FEBRUARY 1980

VOLUME 10 NUMBER 2

MULTIPLE LINEAR REGRESSION VIEWPOINTS

A publication of the Special Interest Group on Multiple Linear Regression

MLRV Abstracts appear in CIJE, the ERIC System,
and microform copies are available from
University Microfilms International
MLRV is listed in EBSCO Librarians Handbook.
ISSN 0195-7171

MULTIPLE LINEAR REGRESSION VIEWPOINTS

TABLE OF CONTENTS

MULTIPLE LINEAR REGRESSION VIEWPOINTS
February, 1980 Volume 10, Number 2

ESTIMATION AND TESTING OF POCKET MEANS USING MULTIPLE LINEAR REGRESSION TECHNIQUES

George P. McCabe and Sharron A. S. McCabe
Purdue University

The problem of predicting a continuous criterion variable from two continuous predictors is considered. Stratification on the predictors is one common procedure for construction of subgroups which are easily labeled and discussed. Through the appropriate use of regression techniques, data can be used more efficiently and inferences regarding carefully selected subpopulations, called pockets, can be made. An example using cognitive styles to predict performance on problem solving tasks is discussed.

INTRODUCTION

We consider the problem of relating a criterion variable Y to two predictor variables X_1 and X_2. If X_1 and X_2 are dichotomous variables, analysis of variance techniques are commonly used. Of course, if the numbers of observations for each of the (X_1,X_2) possibilities are widely disparate, particular care must be exercised in choosing the appropriate form for the unbalanced anova. Generally, the results of such an analysis are readily interpretable since one can discuss estimated means and comparisons among the four groups.

If X_1 and X_2 are continuous variables, regression is usually the method of choice. This technique allows a great deal of flexibility for model building. In general, there is no reason to assume simple linear relationships between predictors and criterion. Quadratic terms, cross-product terms, etc. in X_1 and X_2 can be added to build a model which fits the data reasonably well.

If the fitted regression model is simple and the predictors are orthogonal then the estimated regression coefficients are easily interpreted. Often, however, orthogonality does not hold and models are not simple. As a result, the regression coefficient estimators are correlated and interpretation of single regression coefficients may be misleading.

As an extreme case, consider the "true" regression model $Y = 10X_1 + 10X_2 + \varepsilon$ where X_1 and X_2 are highly correlated (we assume without loss of generality that they are approximately equal.) Due to sampling variation, we can easily get estimated coefficients such as (0,20), (20,0), (30,-10), (100,-80), etc. In such a circumstance, even the sign of the regression coefficient may be suspect.

Researchers, aware of this problem, sometimes avoid the regression framework entirely. One common practice is to dichotomize the continuous predictor variables and proceed with an analysis of variance as described above. Occasionally, cases corresponding to central values of X_1 and X_2 are discarded. Some aspects of this problem have been studied by McCabe (1979). A significant advantage of this approach is that the results can be interpreted in terms of four groups. On the other hand, technical difficulties with the underlying model for this type of analysis are substantial. In addition, if the regression model is appropriate, substantial loss of information can result.

A major goal of statistical analysis is to take a large amount of numerical data and to reduce it to a small number of meaningful statements. In this context, one must have some sympathy for the dichotomizers. The purpose of this paper is to show that the results of a carefully performed regression analysis can easily be transformed into statements about subgroups (which we call pockets), thereby facilitating the interpretation of the data. It should be noted that we are not proposing an alternative to the usual regression analysis calculations. For example, the usual tests on particular coefficients or sets of coefficients are clearly useful. What we propose is a few additional calculations which may help produce the small number of meaningful statements sought by the researcher.

The suggested techniques involve routine application of the general theory of linear models. To illustrate the method, we consider data from a study in which two cognitive style measures were used to predict performance on three separate problem solving tasks.

PROBLEM BACKGROUND

The procedure described in this paper was developed to analyze part of the data from a large study (McCabe, 1976). Relevant aspects of the study are described below.

The predictor variables used were cognitive style measures. The dependent variables were problem solving tasks, namely verbal fluency, syllogistic reasoning and concept identification (French, Ekstrom and Price, 1963).

COGNITIVE STYLES

Cognitive styles are adaptive controls which affect cognitive processes and lead to adaptive solutions (Gardner, Holzman, Klein, Linton & Spence, 1959). Several particular cognitive styles have been identified (Kogan, 1971) and are presumed to coexist within the personality (Gardner, Jackson & Messick, 1960). Research suggests that the combined effects of two or more cognitive syles might better differentiate among persons than the effect of a singular cognitive style (Gardner, et al., 1969). Two cognitive styles were examined as predictor variables for this research. These styles are labelled field-dependence and breadth of categorization.

Field-dependence refers to individual difference in tendency to overcome the influence of conflicting perceptual cues. There are numerous indications that field-dependence level has wide implications for cognitive task performance in females (Barratt, 1955; Ehri & Muzio, 1974; Kogan & Wallach, 1964; Fitzgibbons, Goldberger & Eagle, 1965). For the data reported herein, the Group Embedded Figures Test (Witkin, Oltman, Raskin & Karp, 1971) - a measure which distinquishes field-dependent from field-independent subjects, was used.

Breadth of categorization indicates a style continuum, which encompasses personal preferences for dealing with relatively narrow, exclusive conceptual realms (categories), through preferences for relatively broad, inclusive categories. Aside from the time consuming object sorting tasks, the most widely used measure of breadth of categorization is the Estimation Questionnaire, henceforth denoted EQ (Pettigrew, 1958). Since the EQ is based upon quantitative content, it has been suggested that this measure is biased against females (Sherman, 1967). Such an argument is based upon the relative unfamiliarity of female subjects with quantitative content, and is supported by evidence that subjects tend to be broader in areas which they judge as personally relavant (Glixman & Wolfe, 1967). This particular objection to use of the Estimation Questionnaire could not be raised in connection with another breadth of categorization measure, namely the Synonymity Task (Fillenbaum, 1959), henceforth denoted ST, since the ST is based upon semantic content. Although the ST is listed along with the EQ, as a breadth of categorization measure, the degree of their relationship is a pertinent consideration. The study from which the current data is derived examined performance differences among subjects when breadth of categorization was defined by either the EQ or the ST.

COGNITIVE STYLE POCKETS

Four cognitive style pockets, each defined by a preselected level of field-dependence and breadth of categorization, were examined in relation to their performance on the three problem tasks. Each of the four pockets is denoted by one of the following combinations of the two cognitive styles:

FIBC (field-independent and broad categorizer)
FINC (field-independent and narrow categorizer)
FDBC (field-dependent and broad categorizer)
FDNC (field-dependent and narrow categorizer)

The question was asked: Do different pockets have significantly different dependent variable means? For each problem task, comparisons are made among the pockets.

METHOD

One hundred and six female undergraduates participated in the study for credit in an Introductory Psychology course. All subjects were tested together by a female experimenter during one evening session. Each task was a paper-and-pencil type, group administered. Since the data analyzed herein is part of a larger study involving test anxiety, tasks were administered under a particular type of preperformance instruction (Sarason, 1972) and a concealed stop watch was used for strictly timed tasks.

THE PROCEDURE

Application of regression analysis techniques to produce meaningful statements about the prediction of Y from X_1 and X_2 involves four steps. First a regression equation which fits the data well must be constructed.

Second, appropriate definitions of pockets must be determined. Finally, pocket means are estimated and tests for making comparisons among these means are performed.

CONSTRUCTION OF THE REGRESSION EQUATION

Construction of a regression model which fits the data well is the crucial first step in the proposed procedure. An inappropriate equation is likely to result in, at best, misleading statements about pocket means.

Sophisticated computer programs are no substitute for careful human judgement at this stage. Step-type regression procedures are inappropriate here. Residual plots and transformations are potentially useful tools. A thorough discussion of how to construct regression models is given in Draper and Smith (1966) and Neter and Wasserman (1974).

In general, one should take a rather liberal attitude with regard to inclusion of variables. Hence, marginal terms should be included in the equation and only those which are clearly insignificant should be discarded. The estimation of and comparisons made among subpopulation means (which is the point of this analysis) will not be seriously affected by the inclusion of a useless term or two but deletion of a potentially important term may have serious consequences.

A model with all terms up to order two has worked well with the examples considered. For convenience, this model, i.e.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{1i}^2 + \beta_4 X_{2i}^2 + \beta_5 X_{1i} X_{2i} + \varepsilon_i \qquad (1)$$

will be used in the subsequent discussion. Models of the general form

$$Y_i = \beta_0 + \Sigma \beta_j Z_{ji} + \varepsilon_i \qquad (2)$$

where each Z_j is a known function of X_1 and X_2 are treated in an analogous fashion.

DEFINITIONS OF POCKETS

Four pockets are defined, corresponding to the combinations resulting from considering high and low values for each of the predictor variables X_1 and X_2. In the example used to illustrate this procedure, the pockets are denoted FIBC, FINC, FDBC, and FDNC where FI, FD, BC and NC are abbreviations for field independent, field dependent, broad categorizer, respectively. For notational convenience in the following section, HH, HL, LH and LL will be used interchangeably with their correspondents, namely FIBC, FINC, FDBC and FDNC.

Each pocket corresponds to a particular pair of values for (X_1, X_2). In some cases, a priori reasoning may lead to appropriate choices for these definitions. In the absence of such considerations, we use values of the form

$$\bar{X}_i \pm c_i s_i \qquad (3)$$

where $\bar{X}_i$ and s_i are the mean and standard deviation, respectively, of X_i. In the examples studied, we have used $c_1 = c_2 = 1$. Thus, we have the following correspondences:

$$\begin{aligned} &\text{FIBC:} \quad (\bar{X}_1 + s_1, \bar{X}_2 + s_2) \\ &\text{FINC:} \quad (\bar{X}_1 + s_1, \bar{X}_2 - s_2) \\ &\text{FDBC:} \quad (\bar{X}_1 - s_1, \bar{X}_2 + s_2) \\ &\text{FDNC:} \quad (\bar{X}_1 - s_1, \bar{X}_2 - s_2) \end{aligned} \tag{4}$$

Thus, FIBC denotes the pocket which is one standard deviation above the mean on both field dependence (X_1) and breadth of categorization (X_2). The definitions of the other pockets are similarly translated.

Of course, there may not be any observations at (X_1, X_2) values corresponding to the group definitions. This fact causes no serious difficulties as long as there is some data around these points. If X_1 and X_2 are highly correlated, some difficulties may arise. In such cases, the (X_1, X_2) values corresponding to two of the pockets may appear to be unreasonable and uninteresting. One might consider using the principal components $(Z_1 + Z_2)/\sqrt{2}$ and $(Z_1 - Z_2)/\sqrt{2}$ (where $Z_i = (X_i - \bar{X})/s_i$) or some other means for avoiding this problem. However, care should be taken to avoid pocket definitions which are not easily interpreted. In any case, if the pockets are far from the center of the sample (in the Mahanalobis distance sense), the pocket means will be estimated with large standard errors and no significant useful results are likely to be obtained.

ESTIMATION OF POCKET MEANS

If we write the regression model (1) in matrix form as

$$Y = X\beta + \varepsilon \tag{5}$$

where $Y = (Y_1, Y_2, \ldots, Y_n)'$,

$$X = \begin{pmatrix} 1 & X_{11} & X_{21} & X_{11}^2 & X_{21}^2 & X_{11} & X_{21} \\ 1 & X_{12} & X_{22} & X_{12}^2 & X_{22}^2 & X_{12} & X_{22} \\ \vdots \\ 1 & X_{1n} & X_{2n} & X_{1n}^2 & X_{2n}^2 & X_{1n} & X_{2n} \end{pmatrix}$$

$$\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5)'$$

and

$$\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)'$$

then the least squares estimate of β is

$$\hat{\beta} = (X'X)^{-1} X'Y \ .$$

If the elements of the error vector ε are independently distributed with mean zero and variance σ^2 then $\hat{\beta}$ will have mean β and covariance matrix $\sigma^2(X'X)^{-1}$.
Let

$$G = \begin{pmatrix} g_{00} & g_{01} & \cdots & g_{05} \\ g_{10} & g_{11} & \cdots & g_{15} \\ \vdots & & & \\ g_{50} & g_{51} & \cdots & g_{55} \end{pmatrix}$$

be the usual estimate of this matrix, i.e.

$$G = s^2(X'X)^{-1} \tag{6}$$

where s^2 is the mean squared error (residual mean square) from the regression analysis.

Let x_{HH}, x_{HL}, x_{LH} and x_{LL} denote the designs corresponding to the four pockets. Using (4), this gives

$$x_{HH} = (1, \bar{X}_1+s_1, \bar{X}_2+s_2, (\bar{X}_1+s_1)^2, (\bar{X}_2+s_2)^2, (\bar{X}_1+s_1)(\bar{X}_2+s_2))', \tag{7}$$

$$x_{HL} = (1, \bar{X}_1+s_1, \bar{X}_2-s_2, (\bar{X}_1+s_1)^2, (\bar{X}_2-s_2)^2, (\bar{X}_1+s_1)(\bar{X}_2-s_2))', \tag{8}$$

$$x_{LH} = (1, \bar{X}_1-s_1, \bar{X}_2+s_2, (\bar{X}_1-s_1)^2, (\bar{X}_2+s_2)^2, (\bar{X}_1-s_1)(\bar{X}_2+s_2))', \tag{9}$$

and

$$x_{LL} = (1, \bar{X}_1-s_1, \bar{X}_2-s_2, (\bar{X}_1-s_1)^2, (\bar{X}_2-s_2)^2, (\bar{X}_1-s_1)(\bar{X}_2-s_2))'. \tag{10}$$

The usual estimates of the pocket means are given by

$$\hat{\mu} = x'\hat{\beta} \tag{11}$$

where x is x_{HH}, x_{HL}, x_{LH}, or x_{LL}. The variance of $\hat{\mu}$ is

$$s_{\hat{\mu}}^2 = x'Gx. \tag{12}$$

The estimation can be summarized by tabulating $(\hat{\mu}_{HH}, s_{HH}), (\hat{\mu}_{HL}, s_{HL})$, $(\hat{\mu}_{LH}, s_{LH})$ and $(\hat{\mu}_{LL}, s_{LL})$. If the errors are assumed to be normally distributed, then the $\hat{\mu}$'s are normally distributed and confidence intervals based on the t distribution with n-6 degrees of freedom are appropriate.

Note that, in general, the four estimated pocket means are correlated, since the same regression equation is used for each. Assessment of the exact overall error rate for the four confidence intervals is difficult. A practical solution is to use a Bonferroni approach. Use of 99% intervals for each mean will assure an overall error rate of not more than 4%.

COMPARISON OF SUBPOPULATION MEANS

Due to the already mentioned dependence among the four $\hat{\mu}$'s, the table of $\hat{\mu}$ and s values does not contain sufficient information to construct tests for comparisons among the means. For definiteness, let us consider comparing $\hat{\mu}_{HH}$ and $\hat{\mu}_{HL}$.

It is evident that the coefficients $\hat{\beta}_0, \hat{\beta}_1$ and $\hat{\beta}_3$ are not directly relevant to this comparison. Since

$$\hat{\mu}_{HH} = \hat{\beta}_0 + (\bar{X}_1+s_1)\hat{\beta}_1+(\bar{X}_2+s_2)\hat{\beta}_2+(\bar{X}_1+s_1)^2\hat{\beta}_3 + (\bar{X}_2+s_2)^2\hat{\beta}_4+(\bar{X}_1+s_1)(\bar{X}_2+s_2)\hat{\beta}_5 \tag{13}$$

and

$$\hat{\mu}_{HL} = \hat{\beta}_0 + (\bar{X}_1+s_1)\hat{\beta}_1+(\bar{X}_2-s_2)\hat{\beta}_2+(\bar{X}_1+s_1)^2\hat{\beta}_3 + (\bar{X}_2-s_2)^2\hat{\beta}_4+(\bar{X}_1+s_1)(\bar{X}_2-s_2)\hat{\beta}_5, \tag{14}$$

the difference between the two is simply the following linear combination of the $\hat{\beta}_i$:

$$\hat{\mu}_{HH}-\hat{\mu}_{HL} = (0)\hat{\beta}_0+(0)\hat{\beta}_1+(2s_2)\hat{\beta}_2+(0)\hat{\beta}_3 = (4\bar{X}_2s_2)\hat{\beta}_4+(2(\bar{X}_1+s_1))\hat{\beta}_5. \tag{15}$$

The null hypotheses

$$H_0: \ \mu_{HH} = \mu_{HL} \tag{16}$$

is thus translated into

$$H_0: \ 2s_2\beta_2+4\bar{X}_2s_2\beta_4+2(\bar{X}_1+s_1)s_2\beta_5 = 0. \tag{17}$$

Testing the hypothesis (17) is trivial. Let

$$a = (a_0,a_1,a_2,a_3,a_4)'$$

denote the coefficients of the β's in the null hypothesis of interest. Values of a for the other comparisons are given in Table I. For comparing HH and HL,

$$a = (0,0,2s_2,0,4\bar{X}_2s_2,2(\bar{X}_1+s_1)s_2).$$

the estimated difference between HH and HL is

$$\hat{\mu}_{HH}-\hat{\mu}_{HL} = a'\hat{\beta} \ . \tag{18}$$

The estimated variance of this difference is

$$s^2_{HH-HL} = a'Ga \; . \tag{19}$$

Thus, to test the null hypothesis that the two subpopulation means are equal we calculate

$$t = \frac{a'\hat{\beta}}{\sqrt{a'Ga}} \tag{20}$$

which has a t distribution with n-6 degrees of freedom.

Again problems of multiplicities arise when considering error rates for the six possible tests. Using a Bonferroni approach, one could run each test at the .01 level and have an overall rate not greater than .06. Alternatively, a Scheffe type approach could be used. However, such is likely to be too conservative in the present case.

The possibility of running one-sided tests using (20) should be recognized. If appropriate one-sided hypotheses can be generated a priori, this approach can be profitably exploited.

It should be noted that while most multiple regression package programs do not have the options available for calculating (20), the matrix G is often available. Some multivariate programs which have options for multivariate regression can be used to obtain (20). The output is usually in the form of an F-statistic. If probabilities are given they will usually be Scheffe-type. For two-sided tests use of the F distribution with 1 and n-6 degrees of freedom is appropriate whereas for one-sided test, taking the square root and affixing the proper sign will give (20).

FORMING OTHER POCKETS

Various other criteria can be used to define pockets. For instance one might prefer to make inferences about the average expected value for those subjects in the upper thirds on both X_1 and X_2.

For any subgroup of subjects, the average expected value of the dependent variable can be obtained by evaluating the regression equation at the average values of all predictor terms, e.g. $X_1, X_2, X_1X_2, X_1^2, X_2^2$. This procedure is equivalent to using a pocket defined by the average value for all predictor terms. Moreover, pockets may be defined by integrating all predictor terms with respect to any appropriate probability distribution.

Other extensions are reasonably straightforward. Additional predictor variables can be used. In some cases one might want to define more than two selected values of a particular predictor to generate the pockets.

RESULTS AND INTERPRETATION

Pocket means and standard errors are presented in Table 2 and the statistics for making comparisons among these means are presented in Table 3. Correlations among the variables are given in Table 4. To highlight the differences obtained by using the two different breadth of categorization measures (ST and EQ) graphical displays of the pocket means are provided by Figures 1,2, and 3.

For all three performance measures, field-dependence produces the clearest and largest effects. In all cases, field-independent (FI) pockets outperform field-dependent (FD) pockets. While these differences are statistically significant at the .05 level for only 50% of the cases, the general pattern is apparent from Figures 1,2, and 3.

The breadth of categorization measures add little information to the field-dependence measure for distinguishing pocket performance. In comparing the field-dependent (FD) pockets, there are no significant differences due to either breadth of categorization measure. For the field-independent (FI) pockets only one difference is evident: field-independent broad categorizers (FIBC) perform significantly better on the verbal fluency task than field-independent narrow categorizers (FINC), when breadth of categorization is definded by the Estimation Questionnaire (EQ).

Examination of Figures 1,2, and 3 reveals patterns which, although not statistically significant, are suggestive. When pockets are defined by the Estimation Questionnaire, the broad categorizers (BC) outperform the narrow categorizers (NC) in both field-independent and field-dependent pockets on all three performance tasks. The pattern still holds true for syllogisitc reasoning when the Synonymity Task (ST) is used as the breadth of categorization measure. However, for verbal fluency and concept identification, the pattern is reversed when the ST is used instead of the EQ. Specifically, in these cases, the narrow categorizers (NC) outperform the broad categorizers(BC). To conclude, when predicting the problem solving performance of this female population, field-dependence is a more useful measure than breadth of categorization.

TABLE I

Coefficients of β's for Comparison Tests

Comparison	a_0	a_1	a_2	a_3	a_4	a_5
HH vs HL	0	0	$2s_2$	0	$4\bar{X}_2 s_2$	$2(\bar{X}_1+s_1)s_2$
HH vs LH	0	$2s_1$	0	$4\bar{X}_1 s_1$	0	$2(\bar{X}_2+s_2)s_1$
HH vs LL	0	$2s_1$	$2s_2$	$4\bar{X}_1 s_1$	$4\bar{X}_2 s_2$	$2(\bar{X}_1 s_2+\bar{X}_2 s_1)$
HL vs LH	0	$2s_1$	$-2s_2$	$4\bar{X}_1 s_1$	$-4\bar{X}_2 s_2$	$2(\bar{X}_2 s_1-\bar{X}_1 s_2)$
HL vs LL	0	$2s_1$	0	$4\bar{X}_1 s_1$	0	$2(\bar{X}_2-s_2)s_1$
LH vs LL	0	0	$2s_2$	0	$4\bar{X}_2 s_2$	$s(\bar{X}_1-s_1)s_2$

TABLE 2

Estimated Pocket Means and Standard Errors

Task	Breadth of Categorization Measure	Pockets			
		FIBC	FINC	FDBC	FDNC
Verbal	ST	23.44 ± 1.31	23.44 ± 1.40	19.77 ± 1.24	20.27 ± 1.34
Fluency	EQ	25.42 ± 1.33	21.85 ± 1.36	20.75 ± 1.22	18.89 ± 1.46
Syllogistic	ST	11.39 ± 1.42	10.60 ± 1.51	5.55 ± 1.34	3.54 ± 1.45
Reasoning	EQ	11.80 ± 1.50	10.17 ± 1.53	4.73 ± 1.37	4.51 ± 1.64
Concept	ST	96.27 ± 7.56	104.20 ± 8.05	77.83 ± 7.17	82.48 ± 7.73
Identification	EQ	103.45 ± 7.87	98.17 ± 8.05	85.62 ± 7.21	73.40 ± 8.61

TABLE 3

F-Statistics and Significance Values for Compared Pockets

Task	Breadth of Categorization Measure*	Compared Pockets: FIBC FINC		FIBC FDBC		FIBC FDNC		FINC FDBC		FINC FDNC		FDBC FDNC	
		F**	p	F**	p	F**	p	F**	p	F**	p	F**	p
Verbal	ST	0.00	ns	3.98	.05	2.38	.13	3.03	.09	2.30	.14	0.10	ns
Fluency	EQ	4.07	.05	6.16	.02	9.94	.01	0.26	ns	1.71	.20	1.06	ns
Syllogistic	ST	0.17	ns	8.63	.01	12.49	.01	4.90	.03	9.77	.01	1.37	ns
Reasoning	EQ	0.63	ns	10.66	.01	9.46	.01	0.01	ns	4.90	.03	0.01	ns
Concept	ST	0.59	ns	3.03	.09	1.35	ns	4.70	.04	3.25	.08	0.26	ns
Identification	EQ	0.24	ns	2.46	.12	5.85	.02	0.41	ns	3.41	.07	1.31	ns

*ST designates Fillenbaum's Synonymity Task and EQ designates Pettigrew's Estimation Questionnaire.

**Degrees of freedom for all F-statistics are 1 for the numerator and 100 for the denominator.

TABLE 4

Correlations Among Variables (N = 106)

	1	2	3	4	5
1. ST					
2. EQ	.14				
3. Field-dependence	-.06	-.04			
4. Verbal Fluency	-.06	.19*	.29**		
5. Syllogistic Reasoning	.05	.04	.34**	.24**	
6. Concept Identification	-.09	.11	.23**	.01	.21*

* $p < .05$

** $p < .01$

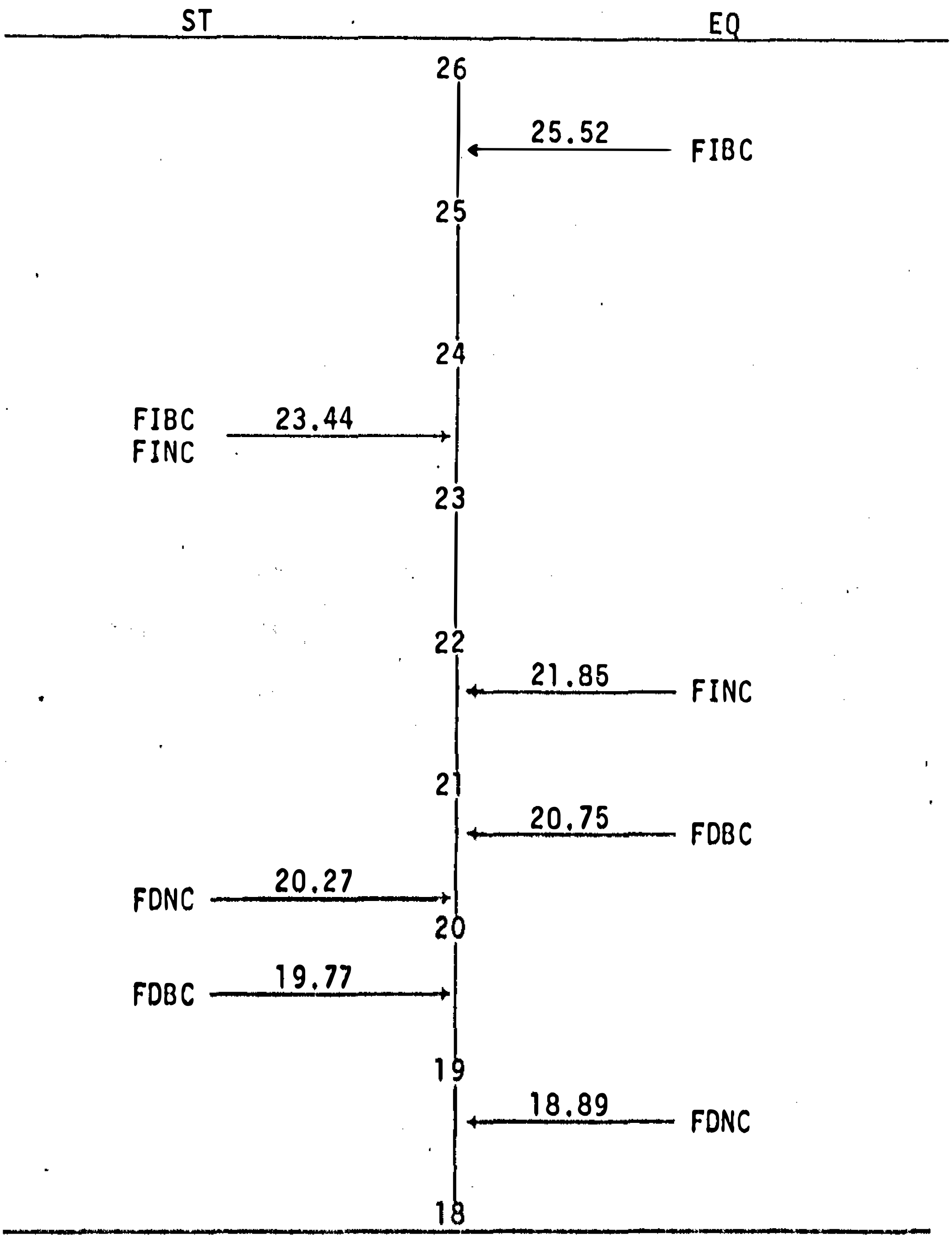

Figure 1. Pocket Means For Verbal Fluency

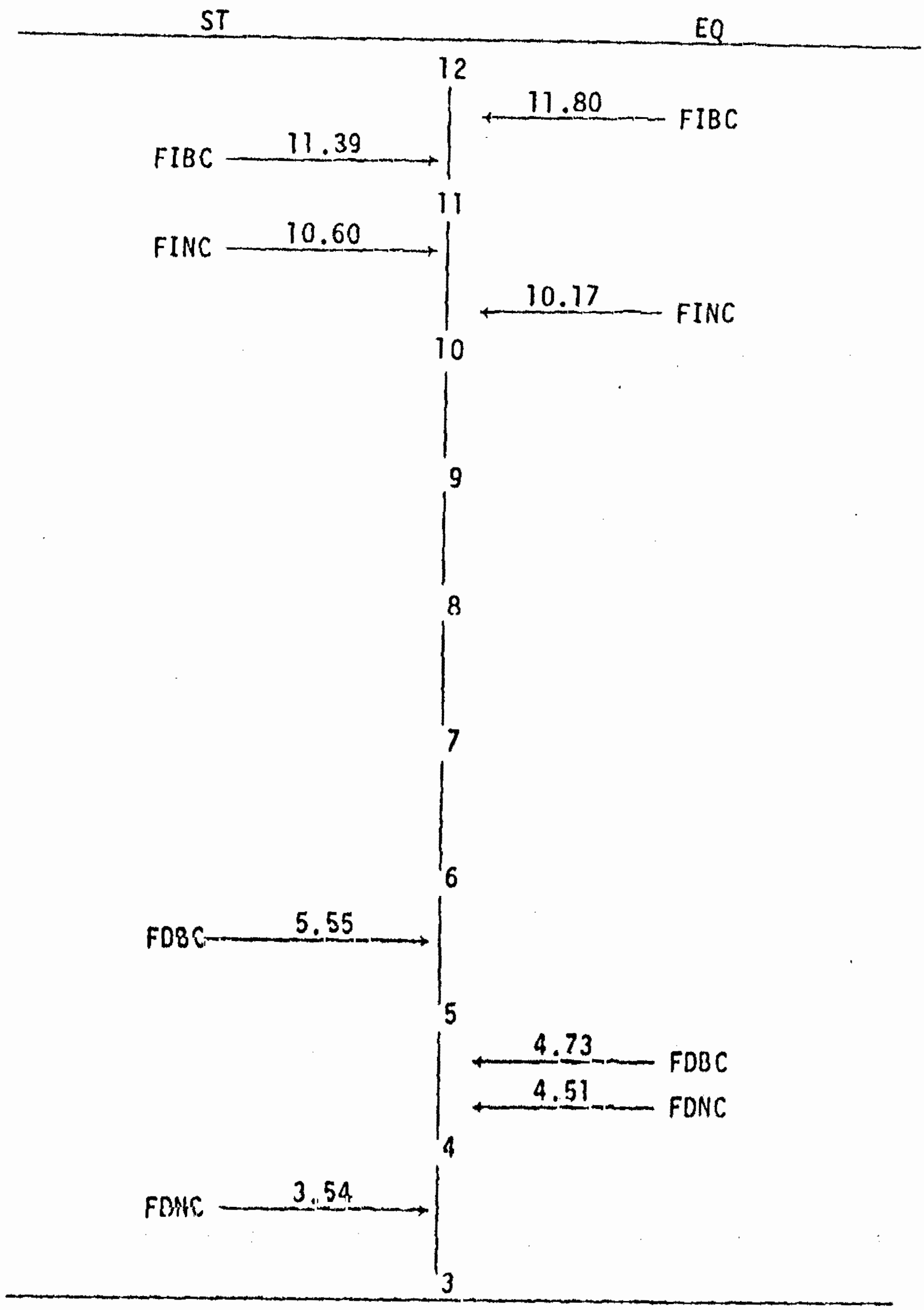

Figure 2. Pocket Means for Syllogistic Reasoning

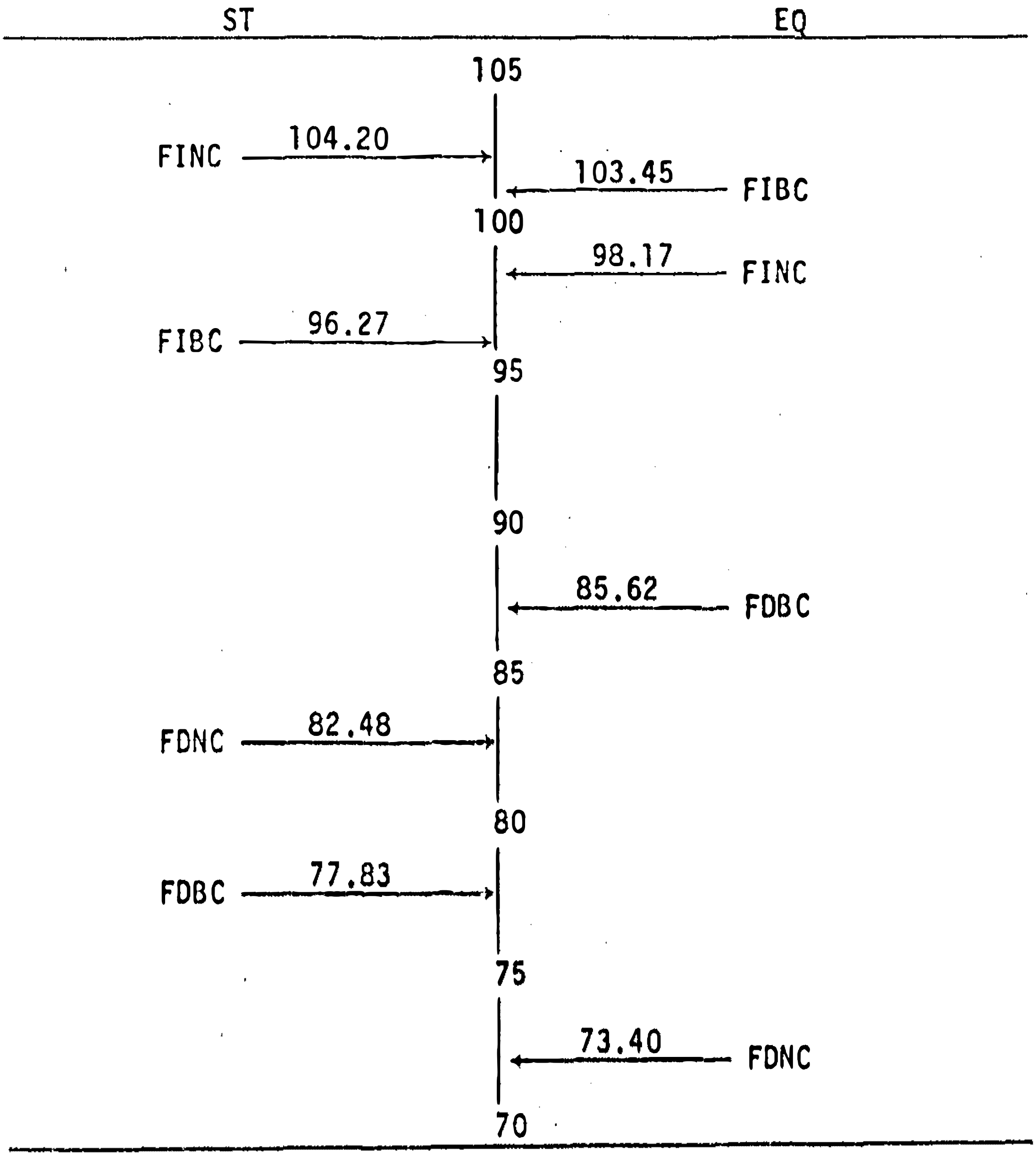

Figure 3. Pocket Means for Concept Identification

REFERENCES

Barratt, E. S. The space-visualization factors related to temperament traits. Journal of Psychology, 1955, 39, 279-280.

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

Ehri, L. C. & Muzio, I. M. Cognitive style and reasoning about speed. Journal of Education Psychology, 1974, 66, 569-571.

Fillenbaum, S. Some stylistic aspects of categorizing behavior. Journal of Personality, 1959, 27, 187-195.

Fitzgibbons, D., Goldberger, L. & Eagle, M. Field-dependence and memory for incidental material. Perceptual and Motor Skills, 1965, 21, 743-749.

French, J. W., Ekstrom, R. B. & Price, L. A. Kit of Reference Tests for Cognitive Factors. New Jersey: Educational Testing Service, 1963.

Gardner, R. W., Holzman, P. S., Klein, G. S., Linton, H. B. & Spence, D.P. Cognitive control: A study of individual consistencies in cognitive behavior. Psychological Issues, 1959, 1, 1-185.

Gardner, R. W., Jackson, D. N. & Messick, S. J. Personality organization in cognitive controls and intellectual abilities. Psychological Issues, 1960, 2, 1-148.

Glixman, A. F. & Wolfe, J. C. Category membership and interitem semantic-space distances. Journal of Personality, 1967, 35, 134-143.

Kogan, N. Educational implications of cognitive styles. In G. S. Lesser (ed.), Psychology and Educational Practice. Illinois: Scott, Foresman and Company, 1971, pp. 242-292.

Kogan, N. and Wallach, M. Risk-Taking. New York: Halt, Rinehart & Winston, 1964.

McCabe, G. P. Use of the 27% rule in experimental design. To appear in Communications in Statistics, 1979.

McCabe, S. A. Prediction of female problem solving skill from cognitive style measures under motivating task-orienting conditions. Ph.D. Dissertation, Purdue University, 1976.

Neter, John and Wasserman, William. Applied Linear Statistical Models. Homewood, Illinois: Richard D. Irwin, 1974.

Pettigrew, T. F. The measurement and correlates of category width as a cognitive variable. Journal of Personality, 1958, 26, 532-544.

Sarason, I. G. Experimental approaches to test anxiety: Attention and the uses of information. In C. D. Spielberger (Ed.), Anxiety: Current Trends in Theory and Research. Vol. 2, New York: Academic Press, 1972, pp. 383-409.

Witkin, H. A., Oltman, P. K., Raskin, E. and Karp, S. A. Manual: Embedded Figures Test, Children's Embedded Figures Test, Group Embedded Figures Test. Palo Alto: Consulting Psychologists Press, Inc., 1971.

George P. McCabe & Sharron A. S. McCabe

Purdue University

MULTIPLE LINEAR REGRESSION VIEWPOINTS
February, 1980 Volume 10, Number 2

HANDLING DISPROPORTIONALITY IN TWO-WAY ANOVAS

Ken Black, University of Houston at Clear Lake City
William K. Brookshire, North Texas State University

Abstract

This paper examines three methods of handling disproportionate cell frequencies in two-way analysis of variance. A Monte Carlo approach was used to study the method of expected frequencies and two multiple regression approaches to the problem as disproportionality increased. Four cases were studied: no effects case, row effects case, interaction effects case, and a row and column effects case. Type I and Type II errors were examined. Several conclusions were reached with regard to the appropriateness of each technique in handling disproportionality.

Introduction

Disproportionate cell frequencies in analysis of variance designs prohibit the researcher from utilizing traditional techniques. Under such conditions, the sums of squares are not additive and yield biased results (Ostle, 1954; Snedecor, 1946; Mood, 1950; Roscoe, 1975; Kendall, 1948; Wert, Neidt, and Almann, 1954).

Factorial designs containing disproportionate cell frequencies can occur in many ways. In multiple-classification of data, unequal representation in each cell is a common occurrence (Tsao, 1946). In the social sciences, unequal subclasses is the rule rather than the exception (Johnson and Jaçkson, 1959; Bessent, 1974). Often unequal sample sizes occur naturally when the variables being observed are such things as classrooms. Experimental subject mortality can occur inadvertantly or purposely (some subjects are dropped from the study as being inappropriate) (Proger, 1972). Failure to record and gross errors in recording can cause missing data (Cochran and Cox, 1950).

There are at least eight different techniques that can be used to solve the problem of disproportionality (Williams, 1972). These can be divided into three categories: data forcing methods, approximate methods, and regression methods. The data forcing methods include: discarding data and estimating missing data. The approximate methods include: unweighted means analysis, method of weighted means, and the method of expected frequencies. The regression techniques include what Overall and Klett (1972) call the "complete linear-model analysis", the "experimental-design analysis", and the "step-down analysis".

The method of discarding data can be wasteful and causes the investigator to lose information (Williams, 1972; Wert, Neidt, and Almann, 1954). It has been found to be a poor alternative to other methods due to the strong tendency to yield Type II errors (Dalton, 1976). Some researchers have recommended this method only if the number of observations differ by a few and if the observations are not cast out permanently (Searle, 1971).

The method of estimating missing data yields treatment effects that are slightly inflated (Williams, 1972; Cochran and Cox, 1950). This method is no longer appropriate with the availability of computer resources, and it can be psychologically unnerving to artificially create data for analysis (Williams, 1972). Neither the method of discarding data nor the method of estimating missing data are exact tests.

The unweighted means analysis may be the most widely used technique for handling disproportionate cell frequencies (Williams, 1972). It is probably the simplest and one of the most justifiable techniques for analyzing disproportionate designs (Glass and Stanley, 1970). It has a minimum of computation and furnishes a short-cut procedure of testing for interaction (Anderson and Bancroft, 1952). Experimenters are warned that the unweighted means analysis should be applied only if n's are not very disparate (Myers, 1972).

The method of weighted means was devised by Yates in 1934 with the assumption that interaction exists (Searle, 1971; Tsao, 1946). It is seldom recommended when there are two or more missing scores per cell (Dalton, 1976). Some researchers claim that the weighted means analysis yields tests for main effects which are not the usual F statistic and which have different power functions. Furthermore, as long as no empty cells appear, the method of unweighted means is more generally usable and offers an analysis similar to what the experimenter is familiar with in the equal or proportional frequency case (Steinhorst and Miller, 1969).

The method of expected frequencies involves multiplication of cell sums by the expected cell frequency to obtain a sum for each cell. Sums obtained in this manner are used in estimating main effects and interactions (Dalton, 1976). This method is appropriate when proportionality is not too great (Myers, 1972). This method has been used largely when cell frequencies would naturally be disproportionate.

The "complete linear-model analysis" is a regression method that involves an estimation of independent effects of each factor adjusted for all others included in the model (Overall and Klett, 1972). Some researchers believe that this method is the best extension of traditional analysis of variance because the same parameters are estimated and the same hypotheses are tested in the orthogonal and the nonorthogonal cases (Carlson and Timm, 1974).

The "experimental-design analysis" involves an estimation of main effects disregarding interactions and then an estimation of interactions adjusted for main effects (Overall and Klett, 1972). Overall and Spiegel (1969) stated that this method seemed to be the most appropriate method for analysis of experimental data involving disproportionate cell frequencies. Later, they reversed this stance in favor of the "complete linear-model analysis" (Overall, Spiegel, and Cohen, 1975). Speed and Hocking (1976) said that the proper method depends on the hypothesis being tested.

The other regression method for handling disproportionate two-way analysis of variance is the "step-down analysis". This method involves an initial ordering of the effects and then estimating each effect adjusting for those preceding it in the ordering and ignoring those following it (Overall and Klett, 1972). This method has been referred to as the hierarchal model. With this approach, a researcher is required to order the variables in relation to their research interest (Williams and Linden, 1972). The requirement of establishing a priori an ordering of variables limits its usage to the researcher (Dalton, 1976). The immediate importance of this method lies in its appropriateness for the mixed model (Searle, 1971).

There has been considerable debate as to which, if any, of the techniques is more appropriate in handling the disproportionality situation. Little research has been done in this area.

The Study

A Monte Carlo simulation study was conducted to determine which of the three of the more promising techniques were more appropriate and effective in handling disproportionate cell frequencies. The three techniques chosen for study were: (1) method of expected frequencies, (2) the "complete linear-model analysis", and (3) the "experimental-design analysis". The data forcing techniques were not included because the literature has shown them to be the poorest alternative solutions to the problem of disproportionality. The method of weighted means was not included because it is seldom recommended when there are two or more missing values per cell. The "step-down analysis", of the regression solutions was not used because it does not test the same hypotheses as orthogonal analysis of variance.

The study considered four cases. The first case was the no effect case which was used to examine Type I errors. To study this case, designs were derived in which there were no built-in row, column, or interaction effects. The other three cases were: (1) the row effects case, (2) the row and column effects case, and (3) the interaction effects case. In each of these three cases, effects were built-in, and Type II errors could be examined.

Procedure

Within each case, the procedure was identical. Forty random numbers were generated into a 2x2 design with ten numbers in each cell. An equal cell ANOVA was performed. Forty random numbers in a disproportionate design were then produced. Disproportionality was measured by a modified chi-square approach (Ferguson, 1971). Table 1 contains the chi-square values and their associated frequencies. Analysis of variance was performed by the method of expected frequencies, the "complete linear-model analysis", and the "experimental-design analysis". Forty more random numbers were generated, and the process was repeated 1000 times. The probability of each F value occurring was calculated, and a frequency distribution of the probabilities was tabulated for each method. The probability distributions for the three techniques were each compared to the probability distribution of the equal cell ANOVA by a Kolmogorov-Smirnov test. The entire process was repeated for eleven levels of disproportionality. An examination was made of the behavior of the three techniques as disproportionality increased.

Table 1

Cell Frequencies for Chi-Square Values of Disproportionality

χ^2 Value	Frequencies		χ^2 Value	Frequencies	
0.0	10	10	8.6	6	13
	10	10		16	5
1.6	8	12	19.4	4	13
	12	8		20	3
2.6	7	13	24.4	5	9
	12	8		3	23
3.6	7	13	26.6	2	22
	13	7		13	2
6.4	6	14	40.6	3	8
	14	6		27	2
7.4	4	16	59.6	5	1
	11	9		3	31

Techniques

The method of expected frequencies involves the multiplication of cell sums by the expected cell frequency to obtain a sum for each cell. Sums obtained in this manner are used in estimating main effects and interactions (Dalton, 1976).

The "complete linear-model analysis", involves an estimation of independent effects of each factor adjusted for all others included in the model. The structural model for this method in a two-way analysis of variance is:

$$X_{ijm} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + e_{ijm}$$

The "experimental-design analysis", involves an estimation of main effects disregarding interactions and then estimating interactions adjusted for main effects. This method makes the assumption that no true interaction exists. The structural model for this method in a two-way analysis of variance is:

$$X_{ijm} = \mu + \alpha_i + \beta_j + e_{ijm}$$

Results

Case 1

Case One was the no effects case. Kolmogorov-Smirnov D values were used to test the goodness-of-fit of the probability distributions of each of the three methods of handling disproportionality against the equal cell situation. The resulting D values for row, column, and interaction effects are presented in Table 2.

The "complete linear-model analysis" and the "experimental-design analysis" do not yield significantly different D values from the equal cell ANOVA until chi-square is 26.6 or greater (df = 1). Based on probabilities, chi-square values greater than 19.4 were considered extreme disproportionality. Thus, the above methods did not yield significantly different results from the equal cell ANOVA until extreme disproportionality was achieved in the no effects case.

The method of expected frequencies yielded some significant results for $\chi^2 \geq 6.4$. Moderate disproportionality was $3.6 \leq \chi^2 \leq 19.4$. Thus, the method of expected frequencies appears to have a strong tendency to commit Type I errors when moderate and extreme disproportionality occurs.

In addition, frequencies of row, column, and interaction F values were tallied at the .10, .05, and .01 levels to aid in determining why significant D values were obtained. These values are also included in Table 2. At $\chi^2 = 3.6$, the method of expected frequencies was producing 66 significant F values at $\alpha = .05$. The expected number is 50. At $\chi^2 = 19.4$, this figure jumps to 147.

Frequency counts of F values at the .10, .05, and .01 levels for the other two methods reveals very little difference from the expected values even at $\chi^2 = 19.4$.

All three methods produced no F values significant at the .10, .05, and .01 levels when $\chi^2 \geq 26.6$. The total frequency distributions are skewed very heavily toward the small levels of probability. Thus, when no effects are present, all three methods are very conservative.

Table 2

D Values and Frequency of F Values for $\alpha \leq .05$ for No Effects Case Comparing the Three Methods of Handling Disproportionality to the Equal Cell ANOVA as χ^2 Increases for Rows, Columns, and Interactions

D Values

χ^2 value		method of expected frequencies D	freq.**	"complete linear-model analysis" D	freq.	"experimental-design analysis" D	freq.
	row	.000	38	.000	38	.000	38
0.0	col.	.000	49	.000	49	.000	49
	inter.	.000	43	.000	43	.000	43
	row	.024	37	.020	36	.020	36
1.6	col.	.014	58	.016	51	.016	51
	inter.	.021	45	.026	43	.026	43
	row	.026	45	.011	39	.013	40
2.6	col.	.031	63	.019	60	.021	58
	inter.	.021	52	.026	46	.026	46
	row	.041	44	.021	37	.021	37
3.6	col.	.023	66	.034	55	.034	55
	inter.	.017	51	.036	45	.036	45
	row	.057*	61	.034	40	.034	40
6.4	col.	.051*	64	.015	47	.015	47
	inter.	.032	65	.030	49	.030	49
	row	.063*	66	.027	46	.025	44
7.4	col.	.031	63	.026	55	.025	46
	inter.	.041	63	.021	45	.021	45
	row	.076*	61	.023	39	.028	41
8.6	col.	.075*	83	.025	48	.027	51
	inter.	.056*	69	.031	42	.031	42
	row	.139*	124	.026	41	.035	42
19.4	col.	.148*	147	.013	63	.019	64
	inter.	.147*	115	.038	49	.038	49
	row	.409*	0	.689*	0	.742*	0
26.6	col.	.435*	0	.693*	0	.643*	0
	inter.	.508*	0	.684*	0	.684*	0
	row	.490*	0	.751*	0	.761*	0
40.6	col.	.517*	0	.741*	0	.693*	0
	inter.	.645*	0	.732*	0	.732*	0
	row	.489*	0	.787*	0	.689*	0
59.6	col.	.539*	0	.787*	0	.846*	0
	inter.	.714*	0	.789*	0	.789*	0

* denotes significant at the .05 level
** the expected frequency is 50.

Case 2

Case Two was the row effects case. Row effects were built in to the model. Three power levels were used in generating the effects. The results are presented in Table 3. These results were obtained while using a power of .60 at $\alpha = .05$.

In this case, significant D values were obtained for the "complete linear-model analysis" and the "experimental-design analysis" for row effects at $\chi^2 \geq 3.6$. The method of expected frequencies produced significant D values for $\chi^2 \geq 19.4$.

An examination of the F frequencies at the .05 level of significance for row effects shows a drop in numbers for the "complete linear-model analysis" and the "experimental-design analysis" compared to the method of expected frequencies at $\chi^2 = 2.6$. At this level of disproportionality for $\alpha = .05$ and $1 - \beta = .60$, these two methods produced 552 and 553 F values when the expected number was 600. Thus, there is a tendency towards Type II errors. As chi-square increased, the number of F's at $\alpha \leq .05$ dropped. At $\chi^2 = 10.0$, both methods produced 449 F values. At $\chi^2 = 19.4$, the figure is around 355. Meanwhile, the method of expected frequencies produced more F values at these levels than either of the other two methods. At $\chi^2 = 19.4$, it still was yielding 536 F values at $\alpha \leq .05$ for a power of .60.

For levels of disproportionality, $\chi^2 \geq 24.4$, all three methods produced no F values with probabilities less than .10. Thus, for extreme levels of disproportionality Type II errors occurred frequently.

Table 3

D Values and Frequency of F Values for $\alpha < .05$ and $1 - \beta = .60$ for the Row Effects Case Comparing the Three Methods of Handling Disproportionality to the Equal Cell ANOVA as χ^2 Increases for Rows

	D Values					
	method of expected frequencies		"complete linear-model analysis"		"experimental-design analysis"	
χ^2 value	D	freq.**	D	freq.	D	freq.
0.0	.000	579	.000	579	.000	579
1.6	.013	566	.024	556	.024	556
2.6	.017	580	.027	552	.027	553
3.6	.016	572	.045*	534	.045*	534
6.4	.026	582	.084*	509	.084*	509
8.6	.034	575	.096*	483	.095*	484
10.0	.037	553	.130*	449	.130*	449
19.4	.072*	536	.228*	351	.224*	355
24.4	.970*	0	.986*	0	.987*	0
26.6	.953*	0	.987*	0	.989*	0
40.6	.942*	0	.987*	0	.989*	0

* denotes significant at the .05 level
** the expected frequency is 600

Case 3

Case three was the interaction effects case. A power of .60 was again used at $\alpha = .05$. Only interaction effects were built-in in case three. The results of the D values are in Table 4.

Significant results were obtained for the "complete linear-model analysis" and the "experimental-design analysis for $\chi^2 \geq 3.6$. These results correspond with results for the row effects case. The frequencies of values at the .05 level of significance are similar to the row effects case. Type II errors are being committed for $\chi^2 \geq 3.6$.

The method of expected frequencies yielded significant D values for $\chi^2 \geq 7.4$. This is at a lower level of disproportionality than for the row effects case. An examination of the frequency of F values at $\alpha = .05$ and power equal .60 shows that this method produced 541 F values at $\alpha \leq .05$ for interaction effects. For $\chi^2 \geq 26.6$, all three methods produced zero F values with probabilities less than .10.

Table 4

D Values and Frequency of F values for $\alpha < .05$ and $1 - \beta = .60$ for the Interaction Effects Case Comparing the Three Methods of Handling Disproportionality to the Equal Cell ANOVA as χ^2 Increases for Interaction

	D Values					
χ^2 value	method of expected frequencies		"complete linear-model analysis"		"experimental-design analysis"	
	D	freq.**	D	freq.	D	freq.
0.0	.000	589	.000	589	.000	9
1.6	.014	598	.015	585	.015	585
2.6	.016	591	.025	564	.025	564
3.6	.014	592	.047*	550	.047*	550
6.4	.038	567	.100*	489	.100*	484
7.4	.074*	541	.131*	458	.131*	458
8.6	.054*	541	.131*	459	.131*	459
19.4	.133*	493	.270*	331	.270*	331
26.6	.961*	0	.978*	0	.976*	0
40.6	.976*	0	.985*	0	.985*	0
59.6	.974*	0	.991*	0	.991*	0

* denotes significant at the .05 level
** the expected frequency is 600

Case 4

Case Four is the row and column effects case. Effects were built in for rows and columns using a power of .60 at $\alpha = .05$. The resulting D values are in Table 5.

Table 5 reveals that for $\chi^2 \geq 3.6$, the "complete linear-model analysis" and the "experimental-design analysis" yielded significant D values for both row and column effects. An examination of frequencies of F values for $\alpha \leq .05$ reveals that for $\chi^2 \geq 3.6$, the gap between the number of F values expected and those yielded widens. These three methods commit Type II errors for $\chi^2 \geq 3.6$.

The method of expected frequencies yielded significant D values for $\chi^2 \geq 19.4$. However, at $\chi^2 = 19.4$, about 530 of the F values have probabilities less than .05. This is a much smaller tendency toward Type II errors than the other two methods which yielded only about 360. For $\chi^2 \geq 24.4$, zero F values are produced for $\alpha \leq .10$ for all three methods.

Table 5

D Values and Frequency of F values for $\alpha < .05$ and $1 - \beta = .60$ for the Row and Column Effects Case Comparing the Three Methods of Handling Disproportionality to the Equal Cell ANOVA as χ^2 Increases for Rows and Columns

		D Values					
χ^2 value		method expected frequencies		"complete linear-model analysis"		"experimental-design analysis"	
		D	freq.**	D	freq.	D	freq.
0.0	row	.000	579	.000	579	.000	579
	col.	.000	576	.000	576	.000	576
1.6	row	.013	566	.024	555	.024	555
	col.	.017	593	.022	581	.021	581
2.6	row	.017	580	.027	552	.026	553
	col.	.011	576	.036	549	.035	547
3.6	row	.016	572	.045**	534	.045*	534
	col.	.011	579	.044*	543	.044*	543
6.4	row	.026	582	.084	509	.084*	509
	col.	.027	577	.068*	508	.068*	508
8.6	row	.033	575	.096*	483	.092*	487
	col.	.036	577	.092*	495	.086*	494
10.0	row	.037	553	.131*	448	.131*	448
	col.	.033	578	.113*	463	.113*	463
19.4	row	.072*	538	.228*	351	.223*	356
	col.	.082*	526	.225*	369	.226*	361
24.4	row	.970*	0	.986*	0	.987*	0
	col.	.924*	0	.970*	0	.930*	0
26.6	row	.953*	0	.987*	0	.989*	0
	col.	.895*	0	.962*	0	.955*	0

* denotes significant at the .05 level
** the expected frequency is 600

Discussion

Several conclusions can be reached from this study. For small levels of disproportionality, all three methods will yield similar nonspurious results; and thus, any of the three methods would be appropriate for use. For moderate levels of disproportionality, the "complete linear-model analysis" appears to be the best method to use to control Type I errors. The method of expected frequencies appears to be the best method to control Type II errors. For extreme levels of disproportionality, all three methods yield spurious results.

The two regression methods produce very similar results. In all cases, at a chi-square with a probability level of less than or equal to .06, at least one of the three methods yields spurious results in all cases.

References

Anderson, R. L. and T. A. Bancroft. Statistical Theory in Research. (New York: McGraw-Hill Book Company, 1952).

Bessent, Authella Ann McGuirre. "A Monte Carlo Study of the Analysis of Variance by Unweighted Means," Doctoral dissertation, The University of Texas at Austin. Austin, Texas, August, 1974.

Carlson, James E. and Neil H. Timm. "Analysis of Nonorthogonal Fixed-Effects Designs." Psychological Bulletin. (September, 1974, 81, 563-570).

Cochran, William G. and Gertrude M. Cox. Experimental Designs. (New York: John Wiley and Sons, Inc., 1950).

Dalton, Starrette. "Regression Approaches and Approximate Solutions to Analysis of Variance with Disproportionality Varied." Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976).

Ferguson, George A. Statistical Analysis in Psychology and Education. (New York: McGraw-Hill Book Company, 1971).

Glass, Gene V. and Julian C. Stanley. Statistical Methods in Education and Psychology. (Englewood Cliffs, New Jersey: Presntice-Hall, 1970).

Johnson, Palmer O. and Robert W. B. Jackson. Modern Statistical Methods: Descriptive and Inductive. (Chicago: Rand McNally and Company, 1959).

Kendall, Maurice G. The Advanced Theory of Statistics. (London: Charles Griffith and Company, Limited, 1948).

Mood, Alexander McFarlene. Introduction to the Theory of Statistics. (New York: McGraw-Hill Book Company, 1950).

Myers, Jerome L. Fundamentals of Experimental Design. (Boston: Allyn and Bacon, Inc., 1972).

Ostle, Bernard. Statistics in Research. (Ames, Iowa: The Iowa State College Press, 1954).

Overall, John E. and C. James Klett. Applied Multivariate Analysis. (New York: McGraw-Hill Book Company, 1972).

Overall, John E. and Douglas K. Spiegal. "Concerning Least Squares Analysis of Experimental Data." Psychological Bulletin. (1969, 72, 311-322).

Overall, John E., Douglas K. Spiegal, and Jacob Cohen. "Equivalence of Orthogonal and Nonorthogonal Analysis of Variance." Psychological Bulletin. (March, 1975, 82, 182-186).

Proger, Barton B. "Unequal Cell Frequencies in Analysis of Variance: A Review and Extension of Methodology for Multiple Missing Observations." Paper published at Montgomery County Schools, King of Prussia, Pennsylvania. Reported in ERIC, 1972.

Roscoe, John T. Fundamental Research Statistics for Behavioral Sciences. (New York: Holt, Rinehart and Winston, Inc., 1975).

Searle, Shayle R. Linear Models. (New York: Wiley Publishing Company, 1971).

Speed, F. M. and R. R. Hocking. "The Use of the R( )-Notation with Unbalanced Data." The American Statistician. (February, 1976, 30, 30-33).

Steinhorst, R. Kirk and C. Dean Miller. "Disproportionality of Cell Frequencies in Psychological and Educational Experiments Involving Multiple Classifications." Educational and Psychological Measurement. (Winter, 1969, 29, 799-811).

Tsao, Fei. "General Solution of the Analysis of Variance and Covariance in the Case of Unequal or Disproportionate Numbers of Observations in the Subclasses." Psychometrika (June, 1946, 11, 107-128).

Snedecor, George W. Statistical Methods. (Ames, Iowa: The Iowa State College Press, 1946).

Wert, James E., Charles O. Neidt, and J. Stanley Almann. Statistical Methods in Educational and Psychological Research. (New York: Appleton-Century Crofts, Inc., 1954).

Williams, John D. "Two Way Fixed Effects Analysis of Variance with Disproportionate Cell Frequencies." Multivariate Behavioral Research. (1972, 7, 66-82).

Williams, John D. and Alfred C. Lindem. "A Computer Program for Two-Way Fixed Effects Analysis of Variance with Disproportionate Cell Frequencies." Educational and Psychological Measurement. (1972, 32, 169-172).

Ken Black
University of Houston at Clear Lake City
William K. Brookshire
North Texas State University

MULTIPLE LINEAR REGRESSION VIEWPOINTS
February, 1980 Volume 10, Number 2

THE PREDICTION OF FACULTY RANK: A COMPARISON OF TWO MULTIVARIATE TECHNIQUES

Sonya L. Blixt
Kent State University

Abstract Multiple regression is advocated as the appropriate method when one wishes to predict a criterion measured on an interval or ratio scale. Discriminant analysis often followed by a classification procedure is recommended in the prediction of a nominal variable. The purpose of this empirical study was a comparison of the two techniques when the criterion is ordinal in nature.

The criterion was the current academic rank (Full, Associate, or Assistant Professor) of a faculty member. The predictors used in this study were salary, age, years at the university, years of professional experience before joining the university, and the year that the faculty member gained his/her current rank at the university. The sample totaled 103 faculty members.

A multiple regression equation and a discriminant function were calculated on one-half of the sample. The weights generated from the two models were then applied to the other half to determine which technique provided the more correct prediction of faculty rank.

It was found that the regression technique was better able to predict the ordinal variable for the cross-validation sample. Over all ranks, the regression technique correctly placed 70.59% of the people and the discriminant technique correctly placed 60.78%. Consequently, it appears that even though a scale is ordinal, multiple regression can prove to be a powerful technique. However, it is possible that the regression technique proved to be more powerful because only one discriminant function was significant.

Introduction

Some recent studies have been devoted to a comparison between discriminant and regression analysis (e.g., Alumbaugh et al., 1978; Bledsoe, 1973). The studies, however, have not used a dependent measure which is clearly ordinal in nature. Discussions of multivariate analysis (e.g., Overall & Klett, 1972; Tatsuoka, 1971) indicate that discriminant analysis followed by a classification procedure is appropriate for assigning individuals to groups, but the literature is unclear regarding the appropriate procedure for groups arranged in an ordinal fashion. The literature on regression (e.g., Guilford & Fruchter, 1978) indicates that, if the dependent variable is not dichotomous, it must be measured on an interval or ratio scale for the results to be "meaningful." This requirement, however, does not suggest how to deal with a dependent measure that is ordinal in nature when one wishes to predict group membership.

The main purpose of the present study was to determine which statistical procedure, discriminant function analysis followed by classification or multiple regression, would provide the most accuracy in assigning subjects to groups if group membership had been measured on an ordinal scale.

Method

Data on salary, age, years at the university, years of professional experience before joining the university, the year that a faculty member gained his/her current rank at the university, and current rank for 103 faculty members in a college of education at a large midwestern university was gathered. All of the variables are at least interval in nature except for faculty rank which was regarded as being ordinal. Information was copied directly from faculty files except for "years of professional experience before joining the university." For this variable, a faculty member, who is a full professor and has had many years of university service, gathered the information from the files but used his judgment as to whether the service was professional in nature. Following this, stratified random sampling was used to assign subjects into two groups. It was found that there were very few instructors ($\underline{n}$ = 3 in each of the two groups); consequently this group was dropped from further analyses.

Data on the first five interval variables served as predictors of the last variable, faculty rank, in a multiple regression analysis and discriminant analysis. For the multiple regression analysis, an assistant professor was assigned a "2"; an associate a "3"; and a full a "4". Following the generation of a regression equation for Group 1, the regression weights derived were used to predict the faculty status of the members of Group 2. The predicted rank "scores" were rounded to the nearest whole number and the correct matches to actual standing were then determined.

For the discriminant analysis, the discriminant weights and centroids of the three faculty ranks were determined for Group 1. The weights on the significant discriminant function were then applied to the data from Group 2 and a χ^2 was calculated for each person from each group centroid. The group from which the individual generated the smallest χ^2 was chosen as the person's predicted rank.

Results

The summary data on predictors and the correlations of those predictors with the criterion for each of the two groups are included in Table 1. As

TABLE 1

Summary Data of Predictor Variables and
Correlations with Criterion
(*n* Group 1 = 52; *n* Group 2 = 51)

Variable	Mean		SD		r	
	Group 1	Group 2	Group 1	Group 2	Group 1	Group 2
Salary	21,263.29	21,379.57	3,071.86	3,537.82	.64**	.76**
Age	49.17	46.45	9.63	9.24	.48**	.61**
Years at University	12.46	9.37	8.16	7.21	.56**	.57**
Years Experience Before University	10.85	11.31	6.68	4.93	.09	.30*
Year of Obtained Rank	70.85	71.73	5.41	5.06	-.27*	-.52**

*$p<.05$
**$p<.01$

can be seen from this table, all of the predictors were significantly linearly related to faculty rank except for the predictor "Years Experience Before the University". As can be seen from Table 2, the predictive power of the equation generated on Group 1 from the combination of all predictors was quite higl

TABLE 2

Group 1 Multiple Regression Results for Predicting Rank from Salary, Years at the University, Year of Rank, Previous Professional Experience, and Age

Variable	Final Beta	Multiple R	Step-wise F
Salary	.48	.64	34.74*
Years at University	.81	.69	5.98*
Year of Rank	.43	.73	6.63*
Previous Experience	.26	.75	2.50
Age	-.19	.75	.84

*p<.05

Table 3 shows the unstandardized discrimination function coefficients. Of the two functions derived, only the first function contributed significantly to group differentation.

TABLE 3

Unstandardized Coefficients for Both Discriminant Functions

Variable	Coefficients Function 1	Coefficients Function 2
Salary	-.00023	.00026
Years at University	-.12189	-.13250
Year of Rank	-.09206	-.18992
Previous Experience	-.05035	-.01183
Age	.02993	-.05181
Constant	11.93475	12.17555
λ	1.50	.05

Table 4 shows the percent of accurately assigned persons for each category. From this, it appears that the regression analysis produced the greatest number of correct placements.

TABLE 4

Percentages of Correct Matches for Group 2 of a Faculty Member's Predicted Rank and Actual Rank

Rank	n	Percent Matches Regression	Percent Matches Discriminant
Total	51	70.59	60.78
Full	12	41.67	83.33
Associate	21	85.71	38.10
Assistant	18	72.22	72.22

Discussion and Conclusions

One occasionally confronts a situation in which the criterion is measured on a scale which is neither clearly nominal nor clearly metric in nature. For example, if "letter grade" in a course is the criterion measurement which one wishes to predict from several measurements gathered on students, one could question if "letter grade" is sufficiently metric in nature so that multiple regression should be applied.

Based on this study, it appears that multiple regression would produce at least as fruitful results as discriminant function analysis. However, in this study the predictors were highly linearly related to the criterion measurement and this may have given the regression model an a priori advantage. For the discriminant analysis, only one significant linear function was derived. If this had not been the case, multiple regression would not have been fruitful but discriminant might have been.

This study seems to indicate that either statistical procedure can produce a high degree of predictive accuracy when predicting data measured on an ordinal scale. However, it is probable that this resulted from the fact that only a single linear discriminant function was significant and that there was a high linear relationship of the group of predictors with the criterion. Consequently, it is suggested that if there is relatively poor predictive power in predicting an ordinal criterion from multiple regression, one might consider employing discriminant analysis.

REFERENCES

Alumbaugh, R. V., & Crigler, M. A. Comparison of multivariate techniques in the prediction of juvenile postparole outcome. Educational and Psychological Measurement, 1978, 38, 96-106.

Bledsoe, J. C. A prediction of teacher competence: A comparison of two multivariate statistical techniques. Multivariate Behavioral Research, 1973, 8, 3-22.

Guilford, J. P., & Fruchter, B. Fundamental statistics in psychology and education. New York: McGraw-Hill Book Co., 1978.

Overall, J. E., & Klett, C. J. Applied multivariate analysis. New York: McGraw-Hill Book Co., 1972.

Tatsuoka, M. M. Multivariate analysis. New York: John Wiley & Sons, Inc., 1971.

Sonya L. Blixt
Kent State University

MULTIPLE LINEAR REGRESSION VIEWPOINTS
February, 1980 Volume 10, Number 2

A NOTE ON "A DEMONSTRATION OF A TYPE VI ERROR: AN APPLIED RESEARCH PROBLEM" BY STEVE ROLL *ET AL.*

Joseph P. Carbonari
University of Houston

The authors, while attempting to identify an instance of and at the same time extricate themselves from a Type VI error, have apparently fallen into an old, well-used, but still relevant trap--that of the probability of making at least one Type I error during the analysis of the data in one study, often called the experiment-wise or family-wise Type I error rate. If the hypotheses modeled were independent, and they are not, the probability of making at least one Type I error within the study would be .26 [$1 - (1 - \alpha)^k$, where k equals the number of hypotheses to be tested]. This value also serves as a lower bound for tests of related hypotheses, therefore, there is at least a one-in-four chance of making this error.

It would then follow that most of the identified gain in power is a function of covertly increasing the alpha and thus power. Although tables are not available, one could easily expect a power level for the full model, using 12 independent vectors, a medium effect size, and alpha set at "greater-than" .26 to be near that of their calculations for three vectors. These second calculations, by the way, estimate the power available to reject the omnibus null hypothesis, i.e., that the population R^2 of Y being predicted by vectors 1, 2, and 3 is zero. This is not an estimate of the power available for the rejection of the null hypotheses in each of the six models presented. Each model seems to represent a point-biserial correlation and if their combined contribution is an expected effect size of $R^2 = .20$, then on the average each of them would contribute .0333 to the total. This results in a power level for alpha = .05 of about .30 (Cohen & Cohen, 1975, p. 479)--not very satisfactory.

The 3 x 2 x 2 design given in Table 1 is incorrect for many reasons, two of which are: (1) it does not provide for the inclusion of the control group, although why it is perceived as being needed in the first place is in itself an interesting question; and (2) it is impossible to conceive of the variables as being fully crossed. The nature of the variables indicates that Scene Presentation can only be thought of as nested within method. One possible set of planned <u>orthogonal</u> contrasts which could, by present day agreement, be each tested at $\alpha = .05$, would be:

<u>Contrast</u>

1. T1 vs. t2 within A

   Does covert reinforcement differ from no covert reinforcement within the Wolpe procedure?

2. T7 vs. T8 within B

   Does covert reinforcement differ from no covert reinforcement within the Goldfried procedure?

3. $\frac{T1 + T2}{2}$ vs. $\frac{T7 + T8}{2}$

   Does the Wolpe treatment differ from the Goldfried procedure?

4. $\frac{T1 + T2 + T7 + T8}{4}$ vs. T11

   Does the average of these treatments differ from the modified Cautela procedure?

5. $\frac{T1 + T2 + T7 + T8 + T11}{5}$ vs. T13

   Does therapy differ from nontherapy? This could also be directional (one tailed).

Contrast coded regression models would provide correct tests.

These or a similar set of planned orthogonal comparisons would indeed maximize power while controlling the Type I error rate. As an aside, the authors must realize that if they were to analyze the hypotheses proposed in their paper, they would be restricting themselves to six distinct interpretations about six related hypotheses and could not pool the results into one overall interpretation as could be done if one model were used.

While it is easy to agree that the indiscriminate inclusion of all possible interactions is wrong, so is the use of multiple unprotected "t" tests (Cohen & Cohen, 1975, p. 162). The issue is real; traditional analysis <u>badly applied</u> often leads to the analysis of interactions or product variables which, if found to be statistically significant, would defy interpretation, but traditional analysis <u>correctly applied</u> would not only protect the researcher from Type VI errors but from the other tigers in the jungle. Perhaps we have come to the point where we can talk of the non-additivity of error types, e.g., the interaction of Type I and Type VI errors, or is it all just poorly conceptualized research and research design?

REFERENCES

Cohen, J., & Cohen, P. <u>Applied multiple regression/correlation for the behavioral sciences</u>. New York: John Wiley & Sons, 1975.

Roll, S., Hoedt, K. C., & Newman, I. A demonstration of a Type VI error: An applied research problem. <u>Multiple Linear Regression Viewpoints</u>, 1979, <u>10</u>(1), 31-38.

Joseph P. Carbonari
University of Houston

A COMMENT ON 'A NOTE ON A DEMONSTRATION OF A TYPE VI ERROR: AN APPLIED RESEARCH PROBLEM BY STEVE ROLL *ET AL*'

Steve Roll, Kenneth C. Hoedt, Isadore Newman
The University of Akron

We would first like to thank Professor Carbonari for the thorough reading and the interest reflected by his writing the note. The Viewpoints audience should be aware that one of the major purposes of MLRV is to disseminate information and clarify issues for applied statisticians and teachers of MLR. Notes like Dr. Carbonari's facilitate the achievement of this goal. In this vein and through this paper we address some of the concerns raised by Dr. Carbonari.

The first point raised in the Note was that the authors, in attempting to control for a Type VI error, made a Type I error. This comment, given the information presented in the original article, is justified. The example developed in the original article was to illustrate a Type VI error that could easily be made. In the actual research a Dunn's correction was used. (See Newman, 1972.) A second comment highly related to the first is more conceptually complicated. Carbonari makes the point that since there were nonorthogonal multiple comparisons the actual gain in power may have been due to the increase in the Type I error. As pointed out in discussing the first point, we did actually control for the Type I error but let's assume we did not. The examples presented assume the alpha level was .05 in both the traditional and the multiple regression approach that there would be a gain in power of .33 in the multiple regression approach since the ratio of subjects to groups increased due to limiting conceptualization to only those comparisons which were of real theoretical and logical interest to the researcher.

This brings us to the third point, in which Carbonari concludes the 3x2x2 design was inappropriate. In the initial article, the authors state on page 33 that this design was considered as a first approach and quickly abandoned because it was an inappropriate design which would result in a Type VI error. In addition, on page 34 the inference is that this is what might have happened if the researcher had used the 3x2x2 design (an inappropriate conceptualization) rather than what was recommended. The intent was to place emphasis on the possibility that someone using traditional analysis of variance in a cookbook fashion with a more subtle research question might make this kind of naive error and not be aware of it.

Point four made in the Note is that one can decrease the probability of making a Type I error without increasing the possibility of making a Type II error by making planned orthogonal comparisons. Again Carbonari is correct; however, the concern is that some of the questions reflected by the orthogonal comparisons may not be the research questions of interest--they may rather be questions formulated to fit a design. Therefore, this approach may lead to controlling for a Type I error but increase the probability of making a Type VI error. We agree that there are certain statistical advantages if one model is used, but since the emphasis of the paper was not on t-tests or controlling for multiple comparisons, a discussion of this was not included although we agree it should have at least been footnoted.

Another potential pragmatic problem associated with a one-model approach is that review editors frequently expect to see interaction reported regardless of whether it is of concern to the researcher and at times they become somewhat irrational in their demand for its presence.

On the first page of Carbonari's Note he suggests that identified gains in power may be a function of the alpha level in that using 12 independent vectors, a medium effect size, and an alpha level of .26 or greater would produce power analysis results comparable to those produced by the three vector models. This comment is of interest and may deserve further study. While the authors of the original paper did not feel this comment was directly related to their paper, they do feel the idea is an intriguing one and should be checked out.

We would like to once again thank Dr. Carbonari for his comments. They were appreciated.

REFERENCES

Carbonari, J. P. A note on "A demonstration of a Type VI error: An applied research problem by Steve Roll et al." Multiple Linear Regression Viewpoints, 1980, 10(2).

Roll, S., Hoedt, K. C., & Newman, I. A demonstration of a Type VI error: An applied research problem. Multiple Linear Regression Viewpoints, 1979, 10(1), 31-38.

Steve Roll, Kenneth C. Hoedt, & Isadore Newman

The University of Akron

ABSTRACTS FOR MULTIPLE LINEAR REGRESSION VIEWPOINTS: VOLUMES 6-9

Carolyn R. Benz, Ronald F. Bobner, Awilda Clemons
The University of Akron

VOLUME 6, No. 1

Jennings, E., & Ward, J. H., Jr. Logical steps in the creation and manipulation of fixed linear models. Multiple Linear Regression Viewpoints, 1975, 6(1), 2-7.

The authors present an 11-step procedure for creating linear models. The first four steps, the most challenging in their opinion, involve translating the research question into a symbolic expression of expected relationships, while the final steps are routine algebraic manipulations.

Karabinus, R. A., & McCormick, C. H. Comparison of regression coefficients in multivariate regression equations. Multiple Linear Regression Viewpoints, 1975, 6(1), 8-20.

The authors describe eight methods of comparing regression coefficients in multivariate regression equations containing the same variables for independent groups. The study investigated the relationship of the variables (Coopersmith Self-Esteem Inventory and Sarason's Test Anxiety Scale for Children) along with certain demographic data in predicting academic success among children in three ethnic groups. The F ratio of the SS residual ($F = \frac{(SSres./df)_1}{(SSres./df)_2}$) was considered the most fair and logical method.

Vasu, E. S., & Elmore, P. B. The effect of multicollinearity and the violation of the assumption of normality on the testing of hypotheses in regression analysis. Multiple Linear Regression Viewpoints, 1975, 6(1), 21-50.

This study investigated the effects of the violation of the assumption of normality coupled with the condition of multicollinearity upon the outcome of testing the hypothesis $\beta' = 0$ in the two-predictor regression equation. A monte carlo approach was utilized in which three different distributions were sampled for two sample sizes over 34 population correlation matrices. The preliminary results indicate that the violation of the assumption of normality has no significant effect upon the outcome of the hypothesis testing procedure. As was expected, however, the population correlation matrices with extremely high collinearity between the independent variables resulted in large standard errors in the sampling distributions of the standardized regression coefficients. Also, these same population correlation matrices revealed a larger probability of committing a type II error. Many researchers rely on beta weights to measure the importance of predictor variables in a regression equation. With the presence of multicollinearity, however, these estimates of population standardized regression weights will be subject to extreme fluctuation and should be interpreted with caution, especially when the sample size involved is relatively small.

McNeil, J. T. Regression analysis for repeated measures designs. Multiple Linear Regression Viewpoints, 1975, 6(1), 52-63.

Based on repeated measures designs using person vectors, this paper focuses on two concerns: (1) a proposed solution to the problem of missing data, and (2) the use of covariates as an alternative to person vectors in controlling for differences between individuals.

Edeburn, C. E., & Ochsner, D. P. STWMULTR: A computer program to expedite the retrieval of residual scores. Multiple Linear Regression Viewpoints, 1975, 6(1), 64-67.

Residual gain analysis was described in general terms and a new computer program, STWMULTR, designed to retrieve and punch residual scores was described. Samples of input and output data cards were included.

Williams, J. D., & Watson, J. G. The analysis of covariance with randomized blocks designs by regression. Multiple Linear Regression Viewpoints, 1975, 6(1), 68-73.

A regression solution is given for a research situation that includes both the analysis of covariance and randomized blocks. Basically, the solution includes the successive use of three linear models. The first model uses the covariate as the predictor while the second model uses both the covariate and the group membership variables; the difference (in R^2 units) between these two models is the proportion of the variance that is attributable to the group membership variables independent of the covariate. The third model includes the covariate, the group membership variables and the blocks. The difference (in R^2 units) between the third model and the second model is the proportion of the variance due to the blocks independent of both the group variables and the covariate.

Williams, J. D. A regression formulation of Dunn's and Scheffé's tests. Multiple Linear Regression Viewpoints, 1975, 6(1), 74-82.

Regression formulations of Dunn's and Scheffé's multiple comparison procedures are presented. The advantages and disadvantages of using the Dunn's, Scheffé's, Dunnett's, and Tukey's tests are explored.

Gillham, J., & Napady, D. Three reasons why percent variance accounted for is important to the development of theory. Multiple Linear Regression Viewpoints, 1975, 6(1), 83-89.

Percent variance accounted for describes the degree of ambiguity in a test of a theory. This percentage is a parsimonious statement of the relative success of each attempt to solve a particular puzzle; it is also a guide to forming still better solutions.

VOLUME 6, No. 2

Root, W., Newman, I., & Novak, E. The relationship between academic performance, test anxiety, race, sex, scholastic ability, and school organization: A multi-variable approach. Multiple Linear Regression Viewpoints, 1975, 6(2), 1-16.

The relationships between academic performance, test anxiety, race, sex, scholastic ability, and school organization were investigated. It was found that the scholastic ability variable was the most predictive factor of academic performance. When covarying the scholastic ability variable, initial differences favoring Caucasian students in graded schools for academic performance and test anxiety became nonsignificant but significant differences between the sexes remained for test anxiety. Caucasians, females, and those from graded schools scored significantly higher; however, when test anxiety was covaried, differences within school organizations became nonsignificant. Only linear significances were found, and all interactions were nonsignificant for the 206 students.

Duff, W., & Houston, S. Parental involvement in the education of their children. Multiple Linear Regression Viewpoints, 1975, 6(2), 17-34.

The investigators surveyed 621 educational professionals as to their perception of parental role involvement, and attempted to determine if the subgroups differed in these perceptions. The responses (yes or no) to each item were used to group the respondents into two clusters for each item. Role variables, district variables, and interaction variables were used as predictors of group membership.

VOLUME 6, No. 2, continued

Weber, D. C. The analysis of incomplete data using regression. Multiple Linear Regression Viewpoints, 1975, 6(2), 35-44.

He compares traditional ANOVA approach and multiple linear regression, pointing out the advantages of the latter. Specific examples are given.

VOLUME 6, No. 3

Jordan, T. E. Influences on preschool cognitive attainment. Multiple Linear Regression Viewpoints, 1975, 6(3), 1-108.

An analysis of measures of cognitive attainment, two at two years, one measure at age three years, two at age four years, and three at age five years is reported. In part one a multiple linear regression analysis examined the contribution of 12 variables to prediction of the eight criteria. In the second part of the analysis the most influential variables were explicated by maximizing their interactions in a second regression analysis. Criteria were the same eight cognitive tests at child ages two to five years. All data were developed through prospective longitudinal case studies begun at birth.

VOLUME 6, No. 4

Newman, I., Deitchman, R., Burkholder, J., & Sanders, R. Type VI error: Inconsistency between the statistical procedure and the research question. Multiple Linear Regression Viewpoints, 1976, 6(4), 1-19.

Type VI error is defined as inconsistency between the statistical procedure and the research question of interest. Six problems associated with Type VI error are explored and techniques for avoiding them are presented. Attention is focused on the impact that poor research has on the field of education.

Williams, J. D. Canonical analysis as a generalized regression technique for multivariate analysis. Multiple Linear Regression Viewpoints, 1976, 6(4), 20-38.

The use of characteristic coding (dummy coding) is made in showing solutions to four multivariate problems using canonical analysis. The canonical variates can be analyzed by the use of multiple linear regression. When the canonical variates are used as criteria in a multiple linear regression, the R^2 values are equal to θ, where θ is the squared canonical correlation coefficient. Several different methods exist for testing multivariate hypotheses. Where the interest is in a two-way disproportionate multivariate analysis of variance, the trace criterion ($\Sigma\theta_i$) seems particularly applicable.

McNeil, K., & Platt, J. Causal inference: Multiple linear regression vs. analysis of variance orthogonal and nonorthogonal designs. Multiple Linear Regression Viewpoints, 1976, 6(4), 39-41.

Artificial categorization of continuous variables and artificial orthogonalizations of correlated variables are discussed as limitations of analysis of variance's ability to make causal inferences. Noting that current methodology (1976) does not permit precise causal inferences using correlated predictors, several methods of limiting the problem are suggested.

Williams, J. D. Should a first course in ANOVA be through MLR? Multiple Linear Regression Viewpoints, 1976, 6(4), 42-45.

Practical and pedagogic concerns that need to be examined prior to deciding on a traditional or MLR approach to a first ANOVA course are presented. Guidelines are suggested concerning the extent to which an MLR approach should be oriented toward a direct translation of ANOVA type questions to MLR solutions.

McNeil, K. Position statement on the roles and relationships between stepwise regression and hypothesis testing regression. Multiple Linear Regression Viewpoints, 1976, 6(4), 46-49.

Hypothesis testing regression and stepwise regression are defined and their roles explained. The relationships between them are explored in terms of the kinds of data analyzed, shrinkage estimates, nonlinear terms, and causal inferences.

Newman, I. Brief note on the justification for using multiple linear regression. Multiple Linear Regression Viewpoints, 1976, 6(4), 50-52.

Noting that the F test is a least square solution and that Multiple Linear Regression is the general case of the least sum of squares solution, Newman presents seven justifications for the use of MLR.

VOLUME 7, No. 1

Williams, J. D. Multiple comparisons by multiple linear regression. Multiple Linear Regression Viewpoints, 1976, 7(1), 1-64.

Several of the more common multiple comparison techniques are explored in a regression approach. Dunnett's test for comparing several groups to a single group, Tukey's(a) honestly significant different test, Newman-Keul's, Tukey's(b), and Duncan's tests are considered. Complex comparisons (contrasts) are shown through Dunn's and Scheffé's tests and through orthogonal comparisons. Orthogonal polynomials are also shown for testing for trend. A method for finding a maximized Scheffé contrast such that the contrast will yield the same R^2 value as the original full model is also included.

The intent of the present monograph is to more fully explore the use of alternate methodologies to the usual multiple F tests when more than one restriction is placed on a full model.

Poynor, H. Spurious aggregation and the units of analysis. Multiple Linear Regression Viewpoints, 1977, 7(2), 1-11.

The author stresses that the choice of unit of analysis (pupil, school district, etc.) is rarely considered a serious issue but should not be ignored. Sampling a population can cause spurious outcomes. Aggregation of units of analysis is often done without thought of its effects. He describes a technique (defining the G variables) for determining sample heterogeneity.

Jennings, E. Comments on Poynor's paper. Multiple Linear Regression Viewpoints, 1977, 7(2), 12-13.

The author offers a critique of Poynor's paper. Claiming it is an abstract concept, he questions the utility of determining the unit of analysis by the G variable.

Poynor, H. Rejoiner to Jennings. Multiple Linear Regression Viewpoints, 1977, 7(2), 14-15.

Poynor clarifies his concept of the G variable. Besides assuring that the models reflect the research questions, the statistical features of data sets cannot be ignored.

Dalton, S. Regression approaches and approximate solutions to analysis of variance with disproportionality varied. Multiple Linear Regression Viewpoints, 1977, 7(2), 16-32.

The degree of nonorthogonality in a factorial design was systematically increased. Five methods of dealing with nonorthogonality were selected and applied: two were least squares solutions (Method 1 and Method 2); two were approximate solutions (the unweighted means analysis and the method of expected frequencies); and the fifth was the alternative of data elimination. Under extreme nonorthogonality all methods converged in yielding conclusions which while erroneous were similar across methods. Under moderate nonorthogonality, however, the unweighted means analysis and Method 1 were superior. Overall, the data elimination alternative was inferior in that it led to more type II errors than any of the other four methods.

Wolfle, L. M. Path analysis and causal models as regression techniques: A comment. Multiple Linear Regression Viewpoints, 1977, 7(2), 33-40.

The author comments on Williams and Kimpel's (1975) paper, "Path Analysis and Causal Models as Regression Techniques." He describes their incorrect designation of indirect effects saying that, in essence, three different effects were occurring. First, there was an indirect causal effect through intervening variables; second, a spurious association due to joint dependence on prior variables; and third, a correlation between predetermined variables.

Williams, J. D., & Klimpel, R. M. Path analysis: A comment on Wolfle's comment. Multiple Linear Regression Viewpoints, 1977, 7(2), 41-42.

The authors react to Wolfle's critique of their original work on path analysis, accepting his breakdown of their "indirect effect" categories into three classifications. They reiterate Wolfle's statement that for MLR practitioners the concept of path analysis can be an assist in writing models.

George, J. D. Multiple regression techniques applied to test the effects of three types of special class placement on the arithmetic achievement of educable mentally retarded pupils. Multiple Linear Regression Viewpoints, 1977, 7 (2), 43-61.

Multiple regression analysis was used to examine the different effects of special class placement on the arithmetic achievement of Educable Mentally Retarded (EMR) pupils. Self-contained classes, selected academic placement programs, and learning resource centers were the types of placement studied. A significant interaction between sex and type of placement was observed with respect to arithmetic achievement. Girls in self-contained classes gained more than boys in the same classes. Boys gained more in selected academic placement programs than in the other two types of placement; girls did best in selected academic placement programs.

VOLUME 7, No. 2, continued

Spaner, S. D. What inferences are allowable with a significant F in regression analysis? Multiple Linear Regression Viewpoints, 1977, 7(2), 62-74.

Spaner reviews the underlying assumptions of the F statistic and those underlying regression. He relates these to model testing in multiple linear regression and discusses inference limitations that can be made from outcomes in both a statistical and practical sense.

Dinero, T. E. An empirical example of the use of interaction terms in the multiple regression model. Multiple Linear Regression Viewpoints, 1977, 7(2), 75-100.

This study compared empirical results from an analysis of variance and a multiple linear regression solution when appropriate interaction terms were included in the regression model. A rationale for deciding which interaction terms should be included was presented.

VOLUME 7, No. 3

Newman, I., & Oravecz, M. T. Solutions to the problem of disproportionality: A discussion of the models. Multiple Linear Regression Viewpoints, 1977, 7(3), 1-51.

This paper has two major purposes. The first is to investigate the usefulness of a χ^2 technique in differentiating between varying degrees of disproportionality and their effects on a Type I error. The second purpose is to present and support the position that the major concern for any research model, whether disproportionate or not, is the research question and how well that question is reflected by the model. Three "exact solutions" for disproportional situations, the hierarchial, unadjusted main effects, and fitting constant methods, will also be discussed in terms of the research question that each reflects, and examples will be presented to demonstrate the most appropriate situation for using each solution.

Dalton, S. Shrinkage in R^2 and unbiased estimates of treatment effects using $\hat{\omega}$. Multiple Linear Regression Viewpoints, 1977, 7(3), 52-59.

The amount of variance accounted for by treatment can be estimated with $\hat{\omega}^2$ or with R^2 (symbolized as R_c^2 after a shrinkage formula has been applied). Monte Carlo methods were employed to compare $\hat{\omega}^2$, R_c^2, and R^2 in terms of bias and precision. R_c^2 and $\hat{\omega}^2$ produced estimates which were negligibly biased. The bias in R^2, while consistently positive, decreased as sample size increased and was too small to be of practical importance when $n \geq 50$. $\hat{\omega}^2$, R_c^2, and R^2 were all most precise with large samples and least precise when treatment effects were moderate in magnitude.

Walton, J. M. The use of multiple regression analysis in predicting success in the counseling practicum. Multiple Linear Regression Viewpoints, 1977, 7(3), 60-66.

The present exploratory study investigated the relationship between several predictor variables and the criterion of success in the counseling practicum among 93 recent graduates of a counselor education program. Forward stepwise regression was used. The investigation revealed that the best predictor of success in the counseling practicum was the square of the graduate grade point average ($ggpa^2$). This suggests the possibility of a curvilinear relationship between this predictor and the criterion. The interaction of female by Miller Analogies Test score (MAT) and the single variable of undergraduate grade point average (Ugpa) also appeared early in the equation. Type of undergraduate institution, type of graduate degree earned, and sex as a single independent variable demonstrated little relationship to the criterion.

Gantner, R. K., George, J. D., & Meadows, M. E. Relationships between results obtained on the Ertl machine and the Wechsler Intelligence Scale for Children (WISC). Multiple Linear Regression Viewpoints, 1977, 7(3), 67-83.

The purpose of this study was to examine relationships between the neural efficiency (NE), symmetry, and time difference (TD) scores on the Ertl machine and WISC scale scores for a group of 22 normal children and a group of 22 children with suspected learning disabilities, all ranging from 8 to 10 years of age. Multiple linear regression techniques were used to analyze the data. Some statistically significant relationships did occur between Ertl machine scores and WISC-V, WISC-P, and WISC-F scale scores for groups 1 and 2. Results supported Ertl's findings that normals and children with learning disabilities (LDs) would have similar NE scores (learning potential). Several symmetry scores (Hemispheric synchronization) and WISC scores correlated significantly in positive directions for both groups. Significant differences occurred between the TD scores (indicator of LDs) but results were in direct contrast to Ertl's claim since group 1 (normals) obtained higher mean scores than group 2.

VOLUME 8, No. 1

Williams, J. D. Full rank and non-full rank models with contrast and BJ binary coding systems for two-way disproportionate cell frequency analyses. Multiple Linear Regression Viewpoints, 1977, 8(1), 1-31.

The two-way non-orthogonal design has been a source of considerable controversy. Several recent publications have emphasized the full rank model solution and discouraged the use of the fitting constants solution, the hierarchical model and the unadjusted main effects solution. By using a cell coding system instead of an effects coding system, a full rank model different from that of Timm and Carlson (1975) is found: this model was first suggested by Jennings (1967). The second full rank solution can be found to be computationally identical to the unadjusted main effects solution.

Williams, J. D. A note on coding the subjects effect in treatments x subjects designs. Multiple Linear Regression Viewpoints, 1977, 8(1), 32-35.

Using a recent innovation described by Pedhazur (1977), a simpler regression solution to the repeated measure design is shown. Instead of coding N-1 vectors to represent the subject effect, the sum of each subject's criterion scores is entered as a vector. This single vector yields the same R^2 value as does the N-1 binary coded subject vectors.

Wolfe, L. M. An introduction to path analysis. Multiple Linear Regression Viewpoints, 1977, 8(1), 36-61.

An introduction to path analysis is posited. The manner in which causal effects can be decomposed is presented. This is followed by a discussion of some recent applications of path analysis to educational topics.

MLR/SIG Annual Meeting (1977). Minutes of annual meeting. Multiple Linear Regression Viewpoints, 1977, 8(1), 63-65.

The official minutes of the 1977 annual meeting of MLR/SIG are reproduced.

VOLUME 8, No. 2

Marquette, J. F., & Dufala, M. M. An interactive approach to ridge regression. Multiple Linear Regression Viewpoints, 1978, 8(2), 1-7.

The use of ridge regression is suggested as a method of limiting the problems caused by multi-collinearity of predictor variables in least squares solutions. An approach to choosing an appropriate ridge value is suggested. Example data are presented and an interactive computer solution (ADEPT) is included.

Walton, J. M., Newman, I., & Fraas, J. W. Ridge regression: A panacea? Multiple Linear Regression Viewpoints, 1978, 8(2), 8-15.

The technique of ridge regression is described along with its advantages and disadvantages. The authors conclude that while it may be an appropriate technique for some analyses, it may not be useful in instances where shrinkage estimates produce little shrinkage, or where the proportion of subjects to variables is sufficient.

Leitner, D. W. A teaching example of a replicable suppressor variable. Multiple Linear Regression Viewpoints, 1978, 8(2), 16-21.

The author describes a procedure to use in teaching the concept of suppressor variable to a statistics class. He uses the prediction of height, using weight and age (the suppressor). A brief review of the literature (and definition of) concerning suppressor variables is included.

Burkholder, J. H. An interactive version of MULR04 with enhanced graphics capability. Multiple Linear Regression Viewpoints, 1978, 8(2), 22-44.

A version of MULR04 employing random access Read/Write to simulate core memory for RT11 configured mini-computers is discussed. This version of MULR04 couples the flexibility of complex multiple regression with the interactive capability of the mini-computer. The program provides the user with the opportunity to enter data and regression models online while allowing examination of results and high quality graphics when desired.

Rakow, E. A. Ridge regression: A regression procedure for analyzing correlated independent variables. Multiple Linear Regression Viewpoints, 1978, 8(3), 1-17.

Ridge regression analysis is presented as a technique to be utilized in multiple linear regression situations when predictor variables are highly correlated. The presence of those variables impose several problems, the solutions to which are are described using ridge regression analysis. The advantages of ridge regression as well as its calculation are offered.

Hick, T. L., & Irvine, D. J. An analysis of the historical regression method of predicting posttest grade equivalents for categorically-aided programs. Multiple Linear Regression Viewpoints, 1978, 8(3), 18-26.

Historical Regression follows directly from the assumption that, without specific intervention, growth will continue at the rate (grade equivalents per year of schooling) obtained at the time of the pretest. When compared with program-level data ($n = 213$) it was found that Historical Regression underestimated final achievement for short programs with older children. It overestimated for younger children in long programs. An alternative method was developed which eliminated the bias, removed half of the error, and eliminated much computation since an expected achievement level for each child was not required.

Kukak, C. R., Levine, D. U., & Meyer, J. K. Neighborhood predictors of reading achievement in six big city school districts: A path analysis. Multiple Linear Regression Viewpoints, 1978, 8(3), 27-43.

The effects of neighborhood characteristics, i.e., race, socioeconomic status, family structure, and density on reading achievement is analyzed using path analysis. Two major hypotheses are analyzed and conclusions drawn from 1970 census sample data. An explanation is given for the selection of multiple regression path analysis.

Morse, P. K. Evaluation of sex-related salary discrimination. Multiple Linear Regression Viewpoints, 1978, 8(3), 44-50.

Using constructed data, the use of multiple regression is demonstrated for "School A," where salaries are fair but where women have been hired only recently, and for "School B," where there is evidence of sex-related bias in salary. The regression analysis identifies the presence or absence of salary bias, although mean salary by sex presents a different picture.

Martin, M. P., & Williams, J. D. Effects of state-wide salary equity provisions on institutional salary policies: A regression analysis. Multiple Linear Regression Viewpoints, 1978, 8(3), 51-65.

Presented is the process whereby a state-wide North Dakota faculty committee seeks to equalize salaries within higher education. Salary discrepancies between the eight North Dakota institutions of higher education prompted the committee to investigate salary inequities. The results are obtained using regression analysis. The impact of the equalization process on one institution's decision-making machinery is interpreted.

Vasu, E. S. The use of prediction intervals in multiple regression analysis. Multiple Linear Regression Viewpoints, 1978, 8(3), 66-81.

Using simulated data, an explanation is offered stressing the advantages of employing prediction intervals rather than predicting for individual cases. The three cases presented use classical regression analysis and vector notation to calculate prediction intervals. A discussion of results is included followed by an appendix with statistical program dataset manipulations.

Rosenthal, W., & Spaner, S. D. A study of three treatments for menstrual difficulties: An analysis using multiple linear regression. Multiple Linear Regression Viewpoints, 1978, 8(3), 82-105.

A regression approach is offered as an alternative procedure to traditional analysis of variance to investigate menstrual distress. Authors state objectives to facilitate the practitioner's understanding of regression solutions for a problem originally posed for analysis of variance solutions for three treatment groups. The tables in the appendix show that directly comparable results are obtained using regression and ANOVA.

Cohen, P. Selecting an appropriate model for data analysis. Multiple Linear Regression Viewpoints, 1978, 8(3), 106-115.

Comments are made on papers presented in this convention issue. The remarks refer to content concerns since the successful application of multiple regression hinges on the selection of relevant data. Each paper review includes a statement of strengths, areas of concerns, and suggestions for improvement.

VOLUME 9, No. 1

Fraas, J. W., & Newman, I. The malpractice of statistical interpretation. Multiple Linear Regression Viewpoints, 1978, 9(1), 1-25.

This paper examines problems that researchers may confront when interpreting statistical research results. The first section of the paper examines the problems associated with the use of gain scores. The second portion of the paper examines why the use of analysis of covariance is superior to the analysis of gain scores in aiding the researcher to avoid misinterpreting the data. The third section of the paper discusses the problem of disproportionality as it produces multicollinearity. The fourth section of the paper examines the difference between the interpretation of research results analyzed by part correlation as opposed to partial correlation. The final section presents a brief discussion of the effect of violating the assumption of rectilinearity in the regression effect.

House, G. D. A three-year ex post facto study of arithmetic achievement for elementary pupils eligible for a remedial arithmetic program. Multiple Linear Regression Viewpoints, 1978, 9(1), 26-48.

This study traced the three-year impact of a remedial arithmetic program on eligible St. Louis Public School pupils. Hypotheses were tested through multiple linear regression models for analyses of covariance. No treatment effects were found. The study reveals that changes in future program evaluation designs are needed.

Ryan, T. P. An approximation technique for variable selection using cost criteria. Multiple Linear Regression Viewpoints, 1978, 9(1), 49-56.

The problem of selecting regression variables using cost criteria is considered. A method is presented which approximates the global minimum of one of several criterion functions which might be employed. Examples are given and the results are compared with the results of other methods. The outcome of a simulation study is also discussed, and suggestions are made as to the practical use of the method.

Huitema, B. E. Univariate nonparametric analysis of variance through multiple linear regression. Multiple Linear Regression Viewpoints, 1978, 9(1), 57-62.

Many methodologists are aware that parametric tests associated with the analysis of variance and the analysis of covariance can be computed using regression procedures. It is shown that multiple linear regression can also be employed to compute the Kruskal-Wallis nonparametric analysis of variance.

Wolfle, L. M. Univariate nonparametric analysis of variance: A comment. Multiple Linear Regression Viewpoints, 1978, 9(1), 63-67.

The relationship between the Kruskal-Wallis H statistic and the multiple R^2 based on regressing ranks on k-1 dummy variables used to identify the groups is explored. A proof is presented and the utility of the regression approach over the traditional computation is considered.

Woehlke, P. L., Leitner, D. W., & Lewis, E. L. A defense of inferential statistics in education. Multiple Linear Regression Viewpoints, 1978, 9(1), 68-74.

Specific refutations to Brown's (1975) and Derrick's (1976) criticisms of inferential statistics and the techniques based on the general linear model are presented.

Huitema, B. E. A closer look at statistical independence, analysis of covariance and directional hypothesis. Multiple Linear Regression Viewpoints, 1978, 9(1), 75-80.

Statistical independence of observations, analysis of covariance, and directional hypothesis are discussed in regard to the inferences that are allowable with a significant F in regression analysis.

Jordan, T. E. On the comparability of multiple linear (MULR-05) and interaction (AID-4) regression techniques. Multiple Linear Regression Viewpoints, 1978, 9(1), 81-89.

Interaction regression and multiple linear regression were compared by analyzing sample data composed of developmental measures on children (N=196). The techniques were compared to see if they identified the same sources of variance and produced comparable R^2 values.

VOLUME 9, No. 2

Williams, J. D. Path analysis from a regression perspective. Multiple Linear Regression Viewpoints, 1978, 9(2), 1-81. (Monograph Series #3)

This monograph presents path analysis to the presumably naive reader who is, on the other hand, a practitioner of multiple linear regression techniques. The major methodological process, recursive structural models, is presented and structural equations are defined, relating these to multiple regression. Sample data sets are used to present practical applications of path analysis to educational research.

Williams, J. D. Contrasts with unequal N by multiple linear regression. Multiple Linear Regression Viewpoints, 1979, 9(3), 1-7.

It is shown that some of the more simplified methods for contrasts with equal N result in erroneous calculations when applied to data sets with unequal N. Instead, the methodology given earlier by Bottenberg and Ward (1963) is effective for finding values for contrasts (where $t = \sqrt{F}$). Also, the unweighted means solution for maximized Scheffé contrasts is shown to fail in finding the maximized contrast with unequal N.

Lewis, E., & Leitner, D. Is the PhD research tool used in the dissertation? Multiple Linear Regression Viewpoints, 1979, 9(3), 8-10.

Students taking Multiple Regression as the PhD research tool from 1970 through 1975 tended to use Multiple Regression as the data analytic tool in the dissertation.

Mouw, J. T., & Nu, V. Increasing power and interpretability in certain repeated measures designs. Multiple Linear Regression Viewpoints, 1979, 9(3), 11-28.

Repeated measures designs offer a relatively powerful procedure for the analysis of behavioral data. In these designs, research questions involve the change of individuals' patterns of responses across time or across a dimension with intervening treatment effects. The addition of one or more between-subject factors allows for the comparison of treatment effects across the repeated measures between groups of subjects. In most of these researches, the grouping variable has been obtained by arbitrarily dichotomizing a continuous variable. This article presents an alternative analysis of data of certain repeated measures designs where the variable is kept in its natural continuous state instead of being dichotomized. Such an analysis is argued to have two advantages: (a) a more realistic interpretation of the results, and (b) a tendency toward an increase in power in the F tests of the repeated dimension and its interaction.

House, G. D. Effects of different types of scores on magnitudes of computed R^2. Multiple Linear Regression Viewpoints, 1979, 9(3), 29-36.

This study compared the magnitudes of R^2 values computed through multiple linear regression models using grade equivalent scores versus raw scores, standard scores, and percentiles as both criterion and predictor variables. It was found that grade equivalent and standard score modes produced similar and higher R^2 values than did raw scores or percentiles.

Fraser, B. J. A multiple regression model for research on teacher effects. Multiple Linear Regression Viewpoints, 1979, 9(3), 37-52.

A description is given of a model for research on teacher effects in which the variance in student outcome posttest performance is attributed to pretest performance, to separate construct domains of student, instructional, and teacher variables, and to interactions between variables in these three construct domains. When the model was applied with a sample of 780 Australian seventh grade pupils, it was found that pretest, an instructional variable, a block of teacher variables, a block of instruction-student interactions, and a block of instruction-teacher interactions were each significant independent predictors of a student attitudinal posttest.

Newman, I., & Thomas, J. A note on the calculation of degrees of freedom for power analysis using multiple linear regression models. Multiple Linear Regression Viewpoints, 1979, 9(3), 53-58.

This note presents 15 examples worked by Cohen in which he uses different formulas to calculate degrees of freedom, depending on the power analysis situation. It is then demonstrated that the same results can be obtained by using a more general formula for calculating degrees of freedom. It was felt that this information may be of pedagogical value.

VOLUME 9, No. 3, continued

AERA Annual Meeting (April 1979): SIG on Multiple Linear Regression. Multiple Linear Regression Viewpoints, 1979, 9(3), 59.

A list of papers presented and their authors is provided.

VOLUME 9, No. 4

Newman, I., & Fraas, J. Some applied research concerns using multiple linear regression. Multiple Linear Regression Viewpoints, 1979, 9(4), 1-49. (Monograph Series #4)

The authors present an examination of the advantages of multiple linear regression as a tool for educational researchers. Concerns for multicollinearity and upward bias and disproportionality R^2 are discussed. Factor regression, component regression, and ridge regression are also discussed.

VOLUME 9, No. 5

McNeil, K., Evans, J., & McNeil, J. Nonlinear transformation of the criterion. Multiple Linear Regression Viewpoints, 1979, 9(5), 1-9.

The utility of a nonlinear transformation of the criterion is established. A well-known law from a field other than education is used as the example to demonstrate the point. The functional relationships may be such (as in the Pythagorean Theorem) that an R^2 of 1.00 cannot be found without making a nonlinear transformation of the criterion. The goal of predictability (R^2 = 1.00) thus may not be reached without making a nonlinear transformation of the criterion.

Clegg, A. A., Jr., Prichard, K., & Weigand, P. Multiple regression as a technique for predicting college enrollment. Multiple Linear Regression Viewpoints, 1979, 9(5), 10-19.

This paper deals with the application of multiple linear regression to the problem of identifying appropriate criterion variables and predicting enrollment in college courses during a period of major rapid decline. Data were gathered on course enrollments for 1972-1978 and organized around five criterion variables. Total college enrollment proved to be the best single predictor with correlations of .89 to .99 with each of 10 departmental course enrollments. The technique has proved to be 96 to 100% accurate in estimating course enrollments in seven of the 10 courses. It is also a valuable means for data-based decision making and long-range planning when faculty committees must advise on administrative decisions.

Wolfe, L. M. Unmeasured variables in path analysis. Multiple Linear Regression Viewpoints, 1979, 9(5), 20-56.

The author discusses measurement error in structural equation models, a potential influence on the explanation of educational phenomena. The author first describes the case of a causal model with a single unmeasured variable: intergenerational occupational mobility from father's socioeconomic status to respondent's educational attainment. Educational attainment in the example is presented as the unobserved variable. Secondly, a more complex example incorporating several unmeasured variables for which multiple indicators were available in a similar situation is presented. A computer program, LISREL, is offered to deal with the latter situation.

Newman, I., Seymour, G. A., & Garver, T. K. A Monte Carlo evaluation of estimated parameters of five shrinkage estimate formuli. Multiple Linear Regression Viewpoints, 1979, 9(5), 57-74.

This study employs a Monte Carlo simulation to determine the accuracy with which the shrinkage in R^2 can be estimated by five shrinkage formuli and cross-validation. The study dealt with the use of shrinkage and cross-validation for different sample sizes, different R^2 values, and different degrees of multi-collinearity.

Williams, J. D., & Wali, M. K. Missing cells and a curious case of degrees of freedom. Multiple Linear Regression Viewpoints, 1979, 9(5), 75-87.

An experimental sampling procedure for plant communities on surface mined areas yielded missing cells and caused a further problem of yielding a "total" number of degrees of freedom equal to N rather than the usual N-1. The discrepancy occurred because the degrees of freedom are not necessarily additive for all missing cell designs. A solution which may circumvent this problem is proposed.

Mouw, J. T., & Vu, N. V. Increasing power and interpretability in certain repeated measures designs. Multiple Linear Regression Viewpoints, 1979, 9(5), 88-106.

Repeated measures designs offer a relatively powerful procedure for the analysis of behavioral data. In these designs, research questions involve the change of individuals' patterns of responses across time or across a dimension with intervening treatment effects. The addition of one or more between-subject factors allows for the comparison of treatment effects across the repeated measures between groups of subjects. In most of these researches, the grouping variable has been obtained by arbitrarily dichotomizing a continuous variable. This article presents an alternative analysis of data of certain repeated measures designs where the variable is kept in its natural continuous state instead of being dichotomized. Such an analysis is argued to have two advantages: (a) a more realistic interpretation of the results, and (b) a tendency toward an increase in power in the F tests of the repeated dimension and its interaction.

Carolyn R. Benz, Ronald F. Bobner, & Awilda Clemons

The University of Akron

ANNUAL MEETING SIG/MULTIPLE LINEAR REGRESSION OF AERA, BOSTON 1980

MULTIPLE LINEAR REGRESSION

SPECIAL INTEREST GROUP MEETING

Thursday, April 10, 1980
4:05 pm - 6:05 pm

Sheraton Hotel - Constitution Room, 2nd Floor

Chair: Bill Connett
Office of Public Instruction
Helena, Montana

Presentations:

"Multivariate Profile Analysis: A Technique for Identifying Moderator Variables in Multiple Regression." David A. Ludwig and Warren D. Dolphin, Iowa State University

"Evaluating Title I Early Childhood Programs: Problems, the Applicability of Model C, and Several Evaluation Plans." Keith A. McNeil and Emily A. Findlay, NTS Research Corporation

"Budget Allocations at MSU: A Linear Regression Policy Capturing Analysis." William Simpson, Michigan State University; Steven Spaner, University of Missouri-St. Louis

"Predictive Validity of Subset Regression Procedures: An Empirical Comparison." Marian L. Thompson, Bureau of the Census, Washington, D.C.

Discussants: Judy McNeil, National Testing Service

Tony Eichelburger, University of Pittsburgh

Chair-elect: Lee Wolfle, Virginia Polytechnic Institute

If you are submitting a research article other than notes or comments, I would like to suggest that you use the following format, as much as possible:

Title

Author and affiliation

Indented abstract (entire manuscript should be single 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.)

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, single-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.

"A publication of the *Multiple Linear Regression Special Interest Group* of the American Educational Research Association, published primarily to facilitate communication, authorship, creativity, and exchange of ideas among the members of the group and others in the field. As such it is not sponsored by the American Educational Research Association nor necessarily bound by the Association's regulations.

"Membership in the *Multiple Linear Regression Special Interest Group* is renewed yearly at the time of the American Educational Research Association Convention. Membership dues pay for a subscription to the *Viewpoints* and are divided into two categories: Individual = $5.00; and Institutional (libraries and other agencies) = $18.00. Membership dues and subscription requests should be sent to the Executive Secretary of the MLRSIG."

THE UNIVERSITY OF AKRON
AKRON, OH 44325

AKRON OHIO U.S. POSTAGE .53 PB METER 672269

TABLE OF CONTENTS

ISSN 0195-7171