Indeed, a behavioral scientist will need to develop skills in other multivariate procedures such as factor analysis, canonical correlation, MANOVA, discriminant analysis, and classification techniques, in which case matrix algebra is a must. Let us not continue to avoid the use of matrix algebra--matrix algebra is usually elegant in its simplicity while the corresponding scalar algebra is involved and messy.

Issue 3--Redundant Models in MLR

The approach to model construction and testing taken by Kelly et al. (4) and Ward and Jennings (8) is one in which appropriate full and restricted linear models are determined and tested with the variance ratio test. If the model involves a classical ANOVA or analysis of covariance design, it is not uncommon to find full and/or restricted models with redundant vectors. (Never let us cease to be thankful for the iterative program which permits a solution when redundancies are present!) Undoubtedly, models which have a vector associated with group membership for each group are easier to conceptualize and formulate initially. However, the researcher must still recognize the dependencies involved if he is to determine the correct degrees of freedom for a test of the significance of the difference of R^2s between two models. So why keep dependencies in the full and/or restricted linear models? Should not the

the Law of Parsimony be employed to remind the scientist that he should keep the formulation of both full and restricted models as simple or brief as possible, i.e., redundant-free? [See the textbook by Mendenhall (5) as an example of a textbook in linear models in which no dependencies appear in either the complete or reduced linear model.]

Issue 4--Orthogonal Coding

Another topic which should be considered in a course in MLR is the use of orthogonal coding. When one does not have access to a computer, the use of orthogonal coding is extremely useful in reducing and simplifying the calculations necessary for the regression analysis. The virtue of orthogonal coding is that in the basic formula $\underline{b} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{Y}$, $\underline{X}'\underline{X}$ is a diagonal matrix. The inverse of such a matrix is a diagonal matrix whose elements are the reciprocals of $\underline{X}'\underline{X}$. It thus becomes easy to solve for the vector of b's, and to obtain the standard errors of the b's. [See Mendenhall (5) for details.]

In addition, the calculation of the over-all R^2 for the full linear model is easily accomplished with the use of orthogonal coding. One first calculates zero-order correlations between each coded vector and the dependent variable. The over-all R^2 will simply be the sum of the squared zero-ordered correlations as the correlation between each of the coded vectors (orthogonal) is zero. Furthermore, when orthogonal comparisons are used for facility of calculations, one does not interpret the individual comparisons, but simply adds their contributions to obtain R^2 or the regression sum of squares. When the F-ratio associated with the R^2 is significant, one may proceed with post hoc comparisons between means.

Issue 5--Data Analysis and MLR

It is the belief of the writers that MLR will be relied on by the behavioral scientist more as a data analytic tool, often suggesting ideas for subsequent research, rather than a procedure for analyzing a classical design. With that in mind, the researcher should develop skills in analyzing regression models. The Coleman Report (1) is an example of a study in which it is imperative that the researcher understand the method of analysis and the appropriate conclusions and implications which can be derived from that method. Examples of methods which should be considered as useful topics in a acourse on MLR are path analysis (9), factor regression analysis (2), stepwise regression analysis (3), setwise regression analysis (10), commonality analysis (6), the backward elimination analysis (3), the forward selection approach (3), and the stagewise regression procedure (3).

Recommendations

It is recommended that the behavioral scientist develop additional skills to include the mechanics of matrix algebra, the ability to express non-redundant regression models, facility with orthogonal coding, and greater exposure to strategies for explaining regression models. We feel that these topics (and others) might well be presented by Joe Ward and Earl Jennings in an advanced workshop to individuals who had completed the beginning workshop (or its equivalent) in an AERA pre-session.

REFERENCES

1. Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A. M., Weinfeld, F.D., & York, R.L. Equality of educational opportunity. Washington, D.C.: U.S. Dept. of Health, Education, and Welfare, Office of Education, U.S. Govt. Printing Office, 1966.
2. Connett, W. W., Houston, S. R., & Shaw, D. The use of factor regression in data analysis. Multiple Linear Regression Viewpoints, 1972, 2, 46-49.
3. Draper, N.R., & Smith, H. Applied regression analysis. New York: Wiley, 1966.
4. Kelly, F.J., Beggs, D.L., McNeil, K.A., Eichelberger, T., & Lyon, J. Research design in the behavioral sciences: Multiple regression approach. Carbondale, Illinois: So. Illinois U. Press, 1969.
5. Mendenhall, W. Introduction to linear models and the design and analysis of experiments. Belmont, California: Wadsworth Publishing Company, Inc., 1968.
6. Mood, A.M. Partitioning variance in multiple regression analyses as a tool for developing learning models. American Educational Research Journal, 1971, 8, 191-202.
7. Tukey, J.W. The future of data analysis. Annals of Mathematical Statistics, 1962, 33, 1-67.
8. Ward, J. H., Jr., & Jennings, E. Introduction to linear models. Englewood Cliffs: Prentice-Hall, Inc., 1973.
9. Williams, J.D.& Linden, A.C. Regression computer programs for setwise regression, analysis of covariance with multiple covariates, and two-way analysis of variance and covariance with disproportionate cell frequencie. (Paper presented at annual American Educational Research Association Meeting, Chicago, Illinois, 1974)
10. Wright, S. Path coefficients and path regressions: Alternative or complementary concepts. Biometrics, 1960, 16, 189-202.

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 5, No. 2, 1974

AN APPLICATION OF THE HIGHER ORDER FACTORIAL ANALYSIS DESIGNS WITH DISPROPORTIONALITY OF CELLS

Deitchman, R., Newman I., Burkholder, J. and Sanders, R. E.

The University of Akron

Abstract: The results of the application of several techniques available for analyzing data when there is disproportionality of cells is reported. How accurately they answer the research questions asked is also discussed.

Williams (1974) has reported a possible solution for the four-way disproportionate fixed effects model. We are just a few of the applied researchers who have been plagued with the disproportionality of cells when using higher order factorial analysis designs. There are a variety of solutions available which can be divided into two major categories: a) approximate solutions, and b) exact solutions. Some examples of approximate solutions are: randomly eliminate some data, or calculate scores just on means thereby reducing the N and therefore power (Edwards, 1967). Since we were interested in an increase in accuracy and power, only exact solutions were of specific interest to us. All the exact solutions use multiple linear regression procedures. The researchers of the following study were confronted with a 4 x 2 x 2 x 5 factorial design with disproportionality and repeated measures. The first attempt was to find a computer program that would adjust for disproportionality that used a traditional analysis of variance solutions--no such program seemed available. A review of the literature dealing with such problems (Jennings, 1967; Overall and Spiegel, 1969; Williams, 1972; Williams, 1974; Applebaum, Elliot and Cramer, 1974) seemed to indicate that there are three major exact solutions that may be appropriate to solving our problem:

1. Hierarchial model (Cohen, 1968; Williams, 1974). This method requires apriori knowledge of which variables in a solution are more important or of greatest interest. With this approach, each effect is only adjusted for the effects that are of greatest interest. Since in our particular problem such apriori knowledge was not totally appropriate, this solution was not attempted.

2. The unadjusted main effects methods (Williams, 1974). This solution requires that the main effects may not be adjusted while the interaction effects would be adjusted for the variance accounted for by the main effects. Since one of our main concerns was with how much variance a particular main effect was accounted for, this solution was rejected.

3. Method of fitting constants. Since our interest was in the independent amount of variance accounted for by a particular treatment, the method of fitting constants was employed. This method as a solution for disproportionality has been suggested and described most frequently by Winer, 1971; Scheffe, 1959; and others. In this procedure each main effect is adjusted for each other main effect and the interactions are adjusted for all main effects. This procedure and the analysis of

covariance procedure which was similar to what has been suggested by Overall and Spiegal (1969) were compared on data obtained in an actual experimental setting.

Method

The experiment consisted of a 4 x 2 x 2 x 5 factorial design. Rat mothers were crowded into four groups, the A effect (8, 4, 2, 1), during the gestation period from days 1-19. During the first 16 days of the litter period mothers were observed two hours in the morning and two hours in the evening with their litters (B effect). A 12 item check list was employed. The behavior of each mother was checked one at a time until all mothers were checked. This cycle was repeated 40 times each session (each 20 recordings allowing for a first vs. second half comparison--C effect) for 15 days (the data was collapsed into five 3-day periods--D effect) twice a day (A.M. vs. P.M.). The full model containing all variables is presented below:

Description of Variables

Y_1 = The criterion score

X_1 = 1 if a member of A_1; 0 otherwise

X_2 = 1 if a member of A_2; 0 otherwise

X_3 = 1 if a member of A_3; 0 otherwise

X_4 = 1 if a member of B_1; 0 otherwise

X_5 = 1 if a member of G_1; 0 otherwise

X_6 = 1 if a member of D_1; 0 otherwise

X_7 = 1 if a member of D_2; 0 otherwise

X_8 = 1 if a member of D_3; 0 otherwise

X_9 = 1 if a member of D_4; 0 otherwise

$$\left.\begin{array}{l} X_{10} = X_1 \cdot X_4 \\ X_{11} = X_2 \cdot X_4 \\ X_{12} = X_3 \cdot X_4 \end{array}\right\} \text{AB interaction}$$

$$\left.\begin{array}{l} X_{13} = X_1 \cdot X_5 \\ X_{14} = X_2 \cdot X_5 \\ X_{15} = X_3 \cdot X_5 \end{array}\right\} \text{AC interaction}$$

$$\left.\begin{array}{l} X_{16} = X_1 \cdot X_6 \\ X_{17} = X_1 \cdot X_7 \\ X_{18} = X_1 \cdot X_8 \\ X_{19} = X_1 \cdot X_9 \\ X_{20} = X_2 \cdot X_6 \\ X_{21} = X_2 \cdot X_7 \\ X_{22} = X_2 \cdot X_8 \\ X_{23} = X_2 \cdot X_9 \\ X_{24} = X_3 \cdot X_6 \\ X_{25} = X_3 \cdot X_7 \\ X_{26} = X_3 \cdot X_8 \\ X_{27} = X_3 \cdot X_9 \end{array}\right\} \text{AD interaction}$$

$$\left.X_{28} = X_4 \cdot X_5\right\} \text{BC interaction}$$

$$\left.\begin{array}{l} X_{29} = X_4 \cdot X_6 \\ X_{30} = X_4 \cdot X_7 \end{array}\right\} \text{BD interaction}$$

$$\left.\begin{array}{l} X_{31} = X_4 \cdot X_8 \\ X_{32} = X_4 \cdot X_9 \end{array}\right\} \text{BD interaction}$$

$$\left.\begin{array}{l} X_{33} = X_5 \cdot X_6 \\ X_{34} = X_5 \cdot X_7 \\ X_{35} = X_5 \cdot X_8 \\ X_{36} = X_5 \cdot X_9 \end{array}\right\} \text{CD interaction}$$

$$\left.\begin{array}{l} X_{37} = X_1 \cdot X_4 \cdot X_5 \\ X_{38} = X_2 \cdot X_4 \cdot X_5 \\ X_{39} = X_3 \cdot X_4 \cdot X_5 \end{array}\right\} \text{ABC interaction}$$

$$\left.\begin{array}{l} X_{40} = X_1 \cdot X_4 \cdot X_6 \\ X_{41} = X_1 \cdot X_4 \cdot X_7 \\ X_{42} = X_1 \cdot X_4 \cdot X_8 \\ X_{43} = X_1 \cdot X_4 \cdot X_9 \\ X_{44} = X_2 \cdot X_4 \cdot X_6 \\ X_{45} = X_2 \cdot X_4 \cdot X_7 \\ X_{46} = X_2 \cdot X_4 \cdot X_8 \\ X_{47} = X_2 \cdot X_4 \cdot X_9 \\ X_{48} = X_3 \cdot X_4 \cdot X_6 \\ X_{49} = X_3 \cdot X_4 \cdot X_7 \\ X_{50} = X_3 \cdot X_4 \cdot X_8 \\ X_{51} = X_3 \cdot X_4 \cdot X_9 \end{array}\right\} \text{ABD interaction}$$

$$\left.\begin{array}{l} X_{52} = X_1 \cdot X_5 \cdot X_6 \\ X_{53} = X_1 \cdot X_5 \cdot X_7 \\ X_{54} = X_1 \cdot X_5 \cdot X_8 \\ X_{55} = X_1 \cdot X_5 \cdot X_9 \\ X_{56} = X_2 \cdot X_5 \cdot X_6 \\ X_{57} = X_2 \cdot X_5 \cdot X_7 \\ X_{58} = X_2 \cdot X_5 \cdot X_8 \\ X_{59} = X_2 \cdot X_5 \cdot X_9 \\ X_{60} = X_3 \cdot X_5 \cdot X_6 \\ X_{61} = X_3 \cdot X_5 \cdot X_7 \end{array}\right\} \text{ACD interaction}$$

Description of Variables, Continued

$X_{62} = X_3 \cdot X_5 \cdot X_8$
$X_{63} = X_3 \cdot X_5 \cdot X_9$
ACD interaction

$X_{64} = X_4 \cdot X_5 \cdot X_6$
$X_{65} = X_4 \cdot X_5 \cdot X_7$
$X_{66} = X_4 \cdot X_5 \cdot X_8$
$X_{67} = X_4 \cdot X_5 \cdot X_9$
BCD interaction

$X_{68} = X_1 \cdot X_4 \cdot X_5 \cdot X_6$
$X_{69} = X_1 \cdot X_4 \cdot X_5 \cdot X_7$
$X_{70} = X_1 \cdot X_4 \cdot X_5 \cdot X_8$
$X_{71} = X_1 \cdot X_4 \cdot X_5 \cdot X_9$
$X_{72} = X_2 \cdot X_4 \cdot X_5 \cdot X_6$
$X_{73} = X_2 \cdot X_4 \cdot X_5 \cdot X_7$
$X_{74} = X_2 \cdot X_4 \cdot X_5 \cdot X_8$
$X_{75} = X_2 \cdot X_4 \cdot X_5 \cdot X_9$
$X_{76} = X_3 \cdot X_4 \cdot X_5 \cdot X_6$
$X_{77} = X_3 \cdot X_4 \cdot X_5 \cdot X_7$
$X_{78} = X_3 \cdot X_4 \cdot X_5 \cdot X_8$
$X_{79} = X_3 \cdot X_4 \cdot X_5 \cdot X_9$
ABCD interaction

$X_{80} = R_1$ 1 if from Rat 1; otherwise 0
$X_{81} = R_2$ 1 if from Rat 2; otherwise 0
$X_{82} = R_3$ 1 if from Rat 3; otherwise 0
$X_{83} = R_4$ 1 if from Rat 4; otherwise 0
$X_{84} = R_5$ 1 if from Rat 5; otherwise 0
$X_{85} = R_6$ 1 if from Rat 6; otherwise 0
$X_{86} = R_7$ 1 if from Rat 7; otherwise 0
$X_{87} = R_8$ 1 if from Rat 8; otherwise 0
$X_{88} = R_9$ 1 if from Rat 9; otherwise 0
$X_{89} = R_{10}$ 1 if from Rat 10; otherwise 0
$X_{90} = R_{11}$ 1 if from Rat 11; otherwise 0
$X_{91} = R_{12}$ 1 if from Rat 12; otherwise 0
$X_{92} = R_{13}$ 1 if from Rat 13; otherwise 0
$X_{93} = R_{14}$ 1 if from Rat 14; otherwise 0
$X_{94} = R_{15}$ 1 if from Rat 15; otherwise 0
$X_{95} = R_{16}$ 1 if from Rat 16; otherwise 0
$X_{96} = R_{17}$ 1 if from Rat 17; otherwise 0
$X_{97} = R_{18}$ 1 if from Rat 18; otherwise 0
$X_{98} = R_{19}$ 1 if from Rat 19; otherwise 0
$X_{99} = R_{20}$ 1 if from Rat 20; otherwise 0
$X_{100} = R_{21}$ 1 if from Rat 21; otherwise 0
$X_{101} = R_{22}$ 1 if from Rat 22; otherwise 0
$X_{102} = R_{23}$ 1 if from Rat 23; otherwise 0
$X_{103} = R_{24}$ 1 if from Rat 24; otherwise 0
$X_{104} = R_{25}$ 1 if from Rat 25; otherwise 0
$X_{105} = R_{26}$ 1 if from Rat 26; otherwise 0
$X_{106} = R_{27}$ 1 if from Rat 27; otherwise 0
$X_{107} = R_{28}$ 1 if from Rat 28; otherwise 0
$X_{108} = R_{29}$ 1 if from Rat 29; otherwise 0
$X_{109} = R_{30}$ 1 if from Rat 30; otherwise 0
$X_{110} = R_{31}$ 1 if from Rat 31; otherwise 0
$X_{111} = R_{32}$ 1 if from Rat 32; otherwise 0
$X_{112} = R_{33}$ 1 if from Rat 33; otherwise 0
$X_{113} = R_{34}$ 1 if from Rat 34; otherwise 0
$X_{114} = R_{35}$ 1 if from Rat 35; otherwise 0
$X_{115} = R_{36}$ 1 if from Rat 36; otherwise 0
$X_{116} = R_{37}$ 1 if from Rat 37; otherwise 0

$E = (Y - \hat{Y})$ Error
U = Unit vector, 1 for each replicate in the study
$b_0, b_1, \ldots + b_{116}$ partial Regression Weight

Some representative models used for testing some of the most pertinent questions are presented below:

$$Y_1 = b_0U + b_1X_1 + b_2X_2 + b_3X_3 + * + E_1 \quad (1)$$

$$Y_1 = b_0U + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + * + E_2 \quad (2)$$

$$Y_1 = b_0U + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + * + E_3 \quad (3)$$

$$Y_1 = b_0U + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 + b_9X_9 + * + E_4 \quad (4)$$

$$Y_1 = b_0U + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 + b_9X_9 + b_{10}X_{10} + b_{11}X_{11} + b_{12}X_{12} + * + E_5 \quad (5)$$

$$Y_1 = b_0U + b_1X_1 + \ldots + b_{13}X_{13} + b_{14}X_{14} + b_{15}X_{15} + * + E_6 \quad (6)$$

$$Y_1 = b_0U + b_1X_1 + \ldots + b_{16}X_{16} + b_{17}X_{17} + b_{18}X_{18} + b_{19}X_{19} + b_{20}X_{20} + b_{21}X_{21} + \ldots + b_{27}X_{27} + * + E_7 \quad (7)$$

$$Y_1 = b_0U + b_1X_1 + \ldots + b_{28}X_{28} + * + E_8 \quad (8)$$

$$Y_1 = b_0U + b_1X_1 + \ldots + b_{29}X_{29} + b_{30}X_{30} + b_{31}X_{31} + b_{32}X_{32} + * + E_9 \quad (9)$$

$$Y_1 = b_0U + b_1X_1 + \ldots + b_{33}X_{33} + b_{34}X_{34} + b_{35}X_{35} + b_{36}X_{36} + * + E_{10} \quad (10)$$

$$Y_1 = b_0U + b_1X_1 + \ldots + b_{37}X_{37} + b_{38}X_{38} + b_{39}X_{39} + * + E_{11} \quad (11)$$

$$Y_1 = b_0U + b_1X_1 + \ldots + b_{40}X_{40} + \ldots + b_{51}X_{51} + * + E_{12} \quad (12)$$

$$Y_1 = b_0U + b_1X_1 + \ldots + b_{52}X_{52} + \ldots + b_{63}X_{63} + * + E_{13} \quad (13)$$

$$Y_1 = b_0U + b_1X_1 + \ldots + b_{64}X_{64} + b_{65}X_{65} + b_{66}X_{66} + b_{67}X_{67} + * + E_{14} \quad (14)$$

$$Y_1 = b_0U + b_1X_1 + \ldots + b_{68}X_{68} + \ldots + b_{79}X_{79} + * + E_{15} \quad (15)$$

Where: Y_1 = the criterion score

$b_0 - b_{116}$ are regression coefficients (likely to be different from equation to equation)

$X_1 - X_{79}$ are as previously defined

$E_1 - E_{15}$ are the errors in prediction with each respective mode

*In all cases add $R_1 \ldots R_{37}$

The schematic representation of the research design and the full regression model that reflects all the variables is presented in Tables 1 and 2.

TABLE 1

$$Y = b_0U + A_1 + A_2 + A_3 + B_1 + C_1 + D_1 + D_2 + D + D_4 + \overset{\text{AxB}}{(A_1B_1 + A_2B_1 + A_3B_1)} + \overset{\text{AxC}}{(A_1C_1 + A_2C_1 + A_3C_1)} +$$

$$\overset{\text{AxD}}{(A_1D_1 + A_1D_2 + A_1D_3 + A_2D_1 + A_2D_2 + A_2D_3 + A_2D_4 + A_3D_1 + A_3D_2 + A_3D_3 + A_3D_4)} + \overset{\text{BxC}}{(B_1C_1)} + \overset{\text{BxD}}{(B_1D_1 + B_2D_2 +}$$

$$B_1D_3 + B_1D_4) + \overset{\text{CxD}}{(C_1D_1 + C_1D_2 + C_1D_3 + C_1D_4)} + \overset{\text{AxBxC}}{(A_1B_1C_1 + A_2B_1C_1 + A_3B_1C_1)} + (A_1B_1D_1 + A_1B_1D_2 +$$

$$\overset{\text{AxBxD}}{A_1B_1D_3 + A_1B_1D_4 + A_2B_1D_1 + A_2B_1D_2 + A_2B_1D_3 + A_2B_1D_4 + A_3B_1D_1 + A_3B_1D_2 + A_3B_1D_3 + A_3B_1D_4)} +$$

$$\overset{\text{AxCxD}}{(A_1C_1D_1 + A_1C_1D_2 + A_1C_1D_3 + A_1C_1D_4 + A_2C_1D_2 + A_2C_1D_3 + A_2C_1D_4 + A_3C_1D_1 + A_3C_1D_2 + A_3C_1D_3} +$$

$$A_3C_1D_4) + \overset{\text{BxCxD}}{(B_1C_1D_1 + B_1C_1D_2 + B_1C_1D_3 + B_1C_1D_4)} + (A_1B_1C_1D_1 + A_1B_1C_1D_2 + A_1B_1C_1D_3 + A_1B_1C_1D_4 +$$

$$\overset{\text{AxBxCxD}}{A_2B_1C_1D_1 + A_2B_1C_1D_2 + A_2B_1C_1D_3 + A_2B_1C_1D_4 + A_3B_1C_1D_1 + A_3B_1C_1D_2 + A_3B_1C_1D_3 + A_3B_1C_1D_4)} +$$

$$R_{80} + \ldots + R_{116} + E$$

	Df
A	3
B	1
C	1
D	4
AxB	3
AxC	3
AxD	12
BxC	1
BxD	4
CxD	4
AxBxC	3
AxBxD	12
AxCxD	12
BxCxD	4
AxBxCxD	12

TABLE 2

A_1																			
B_1										B_2									
C_1					C_2					C_1					C_2				
D_1	D_2	D_3	D_4	D_5	D_1	D_2	D_3	D_4	D_5	D_1	D_2	D_3	D_4	D_5	D_1	D_2	D_3	D_4	D_5

A_2																			
B_1										B_2									
C_1					C_2					C_1					C_2				
D_1	D_2	D_3	D_4	D_5	D_1	D_2	D_3	D_4	D_5	D_1	D_2	D_3	D_4	D_5	D_1	D_2	D_3	D_4	D_5

A_3																			
B_1										B_2									
C_1					C_2					C_1					C_2				
D_1	D_2	D_3	D_4	D_5	D_1	D_2	D_3	D_4	D_5	D_1	D_2	D_3	D_4	D_5	D_1	D_2	D_3	D_4	D_5

A_4																			
B_1										B_2									
C_1					C_2					C_1					C_2				
D_1	D_2	D_3	D_4	D_5	D_1	D_2	D_3	D_4	D_5	D_1	D_2	D_3	D_4	D_5	D_1	D_2	D_3	D_4	D_5

Results and Discussion

One question we were interested in was whether each of the main effects accounted for a significant amount of variance independent of the other variables. One way this was tested (due to disproportionality of cells) was to use the fitting constants method and the covariance method. Presented below are three examples and the models used in contrasting the results obtained using each method:

1. Response: Neonatal Grooming

(A) Main Effect (Crowding)

Result:

Covariance--Significant Fitting Constant--Non-Significant

Covariance:

Full Model

$$\text{Model } 1 = b_0U + b_1X_1 + \ldots + b_{116}X_{116} + E$$

Restricted Model Restriction = $(b_1 = b_2 = b_3)$

$$Y_1 = b_0U + b_4X_4 + \ldots + b_{116}X_{116} + E$$

$F = 8.29$ df = 3,661 $R^2_F = .509$ $R^2_r = .490$ $P = .00002$

Fitting Constants:

Full Model

$$\text{Model } 2 = b_0U + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 + b_9X_9 + b_{80}X_{80} + \ldots + b_{116}X_{116} + E$$

Restricted Model

Restriction = $(b_1 = b_2 = b_3)$

$$\text{Model } 3 = b_0U + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 + b_9X_9 + b_{80}X_{80} + \ldots + b_{116}X_{116} + E$$

$F = .0015$ $\quad$ $df = 3,730$ $\quad$ $R^2_F = .326$ $\quad$ $R^2_r = .326$ $\quad$ $P = .983$

This may be the case since the covariance method produced a significantly lower error term while the covariates may not be highly correlated with crowding.

2. Response: Self Grooming

(B) Main Effect (A.M. vs. P.M.)

Result:

Fitting Constants--Significant $\quad$ Covariance--Non-Significant

Covariance:

Full Model

$$\text{Model } 4 = b_0U + b_1X_1 + \ldots + b_{116}X_{116} + E$$

Restricted Model

Restriction = $(b_4 = 0)$

$$\text{Model } 5 = b_0U + b_1X_1 + \ldots + b_3X_3 + b_5X_5 + \ldots + b_{116}X_{116} + E$$

$F = .787$ $\quad$ $df = 1,661$ $\quad$ $R^2_F = .421$ $\quad$ $R^2_r = .420$ $\quad$ $P = .376$

Fitting Constants:

Full Model

$$\text{Model } 6 = b_0U + b_1X_1 + \ldots + b_9X_9 + b_{80}X_{80} + \ldots + b_{116}X_{116} + E$$

Restricted Model Restriction = ($b_4 = 0$)

Model 7 = $b_0U + b_1X_1 + \ldots + b_3X_3 + b_5X_5 + \ldots + b_9X_9 + b_{80}X_{80} + \ldots + b_{116}X_{116} + E$

$F = 424.34$ $df = 1{,}730$ $R^2_F = .608$ $R^2_r = .380$ $P = 0.0$

One possible explanation for the nonsignificant result using the covariance method might be the correlation between the other covariates and this main effect. An examination of the correlation matrix by the E's showed this to be the case.

3. Response: Drinking Behavior

(A) Main Effect (Crowding)

Result:

Fitting Constant--Non-Significant Covariance--Non-Significant

Covariance:

Full Model

Same as Neonatal Grooming

Restricted Model

Same as Neonatal Grooming

$F = .119$ $df = 3{,}661$ $R^2_F = .4368$ $R^2_r = .4365$ $P = .950$

Fitting Constants:

$F = .013$ $df = 3{,}730$ $R^2_F = .29878$ $R^2_r = .29875$ $P = .985$

References

Applebaum, M. I. and Cramer, E. M. Some problems in the nonorthogonal analysis of variance. Psychological Bulletin, 1974, 81, 335-343.

Cohen, J. Multiple regression as a general data-analytic system. Psychological Bulletin, 1968, 70, 426-443.

Edwards, A. L. Statistical methods, (Second Edition). New York: Holt, Rinehart & Winston, 1967.

Jennings, E. Fixed effects analysis of variance by regression analysis of experimental data. Psychological Bulletin, 1969, 72, 311-322.

Overall, J. E. and Spiegel, D. K. Concerning least squares analysis of experimental data. Psychological Bulletin, 1969, 72, 311-322.

Scheffe, H. The analysis of variance. New York: John Wiley & Sons, 1959.

Williams, J. D. Two way fixed effects analysis of variance with disproportionate cell frequencies. Multivariate Behavioral Research, 1972, 7, 67-83.

Williams, J. D. Four-way disproportionate hierarchial models. Multiple Linear Regression Viewpoints, 1974, 5, #2.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 5, No. 2, 1974

ABSTRACTS OF VOLUMES 1-4 OF MULTIPLE LINEAR REGRESSION VIEWPOINTS

Keith A. McNeil

Educational Monitoring Systems
3449 Rentz, Ann Arbor, Michigan

ABSTRACT

Early issues of "Viewpoints" are not on all bookshelves (including the author's until John Williams who was the first editor provided copies). Many of the "Viewpoints" have found their way to such texts as Ward and Jennings, Williams, and McNeil, Kelly and McNeil. Others lay dormant. An updated address for all authors appears at the end of this article.

*The publishing of this paper has been paid for by the <u>Multiple Linear Regression Viewpoints</u> Special Interest group.

McNeil, K. A. The negative aspects of the eta coefficient as an index of curvilinearity. Viewpoints, 1970, 1, 7-17.

The objective of this paper is to discuss the eta coefficient and to point out some limitations and misconceptions about the coefficient. Specifically, we will discuss the fact that; 1) the eta coefficient is a global measure of curvilinearity; 2) the eta coefficient has limited interpretability; 3) there are a number of other curvilinear relationships that might be of more significance and of more interpretability; 4) these other curvilinear relationships do not suggest nor encourage grouping of data as does the eta coefficient; and 5) these other curvilinear relationships may tend to be more amenable to replication than is the eta coefficient.

Stone, L. A. and Skurdal, M. A. Estimation of product moment correlation coefficients through the use of the ratio of contingency coefficient to the maximal contingency coefficient. Viewpoints, 1970, 1, 19-25.

The authors attempt to demonstrate that the ratio, contingency coefficient/maximal contingency coefficient (C/C_{max}), is directly comparable to the product moment correlation coefficient. Inspection of all of the computed C/C_{max} ratios, from 2 X 2 tables, showed that the ratios which correspond most closely to the product moment correlation coefficients were not always the ones which were associated with fourfold tables having dichotomies nearer to .50 - .50 proportions. However, the authors were lead to believe tha the C/C_{max} ratios which best approximated the product moment correlation coefficients generally were from the fourfold tables where there was an approximate 50 - 50 split. The C/C_{max} ratio may be used as a "quick and dirty" estimate of the relationship measure provided by the product moment correlation model. No mathematical justification is offered for this contingency coefficient ratio.

Williams, J. D. Multiple comparisons in a regression framework. *Viewpoints*, 1970, 1, 26-39.

In using multiple regression as a problem solving technique, one problem that might arise is the overuse of a full model with several restricted models, without adjusting the probability level. Such an approach would violate the apparent probability level. This has long been a concern in statistics. Several multiple comparison procedures have been developed for different situations.

The intent of the present paper has been to extend some of the better known multiple comparison procedures (Duncan's Multiple Range Test, Dunn's "C" Test, and Scheffe's Test) to a multiple regression approach. The major change in the regression approach is to assess the result of multiple uses of a full model to a correct distribution, rather than a straight-forward usage of the F distribution.

Goff, A. F. and Houston, S. R. Concurrent Validity of the Koppitz scoring system for the bender visual motor gestalt test. Viewpoints, 1971, 1, 45-52.

The study examined correlations between assessed visual-motor perception, intelligence, and academic achievement. In addition, efficiency of prediction for criterion variables was investigated by employing two approaches of analysis: (a) regression model and (b) Bender *z* model. The following conclusions were formulated on the basis of the obtained data and from the comparison of the two predictive models.

(1) The significant negative correlation found between age and the Bender error score adds further substantiation to the fact that the ability to correctly execute the Bender protocol improves with increased age.

(2) The Bender *z* score correlated to a greater degree with intelligence, reading, and arithmetic achievement than did the Bender error score with the three specified variables.

(3) The obtained correlations of the Bender *z* score with the three criterion variables agrees with the literature in directionality and in significance with assessed intelligence.

(4) However, efficiency is enhanced by using the Bender error score and age rather than the single variable of the Bender *z* score to predict achievement in reading and arithmetic and assessed intelligence.

Jordan, T. E. Curvilinearity within early developmental variables.

Viewpoints, 1971, 1, 53-77.

Squared and cubed vectors were introduced into eighteen regression models each applied to nine criteria. Data came from study of several hundred children in the first three years of life. Departure from linearity did not provide better accounts of the relationship between five predictors and development at 12, 24, and 36 months of age. Illustrations of various patterns of squared vectors and cubed vectors were presented from the data.

Reed, C. L., Feldhusen, J. F., and Van Mondfrans, A. P. Regression models in educational research. *Viewpoints*, 1971, 1, 78-88.

Results of this study do not support the findings of Rock (1965) that the interaction term regression was superior to the quadratic form in predictive efficiency. The most efficient regression model will depend upon: 1. how the variables and criterion are related, 2. the reliability of the predictor variables, and 3. the research question asked.

The studies reviewed in this paper seem to indicate that complex regression models are in some cases more efficient predictors of complex behavior than the most frequently assumed first order model. When quadratic and interaction terms are significant, however, interpretation is made more difficult. Still, an attempt at interpretation seems somewhat better than ignoring the problem or assuming it does not exist.

McNeil, K. A. and Beggs, D. L. Directional hypotheses with the multiple linear regression approach. *Viewpoints*, 1971, 1,

89-102.

Two well known directional tests of significance are presented within the multiple linear regression framework. Adjustments on the computed probability level are indicated. The case for a directional interaction research hypothesis is defended. Conservative adjustments on the computed probability level are offered and a more precise computation is requested of statisticians. Emphasis is placed more on the research question being asked than on blind adherence to conventional formulae.

McNeil, K. A. On the unit vector. Viewpoints, 1971, 2, 2.

Short article presents several interpretations of the unit vector. One of which conceptulizes the unit vector as any predictor to the zero power.

McNeil, K. A. On attenuating a multiple R. Viewpoints, 1971, 2, 3.

A brief conceptual arguement questioning the value of attenuating a multiple R.

Duff, W. L., Jr., Houston, S. R. and Bloom, S. A regression/principal components analysis of school outputs. Viewpoints, 1971, 2, 5-18.

The objective of this study is to identify the correlates of student performance and teacher retention in an inner-city elementary school district. The purpose is to provide urban school administrators with information necessary to cope with the special problems they face in organizing and administering their educational resources.

The study is divided into two parts: a descriptive section and an analytic section. In the descriptive section the writers are concerned with describing the inner urban school system. Here the data to be analyzed are presented and classical regression techniques are used to specify the three basic teacher retention and student performance models. In the second section the data are further analyzed in terms of the unique contribution of a priori specified subsets of predictor variables. This section ends with a comparison of a principal component regression approach to the a priori grouping of predictors used in the unique analysis.

Bolding, J. T. Empirical exercises for the study of multiple regression. Viewpoints, 1971, 2, 21-22.

Six short exercises are presented which utilize random data and illustrate several properties of MLR. Specifically, if ten random variables are added together to produce a criterion, then the 10 variables will yield an R^2 of 1.00 when predicting that criterion. Variations of the above situation are presented.

Connett, W. A note on multiple comparisons. Viewpoints, 1971, 2, 23-24.

Two general methods are available for controlling the overall alpha level. One method is to adjust the F value required for significance. This method in relation to multiple regression was recently discussed in Viewpoints (Williams, 1970). The second way to maintain an overall alpha level is by proper choice of the alpha levels for the individual comparisons. The purpose of this note is to review a method for determining the individual alpha levels necessary to maintain some selected overall alpha level.

Williams, J. D. and Lindem, A. C. Setwise regression analysis-A new data-analytic tool. Viewpoints, 1971, 2, 25-27.

The setwise procedure drops one set at a time in a stepwise fashion. There will be as many steps as there are sets. Statistically, the steps are accomplished by an iterative procedure that allows the R^2 term to be maximized at each stage in a backward stepwise procedure. Once a set is discarded, the set is no longer considered at later stages.

While the difficulty regarding the use of binary coded predictors has been at least partially solved, other difficulties in regard to the stepwise procedure are also involved in the setwise procedure; additionally, the setwise procedure has a new problem unique to itself.

It has been pointed out several times that probability levels in the stepwise procedure are usually violated. Further, when k of the N variables have been dropped, the N - k remaining variables are not necessarily the set of N - k variables that would yield the highest R^2 value. These criticisms would also be valid in regard to the setwise procedure. Additionally, the differences in the number of variables in a set will have some effect upon when that set of variables would be dropped. Other things being equal, a set with 6 variables will be retained longer than a set with 3 variables. Notwithstanding these difficulties, if the setwise procedure is judiciously employed by researchers, then addition data analysis power can be obtained.

The program and sample printout are available on request.

Greenup, H. Watch that first step. Viewpoints, 1972, 2, 32-33.

Author discusses the interactive system at UCLA. "Biomed" programs can be directed by (a) the 2250's typewriter - like keyboard, (b) an attached lightpen, or (c) some specially programmed keys located near the keyboard.

McClaran, R. V. and Brookshire, W. K. A comment on multiple comparisons in a regression framework. Viewpoints, 1972, 2, 34-35.

Authors clarify an error in an earlier article in Viewpoints. The error occurred in the Williams paper dealing with multiple comparisons. One restricted model was inaccurate.

Newman, I. and Fry, J. A response to "a note on multiple comparisons and a comment on shrinkage". Viewpoints, 1972, 2, 36-38.

Connett (1971) presented Kimball's (1951) formula for keeping α levels constant when making a number of comparisons.

A simplified formula, developed by the authors, for computing these α levels is presented along with a brief mathematical proof that Kimball's technique and that of the authors' are approximately equivalent.

The method simply takes the desired α level desired to keep constant across a number of comparisons (α_0) and divides that specific α_0 by the number of comparisons one wishes to make (N). Authors suggest that when the ratio of subjects to variables is 10:1 or less, a shrinkage estimation should be used and reported.

If this is done, research based on multiple correlations will tend to be more replicable and therefore more desirable and useful.

Newman, I. A suggested format for the presentation of multiple linear regression. Viewpoints, 1972, 2, 42-45.

Multiple regression, when presented in the literature, has usually been formated in an idiosyncratic manner. In addition, the format rarely presents all the relevant information in a concise, easy to inspect manner. Instead, one tends to find himself thumbing through pages of the article to find the relevant information.

Author presents a format for the presentation of multiple regression models and the information required for their interpretation.

Connett, W. E., Houston, S. R. and Shaw, D. G. The use of factor regression in data analysis. Viewpoints, 1972, 2, 46-49.

Suppose that it is desired to express the criterion variable as a function of a set of independent variables in which the intercorrelations between the various independent variables is near zero. The procedure involves restructuring the full regression model in such a way that the criterion variable is expressed as a function of several mutually orthogonal factor variables. This principal components-regression approach permits one then to investigate the unique contribution of each of the factor variables to explaining the dependent variable.

The "factor regression" procedure begins with the complete orthogonal factoring of the set of predictors.

Factor scores are then computed.

If a regression model is cast, regressing some criterion variable onto the set of factor score predictors, several interesting properties are noted. The beta weight for a predictor is equal to the validity for that predictor. The R^2 value for any model is equal to the sum of the squares of the beta weights for the model. The exclusion of a factor score variable from the predictor set will result in a drop in R^2 equal to the square of the beta weight for the variable dropped. And, perhaps most important, the dropping of any predictor variable from the predictor set will not affect the beta weights (predicitve contribution in this case) of any of the other variables. These properties are demonstrated in the following example.

This procedure is not a substitute for the long established principals of statistical inference, those being hypothesis building and testing; rather, it provides data organization preliminary to hypothesis development and testing.

Williams, J. D., Maresh, R. T., and Peebles, J. D. A comparison of raw gain scores, residual gain scores, and the analysis of covariance with two modes of teaching reading. Viewpoints, 1972, 3, 2-16.

It should be abudantly clear from the 16 tables that the three approaches to psycho-educational change are different. While this set of data does not exhibit strong relationships between the dichotomous predictor and the various criteria, the use of the statistical significance approach would occasionally yield different interpretations. Perhaps the most objective comparison between the three measures would be the R^2 term (for the analysis of covariance, or adjusted means approach, $R^2_{FM} - R^2_{RM}$). Only one significant difference is found in the three measures. In Table 6, the raw gain is significant ($p < .05$), <u>but</u>, under exactly the conditions that would tend to make this occur, the vertical group was significantly smaller than the graded group on the pre-test, but this difference was almost erased on the post-test. In terms of the raw gains score, this produced a significant difference in favor of the vertical group.

Jennings, E. Linear models underlying the analysis of covariance, residual gain scores and raw gain scores. <u>Viewpoints</u>, 1972, 3, 17-24.

The problem of investigating "change" or "gain" that can be attributed to "treatments" has been discussed extensively over a number of years without a noticable concensus emerging. Cronbach (1970) has even suggested that many questions that appear to involve "change" can be effectively resolved without reference to the concept of change.

Koplyay, J. B. Automatic interaction detector AID-4. <u>Viewpoints</u>, 1972, 3, 25-38.

The primary value of AID-4 to the task scientist is its ability to identify the maximum amount of variance in the criterion which can be accounted for by the predictors available; it relieves the task scientist of the trial-and-error task of attempting to identify the various relevant combinations of linear and non-linear interaction terms presently required by the multiple linear regression technique. The splitting process of AID-4, being based upon maximizing the between sums-of-squares and minimizing the within-sums-of-squares, automatically takes all present interactions into account, indicating the maximum variance predictable in the criterion from the predictors. The interactions and patterns or trends are identifiable from the AID-4 output.

Newman, I. Some further considerations of using factor regression analysis. Viewpoints, 1972, 3, 39-41.

Presents four major points concerning factoring the data before using MLR. When using factor regression procedures, it is important to keep in mind that if one does not use all of the factors (that is, accounting for 100% of the trace) he may be overlooking a suppressor factor (suppressor variable).

If one is interested in improving the multiple regression equation by using factor techniques, there is only one way this can be done--by using fewer factors than the number of original variables. This will increase the df and also possibly decrease shrinkage-estimates. However, when this is done one may be losing information that can account for criterion variance by eliminating a factor that accounts for very little trace of the factored matrix but is highly correlated with the criterion scores.

Using only the factors that account for most of the trace should be avoided when the predictor variables that are being factored are likely to be highly reliable. Under these conditions a variable that accounts for little of the trace variance may be a good and highly reliable predictor of criterion variance.

Newman, I., and Fry, J. Proof that the degrees of freedom for the traditional method of calculating analysis of covariance and the multiple regression method are exactly the same. Viewpoints, 1972, 3, 42-45.

Authors provide a proof that the degrees of freedom for traditional analysis of covariance and for the "over and above" analysis with MLR are exactly the same.

Brebner, M. A. Conditions for no second order interaction in multiple linear regression models for three factor analysis of variance. Viewpoints, 1972, 3, 46-57.

Author provides pictorial, algebraic, and regression models for investigating second order interaction in a three way design.

Newman, I., and McNeil, K. A note on the independent variance of each criterion in a set. Viewpoints, 1972, 3, 58-60.

Frequently a researcher is interested in a number of criterion variables which may not be uncorrelated with each other. The chances are that these variables are likely to be significantly correlated with each other.

If one is interested in accounting for the independent piece of a criterion variable's variance, then the procedures outlined here should be used. By covarying other nonindependent criteria as explained in the paper, one can actually treat each test of each criterion as if the criterion variable were independent of each other.

Newman, I., Lewis, E. L., and McNeil, K. A. Multiple linear regression models which more closely reflect bayesian concerns. Viewpoints, 1972, 3, 71-77.

The purpose of this paper is to discuss several relationships between the Bayesian approach and the multiple linear regression approach.

The testing of interaction is what distinguishes the regression procedure outlined in the present paper from that commonly used in prediction studies. Too often, even plausible interactions are ignored and all subjects are lumped together and, hence, treated as similar. Our conceptual theories have long ago turned to distinct groupings, and it is about time that our statistical procedures reflect this empirical possibility, whether they be Bayesian or multiple linear regression. Until the Bayesian methodology has empirically been shown to be more predictive than multiple regression analysis, the availability and relative mathematical simplicity of multiple regression analysis would seem to indicate preference for its utilization rather than the Bayesian approach.

Brookshire, W. K., and Bolding, J. T. Using coefficients of orthogonal polynomials as predictor variables in multiple regression. Viewpoints, 1973, 4, 1-6.

Coefficients of Orthogonal Polynomials are presented by some authors (Snedecor and Cochran) as a means of simplifying the computation required in trend analysis. Linear regression addicts who are computer oriented can still make good use of such coding in the analysis of complicated designs.

Consider a two factor design where the factors are assumed to be quantitative with levels selected at equal intervals. Testing for main effects and trend analysis can both be simplified by the use of coefficients of orthogonal polynomials as predictor vectors.

McNeil, K. A. Testing an hypothesis about a single polulation mean with multiple linear regression. Viewpoints, 1973, 4 7-14.

The recent emphasis on criterion referenced testing and on the explicit stating of objectives implies that more researchers will be testing hypotheses about a simple population mean. The generalized regression procedure is one way to test such an hypothesis. The appropriate regression models are presented in this paper.

Halasa, O. Identification of significant predictors of children's achievements and attendance. Viewpoints, 1973, 4, 15-22.

The identification of variables other than the treatment process, which are affecting the criterion measure variance has always been a problem.

Most of the regression coefficients failed a statistical test of significance. Of the six predictors, the pre-test score evidenced consistent significant contributions to the criterion variance. The per cent of predictable variance, however, indicates that a significant proportion of the variance remains unaccounted for.

Newman, I., and McNeil, K. Application of multiple regression analysis in investigating the relationship between the three components of attitude in Rosenberg and Hovland's theory for predicating a particular behavior. Viewpoints, 1973, 4, 23-39.

Multiple regression and factor analysis techniques were used to investigate the relationship between the components of attitude and their differential predictive power. It was found that the different components of attitude and their linear interaction are more likely to be predictive for intimate rather than non-intimate behaviors. The cognitive component was found to be significantly predictive of intimate behavior but not predictive for non-intimate behavior. Out of the three measures used, the behavioral differential was the most predictive scale for both intimate and non-intimate behavior.

Ward, J. H., Jr. Guidelines for reporting regression analyses. Viewpoints, 1973, 4, 40-41.

Suggests including in a report of regression analyses:

1. General Comments
2. Regression Analysis Discussion

a. natural language statements of the hypotheses

b. identification of the assumed model

c. hypotheses in terms of assumed model

d. identification of the restricted model

e. results of the test

3. Vector definitions

4. Analyses

a. model specifications

b. model comparisons

5. Regression Computer Output

McNeil, K. Reaction to Ward's "Guidelines for reporting regression analysis" and some alternatives. Viewpoints, 1973, 4, 42-44.

Suggests guidelines other than Ward's:

1. Statement of research hypothesis
2. Statement of statistical hypothesis
3. Statement of alpha
4. Formulation of full model
5. Statement of restrictions
6. Formulation of restricted model
7. Definition of vectors
8. Reporting of the probability of the calculated F, and the subsequent decision.

Newman, I. A revised "Suggested format for the presentation of multiple regression analyses". Viewpoints, 1973, 4, 45-47.

Presents a format for displaying the hypotheses and results for those hypotheses. This format could be used in conjunction with either Ward's or McNeil's earlier stated formats.

Williams, J. D. Applications of setwise regression analysis. Viewpoints, 1973, 4, 1-7.

One of the earlier applications of the setwise technique was made by Grooters (1971). Grooters was interested in predicting costs per student credit hour in four state colleges. The input data were means by department for 16 variables, forming nine sets. Four of the sets were single variables, four sets were logical sets and one set was formed among mutually exclusive binary sets similar to the religious set described earlier. Such a situation typically involves a linear dependency within the total set. To remove the dependency, any one of the variables within the set can be excluded, and the analysis can be performed. Using the setwise technique, Grooters was able to isolate a rather intriguing result; student costs are in some measure higher in departments that have a higher incidence of outside of school professional activity (consulting, speaking, local community work, artistic endeavors outside the college setting, etc.). Yet another interesting result was that the average salary paid per member in the department was the first set to drop out.

Cummins, M. E. Utilization of multiple regression analysis in changing the verbal behavior patterns of elementary classroom teachers through self-evaluation. Viewpoints, 1973, 4, 8-17.

Multiple regression analysis was used to investigate the effectiveness of a self-evaluation technique in changing the verbal behavior patterns of elementary classroom teachers. Three ratios were employed: the I/D ratio, or ratio of indirect to direct statements by the teacher; the TRR, or teacher response ratio which eliminates questioning and lecturing from the total I/D ratio; and the PIR, or the ratio of pupil talk-response to pupil talk-initiated. The self-evaluation technique for changing the verbal behavior patterns of elementary classroom teachers was not found to be significant for any of the three measures.

Pohlmann, J. T. Incorporating cost information into the selection of variables in multiple regression analysis. Viewpoints, 1973, 4, 18-26.

The problem of finding the best regression equation is considered from the standpoint of predictor costs. Typically, variables are selected for inclusion in prediction equations on the basis of their unique contribution to the prediction of a criterion. A method is presented whereby losses due to lack of predictability and predictor costs are combined in a loss function. The best predictor set is then chosen that simultaneously minimizes losses incurred in measuring the predictors and losses incurred from lack of predictability of a criterion variable.

Lewis, E. L., and Mouw, J. T. The use of contrast coding to simplify anova and ancova procedures in multiple linear

regression. Viewpoints, 1973, 4, 27-44.

Multiple regression (MR) is a powerful and flexible technique for handling data analysis. The present paper presents a discussion of the use of "contrast coding" in performing analysis of variance and analysis of covariance procedures in MR. Contrast coding provides a method for coding nominal variable in the set of predictor vectors in MR so that such vectors reflect a set of orthogonal comparisons. As a result, one is able to test hypotheses concerning more specific research questions than those usually tested in more traditional MR coding procedures. By adding more components to the general linear model, contrast coding provides a relatively simple and logical basis for extending analysis of variance to its various subclassifications.

Newman, I. Variations between shrinkage estimation formulas and the appropriateness of their interpretation. Viewpoints, 1973, 4, 45-48.

This is a discussion paper dealing with the use of shrinkage, different methods for estimating shrinkage, and the accuracy of shrinkage estimates when variables are preselected as in stepwise regression and when variables are not preselected.

Pyle, T.W. The analysis of split-plot and simple hierarchical designs using multiple linear regression. Viewpoints, 1973, 4, 1-9.

For those researchers who are interested in using multiple linear regression (MLR) as a general data analytic technique, it is of some importance to recognize what types of models and re-

strictions on models generate F ratios that correspond to classical hypothesis testing procedures. One of these is the split-plot factorial design and another is the completely randomized hierarchical design. What will be demonstrated here is that MLR models can be constructed for these designs which generate F ratios that are equivalent to those obtained by traditional computing formulas.

Buzahora, R.C. and Williams, J.D. An empirical comparison of residual gain analysis and the analysis of covariance. Viewpoints, 1973, 4, 10-17.

Summary - An extensive comparison of the analysis of covariance and the residual gain analysis was made. Using subtests of the Iowa Tests of Basic Skills, the School Attitude Inventory and School Sentiment Index independently at each grade level, grades 3-8, 258 analyses were compared. While the analysis of covariance gave some indication of being more powerful, this result was not uniform over the various analyses.

Houston, J.A. and Houston, S.R. Judgement analysis and pornography. Viewpoints, 1973, 4, 18-31.

Judgement Analysis (JAN) was used as a methodology for determining what is pornographic by testing this technique with three groups concerned with this issue. These groups included doctoral students majoring in Psychology, Counseling and Guidance (PCG) at the University of Northern Colorado, lawyers and police officers from the city of Greeley, Colorado. JAN proved to be an effective technique in the identification of policies.

The problem of what is pornographic is indeed a complex one as evidenced by the many specific categorical and complex policies present in the PCG judges, lawyers and police officers.

Houston, S. R. Identifying faculty policies of teaching effectiveness. Viewpoints, 1973, 4, 32-38.

Judgement Analysis (JAN) was employed to capture the teacher effectiveness policy (ies) of College of Education faculty at the University of Northern Colorado (UNC). Fifty-seven judges evaluated 60 hypothetical faculty members, each on four characteristics. Results indicated that possibly three different judgemental systems or policies existed.

Brebner, M. A. Multiple linear regression models for analysis of covariance including test for homogeneity of regression slope Part 1- oneway designs with one covariate. Viewpoints, 1973, 4, 39-50.

The objective of the paper is to describe the multiple linear regression models which correspond exactly in all respects to the classical analysis of covariance and tests for homogeneity of regression slope. It is the author's opinion that the approach pioneered by Bottenberg and Ward gives a much clearer and more meaningful understanding of analysis of covariance than the classical method.

Pohlman, J. T. and Newman, I. Regression effects when the assumption of rectilinearity is not tenable. Viewpoints, 1973, 4, 51-59.

When analyzing data which deals with repeated testing, one may find that extreme scorers on a pretest regress away from the mean upon post test, contrary to what one would expect from the regression effect. This paper discusses regression effects and presents the argument that when contrary results occur, they are indicative of violation of an underlying assumption of rectilinearity for the Pearson r. Therefore it is recommended that one should look for nonlinear relationships when interpreting such data. In addition, three methods for determining if nonlinear relationships exist in data, are suggested and briefly discussed.

Olson, G.H. A general least squares approach to the analysis of repeated measures designs. Viewpoints, 1973, 4, 60-69.

Two procedures for computing general least squares analyses of repeated measures designs are discussed. The first procedure, appropriate for small N, is a straight-forward application of the usual regression approach to the analysis of variance. The second approach, appropriate for large N, also utilizes the regression approach but requires some minor calculation in addition to that typically performed by most computer programs.

Williams, J. D. A note on contrast coding vs. dummy coding. Viewpoints, 1973, 4, 1-5.

Abstract--A comparison is made between the contrast coding system for solution to the analysis of variance design presented by Lewis and Mouw (1973), and the use of dummy coding for solution to the analysis of variance designs. Some of the limitations and advantages of each approach are given.

Klein, M. and Newman, I. Estimated parameters of three shrinkage estimate formuli. Viewpoints, 1973, 4, 6-11.

This paper examines the shrinkage formuli of Wherry, McNemar and Lord in relation to overcorrection. A table is given which shows the number of times that each formula resulted in a negative value of R^2 for different numbers of variables and sample sizes.

McNeil, K. and McShane, M. Complexity in behavioral research as viewed within the multiple linear regression approach. Viewpoints, 1973, 4, 12-15.

The paper attempts to clarify the notion of complexity in research. Perhaps, in the past, researchers have over simplified the variables and their interrelationships. Thinking of interaction variance as bad variance or as producing an undesired result, or lamenting the complexity of the phenomena under consideration does little or nothing to advance research in the behavioral sciences. The authors examine two views of complexity which seem to exist: (1) complexity as indicated by the number of predictor variables needed to account for a criterion behavior, and (2) complexity as indicated by the nature of the predictor variables. The authors argue that the second view of complexity is not valid, and that the multiple linear regression technique provides an easy way to index the first view.

Wyshak, G. Modification of multiple regression when an independent variable is subtracted from the dependent variable. Viewpoints, 1974, 4, 16-20.

Behavioral scientists are often concerned with regressing some dependent or outcome variable, Y, on a number of independent or explanatory variables, X_i, i = 1,2,...,k. (Model 1). If Y is a final score or measurement and X_1 an initial score, the investigator may be interested in some measure of change, say $Y - X_1$, and its relation to the several explanatory variables including X_1. (Model 2). Analyses would be based on two multiple regression equations, one relating to the regression of Y on X_1, $X_2,\ldots,X_k$; and the other to the regression of $(Y - X_1)$ on the same X's.

The note calls attention to the fact that one analysis would suffice for the two models because the regression coefficients, the total sums of squares, deviations sums of squares and regression sums of squares are readily obtained for Model 2 once the calculations have been made under Model 1.

St. Pierre, R. G. Vargen: A multiple regression teaching program. Viewpoints, 1974, 4, 21-29.

VARGEN (Variable Generator) creates sets of data with known statistical properties by generating user-specified variables which are functions of uniformly distributed random numbers for each of a group of subjects. The user specifies the relative size and location of from one to nine predictor variables and one criterion variable within a ten by ten matrix (hereafter called the "universe") and, therefore, the amount of variance accounted for by each variable. A visual display is produced showing the size and location within the universe of any five variables.

Williams, J. D., and Lindem, A. C. Regression computer programs for setwise regression and three related analysis variance techniques. Viewpoints, 1974, 4, 30-46.

Four computer programs using the general purpose multiple linear regression program have been developed. Setwise regression analysis is a stepwise procedure for sets of variables; there will be as many steps as there are sets. COVARMLT allows a solution to the analysis of covariance design with multiple covariates. A third program has three solutions to the two-way disproportionate analysis of variance: (a) the method of fitting constants, (b) the hierarchical model and (c) the unadjusted main effects solution. The fourth program yields three solutions to the two-way analysis of covariance, with or without proportionality, and with multiple covariates. The three solutions are similar to those described for a two-way analysis of variance with disproportionate cell frequencies.

Mason, R. L., and McNeil, K. A. Mashit - for ease in regression program communication. Viewpoints, 1974, 4, 47-64.

This regression system is an intermediate result of a project to develop a comprehensive regression computer system as a foundation for a complete statistical man-machine interface. The outstanding features of the system can be condensed into two principle concepts. First, the program dynamically allocates core resulting in no limit on title cards, question cards, etc. Secondly, "English type" user commands are used in a free format mode to save computer instruction time. The resulting system is two phase constructed in such a manner that additional capabilities can be added efficiently.

Fanning, F. W., and Newman, I. The development and demonstration of multiple regression models for operant conditioning questions. Viewpoints, 1974, 4, 65-87.

Based on the assumption that inferential statistics can make the operant conditioner more sensitive to possible significant relationships, regression models were developed to test the statistical significance between slopes and Y intercepts of the experimental and control group subjects. These results were then compared to the traditional operant conditioning eyeball technique analysis.

One major advantage of using the regression procedure, rather than the traditional eyeball technique is that probability estimates can be attributed to the accuracy of the statements.

Another advantage of the regression procedure used is the ability to test the curvilinear relationships above and beyond linear ones, which is not feasible with the eyeball technique on multiple baseline analysis. Similarily, one cannot test to see if the slopes of the control group are significantly different statistically.

In addition, as demonstrated in this paper we can also test to see if the functional relationship of one treatment is significantly different from the functional relationship of some other treatment (across some area of interest).

Haynes, J. R. and Swanson, R. G. Method for comparison of non-independent multiple correlations. Viewpoints, 1974, 5, 1-6.

The technique for the comparison of related *R*s is based on the residual criterion scores rather than directly on the *R* or R^2 values. The procedural steps are: (a) generate predicted criterion scores for each set of predictor variables; (b) for each set of predicted scores, obtain the absolute difference between each predicted and corresponding actual criterion scores; (c) analyze these difference scores by a single classification analysis of variance for repeated measures. If a significant *F* is obtained from the ANOVA, some or all of the multiple correlations would be considered different. Multiple comparisons to determine which *R*s differ from each other are then conducted using techniques such as the Newman-Keuls studentized range statistical test.

Maola, J. F., The multiple regression approach for analyzing differential treatment effects - the reverse gestalt model. Viewpoints, 1974, 5, 7-10.

This paper was written in order to demonstrate a method of using multiple regression for determining the effect of independent differences of providing treatment. The procedures involve a model of comparing differential treatment effects to the total treatment and was, therefore, named the "Reversed Gestalt Model" because of its theoretical base.

Byrne, J. The use of regression equations to demonstrate caus-

ality. Viewpoints, 1974, 5, 11-22.

A universal objective of scientists involved with explanations of behavior or phenomenon is to demonstrate their knowledge of what is causing it. Explanations of causality usually entail knowledge of a number of elements or underlying variates that interrelate to produce the phenomenon.

An R^2 of 1.00 means that 100% of the change that takes place in the phenomenon has been numerically connected to changes that take place among the variables and experimental manipulations controled by the scientist. As the R^2 value approaches 0.0, it means that more and more things outside of the comprehension of the scientist are appearing which cause the phenomenon.

Pyle, T. W. Classical analysis of variance of completely within-subject factorial designs using regression techniques. Viewpoints, 1974, 5, 23-32.

The use of regression analysis in analyzing data obtained from Between-subject experimental designs is well-documented and easily obtainable. In addition, some texts cover the classical analysis of variance (ANOVA) of single-factor repeated measurements designs, and an identification has also been made of regression models which yield *F* ratios (for Main and Interaction effects) equivalent to those obtained by standard computing formulas for split-plot factorial and simple hierarchical designs (Pyle, 1973). However, documentation for the standard ANOVA of completely Within-subject factorial designs using regression analysis is not available as far as this author is aware. The purpose of this article is to provide such documentation by means of an example.

Gloeckler, T. L. Use of multiple linear regression on analysis of intelligence test score changes for visually handicapped adults. Viewpoints, 1974, 5, 33-39.

Multiple linear regression was used to analyze complex data assessing longitudinal changes in IQ test performance of visually handicapped adults. Results indicated: (1) patterns of performance similar to those found in sighted populations, and (2) no influence on IQ changes by a variety of ontological factors.

501-575-2000
X4203

Bolding, James
Ed. Foundations
College of Ed.
Univ. of Arkansas
Fayetteville, AR 72701

Brebner, Michael A.
Dept. of Ed. Psych.
Univ. of Calgary
Calgary 44
Alberta
CANADA

Brookshire, William
North Texas State Univ.
13841
Denton, TX 76203

Buzahora, R.C.
University of North Dakota

Connett, W.
State Dept. of Ed.
and Devel. and peval.
State Capital, MT 59601

Cummins, M.E.
Medina City School District
130 North State Rd.
408
Medina, OH 44256

Duff, W.L.

Fanning, F.W.
Dept. of Special Ed.
Univ. of Akron
Akron, OH 44325

Gloeckler, T.L.
University of Akron
Akron, OH 44325

Goff, Ann
Dept. of Research & Statistical
Methodology
Univ. of Northern Colorado
Greenly, CO 80631

Greenup, H.
5406 Rhea Ave.
Tarzana, CA 91356

Halasa, O.
Div. of Research
and Development
Cleveland Public School
Cleveland, OH

Haynes, J.R.
Dept. of Psychology
Box 13587 NT Station
Denton, TX 76203

Houston, J.A.
Dept. of Research &
Statistical Methodology
Univ. of Northern Colorado
Greenly, CO 80631

Houston, S.R.
Research & Stat.
Methodology
49 Bayshore Dr.
Milford, CT 06460

Jennings, E.
Univ. of Texas Station
7246
Austin, TX 78712

Jordon, T.E.
Behavioral Studies &
Research
Univ. of Missouri
St. Louis, MO 63121

Klein, M.
Univ. of Akron
Akron, OH 44325

Koplyay, Janos B.
Personnel Research Division
Airforce Human Res.
Lackland AFB, TX 78236

Lewis, E.L.
SIU
WHAM
Carbondale, IL 62901

Maola, J.F.
Univ. of Akron
Dept. of Counseling
Akron, OH 44325

Mason, R.L.
2012 Woodmore
Huntsville, AL

McNeil, K.
Dept. of Guidance and Ed.
Pshyc.
3449 Rentz Road
Ann Arbor, MI 48103

Newman, I.
Dept. of Ed. Foundations
Univ. of Akron
Akron, OH 44325

Olson, G.H.
Dallas Ind. School Dist.
3700 Ross Ave.
Dallas, TX 57204

Pohlman, John
Testing Center 618-453-5311
SO Illinois Univ.
Carbondale, IL 62901

Pyle, T.W.
Dept. of Psych.
E. Wash. State College
Cheney, WA 99004

Reed, C.
Rt. 9
72
W. Lafayette, IN

ST. Pierre, R. G.
224 Newtonville Ave.
Newton, MA 02158

Ward, J.H.
Southwest Educ. Laboratory
167 E. Arrowhead Dr.
San Sntonio, TX 78228

Williams, J.D.
Center of Teaching and Learning
Univ. of N. Dakota
Grand Forks. ND 58201

Wyshak, G.
Yale University
Dept. of Epidemiology &
Public Health
Princeton, NJ

IT IS THE POLICY OF THE SIG-MULTIPLE LINEAR REGRESSION AND OF VIEWPOINTS TO CONSIDER FOR PUBLICATION ARTICLES DEALING WITH THE THEORY AND THE APPLICATION OF MULTIPLE LINEAR REGRESSION. MANUSCRIPTS SHOULD BE SUBMITTED TO THE EDITOR AS AN ORIGINAL, DOUBLE-SPACED TYPED COPY. A COST OF $1 PER PAGE SHOULD BE SENT WITH THE SUBMITTED PAPER. REPRINTS ARE AVAILABLE TO THE AUTHORS FROM THE EDITOR. REPRINTS SHOULD BE ORDERED AT THE TIME THE PAPER IS SUBMITTED AND 20 REPRINTS WILL COST $.50 PER PAGE OF MANUSCRIPT. PRICES MAY BE ADJUSTED AS NECESSARY IN THE FUTURE.

THE UNIVERSITY OF AKRON
AKRON, OHIO 44325

AKRON OHIO
U.S. POSTAGE
18
PB METER 337321

SPANER STEVE
BEHAVIORAL STUDIES
UNIV OF MO-ST LOUIS
ST LOUIS MO 63121

BOOKS-SPECIAL

TABLE OF CONTENTS