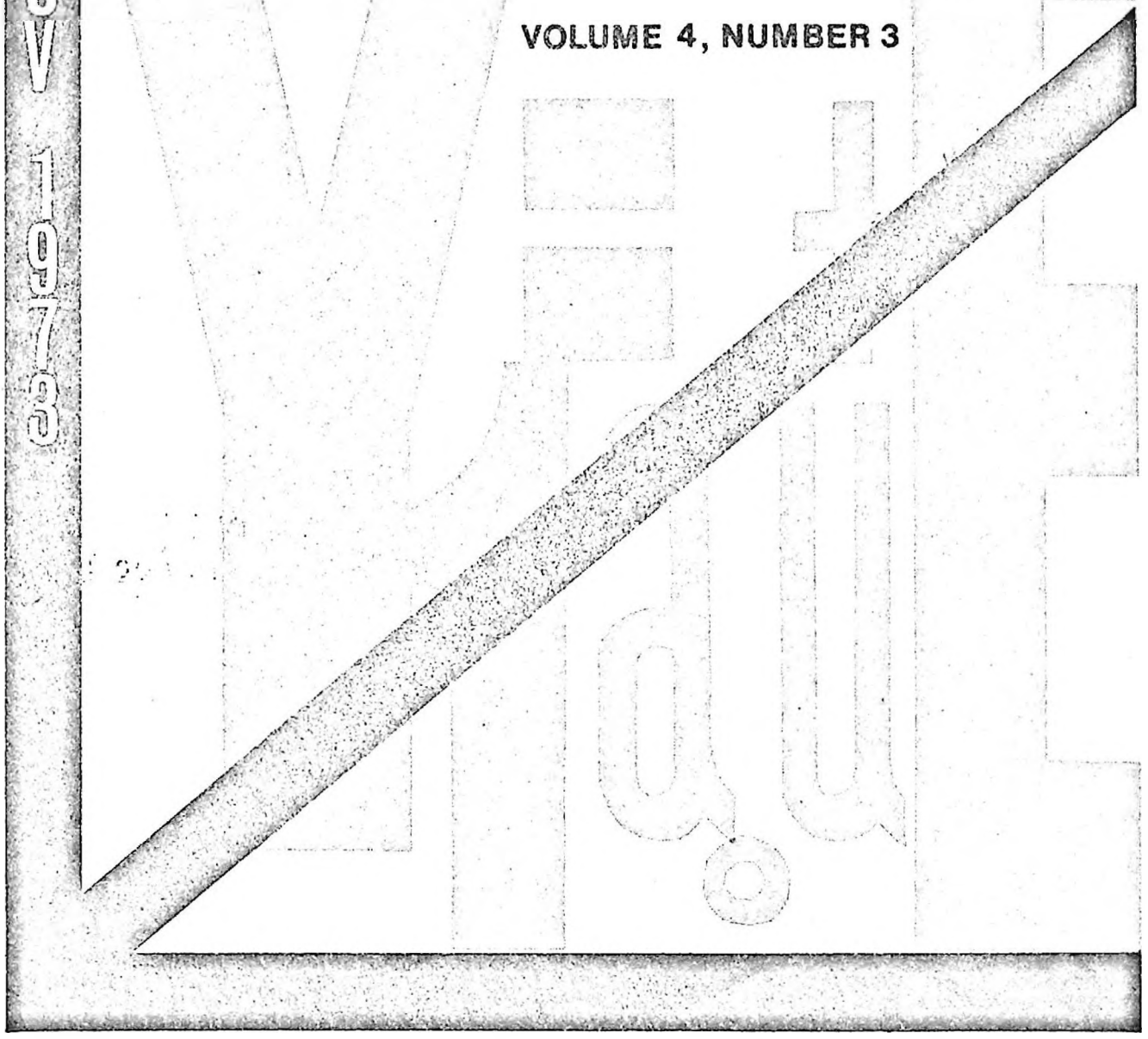
NOV 1973

VOLUME 4, NUMBER 3



MULTIPLE LINEAR REGRESSION VIEWPOINTS

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

If you are submitting a research 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
Single-spaced indented abstract
Introduction (purpose—short review of the literature, etc.)
Method
Results
Discussion (conclusion)
References using APA format

All manuscripts should be sent to the editor at the above address.

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 4, No. 3 1973

THE ANALYSIS OF SPLIT-PLOT AND SIMPLE HIERARCHICAL DESIGNS USING MULTIPLE LINEAR REGRESSION

Thomas W. Pyle

Eastern Washington State College

For those researchers who are interested in using multiple linear regression (MLR) as a general data analytic technique, it is of some importance to recognize what types of models and restrictions on models generate F ratios that correspond to classical hypothesis testing procedures. Many of these equivalences have already been demonstrated for a variety of experimental designs (e.g., Kelly, Beggs, McNeil, Eichelberger, and Lyon, 1969; Williams, 1970; Jennings, 1967), but some appear to have been overlooked. One of these is the split-plot factorial design and another is the completely randomized hierarchical design (Kirk, 1968). What will be demonstrated here is that MLR models can be constructed for these designs which generate F ratios that are equivalent to those obtained by traditional computing formulas. An example of each of these follows.

Split-plot Factorial

A x B Interaction and B Main Effect

In this design n subjects are randomly assigned to each of p levels ($p \geq 2$) of a between-subjects variable (A). Each subject's performance is measured at q time periods (i.e., there are q levels of the within-subjects, B, variable). A total of pnq dependent variable scores are obtained. The following data are from Myers (1972, p. 200).

		B_1	B_2	B_3
A_1	S_1	7	1	7
	S_2	9	2	10
A_2	S_3	11	6	7
	S_4	16	14	9

Consider the following regression model:

$$Y_1 = a_o U + a_1 X_{1,1} + a_2 X_{1,2} + a_3 X_{1,3} + a_4 X_{2,1} + a_5 X_{2,2} + a_6 X_{2,3} + b_1 P_1 + b_2 P_2 + b_3 P_3 + b_4 P_4 + E_1 \quad [1]$$

where:

Y_1 is the vector of dependent variable scores;

U is the unit vector:

$X_{j,k}$ is a vector which contains a one if the corresponding value in Y_1 was made by a person in treatment condition A_jB_k, zero otherwise;

P_i is a vector which contains a one if the corresponding value in Y_1 was made by person i (i = 1, 2, 3, 4,) zero otherwise;

$a_o, a_1, \ldots, b_4$ are weighting coefficients, the values of which are chosen so as to minimize the sum of the squared values of E_1; and

E_1 is the error vector.

It can be shown that when the following restrictions ($\mathcal{R}_1$) are imposed on Model 1,

$$\mathcal{R}_1: \quad \frac{(a_o + a_1 + b_1 + a_o + a_1 + b_2)}{2} - \frac{(a_o + a_4 + b_3 + a_o + a_4 + b_4)}{2} = k$$

$$\frac{(a_o + a_2 + b_1 + a_o + a_2 + b_2)}{2} - \frac{(a_o + a_5 + b_3 + a_o + a_5 + b_4)}{2} = k$$

$$\frac{(a_o + a_3 + b_1 + a_o + a_3 + b_2)}{2} - \frac{(a_o + a_6 + b_3 + a_o + a_6 + b_4)}{2} = k$$

Model 2 obtains:

$$Y = a_o U + c\ (X_{1,1} + X_{1,2} + X_{1,3}) + a_4\ (X_{1,1} + X_{2,1}) + a_5\ (X_{1,2} + X_{2,2}) + a_6\ (X_{1,3} + X_{2,3}) + b_1 P_1 + b_2 P_2 + b_3 P_3 + b_4 P_4 + E_2 \qquad [2]$$

where $c = 1/2\ (2k - b_1 - b_2 + b_3 + b_4)$.

The restrictions may look somewhat complicated, but they are actually rather straightforward. In terms of Model 1 the expression in the first set of parentheses is the average predicted value for treatment combination A_1B_1: $a_o + a_1 + b_1$ is the predicted value for Person 1 at A_1B_1 and $a_o + a_1 + b_2$ is the predicted value for Person 2 at A_1B_1. The meaning of the remaining expressions is completely analogous and should be obvious. It may be of interest to note that if one imposes the set of restrictions $\mathcal{A}'_1$ on Model 1, Model 2 also obtains:

$\mathcal{A}'_1$: $a_1 - a_4 = k$

$a_2 - a_5 = k$

$a_3 - a_6 = k$

When the least squares solutions of Models 1 and 2 are obtained, $R_1^2 = .95150$ and $R_2^2 = .74788$. An F ratio may be computed using the following formula (Bottenberg and Ward, 1963):

$$F = \frac{(R_1^2 - R_2^2) \,/\, df_1}{(1 - R_1^2) \,/\, df_2}$$

where:

df_1 = the number of linearly independent vectors in Model 1 (m_1) minus the number of linearly independent vectors in Model 2 (m_2);

df_2 = the dimension of the vectors (N) minus m_1.

R_1^2 = the squared multiple correlation coefficient of the Full model.

R_2^2 = the squared multiple correlation coefficient of the Restricted model.

In Model 1 there are 8 linearly independent vectors and in Model 2 there are 6 linearly independent vectors. Since N = 12, $df_1 = 2$ and $df_2 = 4$. The obtained value of F is 8.3967.

The restrictions imposed on Model 1 generate an F ratio that is equivalent to that obtained by traditional computing formulas for the test of "no interaction" (Myers, 1968). Note also that the df's are the same as those obtained by classical "rules of thumb": $df_1 = (p-1)(q-1)$, $df_2 = p(n-1)(q-1)$.

The restrictions on Model 1 that generate the F ratio associated with the B main effect are as follows:

$$\mathcal{R}_2: \quad \frac{a_o + a_1 + b_1 + a_o + a_1 + b_2}{2} + \frac{a_o + a_4 + b_3 + a_o + a_4 + b_4}{2} = k$$

$$\frac{a_o + a_2 + b_1 + a_o + a_2 + b_2}{2} + \frac{a_o + a_5 + b_3 + a_o + a_5 + b_4}{2} = k$$

$$\frac{a_o + a_3 + b_1 + a_o + a_3 + b_2}{2} + \frac{a_o + a_6 + b_3 + a_o + a_6 + b_4}{2} = k$$

Imposing $\mathcal{R}_2$ on Model 1 (by solving for a_1, a_2, and a_3) yields Model 3:

$$Y = a_oU + c(X_{1,1} + X_{1,2} + X_{1,3}) + a_4(X_{2,1} - X_{1,1}) + a_5(X_{2,2} - X_{1,2}) \\ a_6(X_{2,3} - X_{1,3}) + b_1P_1 + b_2P_2 + b_3P_3 + b_4P_4 + E_3 \qquad [3]$$

where $c = 1/2\ (2k - b_1 - b_2 - b_3 - b_4 - 4a_o)$.

The R^2 obtained from Model 3 is .70909, and when Model 3 is compared to Model 1 using the F ratio as defined above, F = 9.9962 with $df_1 = 2$ and $df_2 = 4$. This is the same F value, df_1, and df_2 obtained by traditional computing formulas for the B main effect.

A Main Effect

A problem occurs when one attempts to compute the A main effect. It may seem reasonable to generate from Model 1 a restriction which forces the average predicted value for A_1 to equal the average predicted value for A_2. However, this procedure does not work. It is easy to see that it will not work because whatever restriction one imposes on Model 1 df_2 will always be 4, whereas df_2 must be $(n-1)p = 2$ if the MLR procedure is to be strictly equivalent to the traditional procedures. However, consider Model 4:

$$Y_2 = a_oU + a_1A_1 + a_1A_2 + E_4 \qquad [4]$$

where:

Y_2 is a vector of criterion scores. The ith element of Y_2 is the sum (or average) of the q scores in the original data matrix for Person i, i = 1, 2, 3, 4;

U is the unit vector;

A_j contains a one if the corresponding element in Y_2 was obtained by a person in treatment A_j, zero otherwise;

a_o, a_1, and a_2 are the least squares weights; and

E_4 is the error vector.

If the restriction $a_1 = a_2$ is imposed on Model 4, the unit vector model obtains. This model has an associated R^2 of 0.0. By comparing the R^2 of Model 4 (.58273) to an R^2 of 0.0 using the F ratio, one obtains an F value of 2.7931, where $df_1 = 1$ and $df_2 = 2$. This result is identical to that obtained by traditional computing procedures when computing F for the A main effect.

A Hierarchical Model

Groups Within Treatments

In the type of hierarchical design considered here, p levels of the treatment variable (A) are administered to groups of Ss. At each level of A there are g groups; each group is composed of n Ss. It should be noted that Groups is not factorial to Treatments because the g groups at treatment level j are not the same groups at treatment level k.

In the usual ANOVA procedure, the Treatment Mean Square is divided by an error term called Groups within Treatments (G/A). However, under certain conditions (Myers, 1972, Chap. 11) the significance of Treatments is assessed relative to a pooled error term. Both procedures can also be accomplished with the framework of MLR.

The following data matrix is from Myers (1972, p. 233).

A_1			A_2		
b_{11}	b_{21}	b_{31}	b_{12}	b_{22}	b_{32}
5	7	16	24	9	17
6	18	5	21	23	26
18	4	9	12	28	24
12	11	14	16	19	19
41	40	44	73	79	86

Following the standard procedure for the assessment of significance of treatment effects (i.e., G/A is considered the proper error term), we define Model 5:

$$Y_3 = a_0U + a_1A_1 + a_2A_2 + E_5 \qquad [5]$$

where:

Y_3 is a vector of sum scores. The elements of Y_3 are the sums of the individual subject scores in the six groups (i.e., the dimension of Y_3 is 6):

U is the unit vector;

A_j contains a one if the corresponding element in Y_3 came from a group belonging to treatment level j; zero otherwise;

a_o, a_1, and a_2 are least squares weights; and

E_5 is the error vector.

If the restriction $a_1 = a_2$ is imposed on Model 5, the unit vector model obtains. The R^2 for Model 5 is .95799. If Model 5 is compared with the unit vector model using the F ratio as defined previously, the resulting value of F is 91.208 with $df_1 = 1$ and $df_2 = 4$. One obtains the same F, df_1, and df_2 using traditional formulas (the value of F that Myers reports in his Table 9-4 is wrong due to arithmetic and round-off errors.)

Subjects within Groups within Treatments

One may also determine if the variability of groups within treatments contributes significantly to the total variability. To do this Model 6 is created:

$$Y_4 = a_oU + a_1G_{1,1} + a_2G_{2,1} + a_3G_{3,1} + a_4G_{1,2} + a_5G_{2,2} + a_6G_{3,2} + E_6 \quad [6]$$

where:

Y_4 is a vector of criterion scores. Each element of Y_4 is the score obtained by an individual S;

U is the unit vector;

$G_{k,j}$ contains a one if the corresponding element in Y_4 was made by a subject in Group k, Treatment j (k = 1, 2, 3; j = 1, 2), zero otherwise;

a_o, a_1,..., a_6 are least squares weights; and

E_6 is the error vector.

If the following restrictions are made,

$$a_1 = a_2 = a_3$$

$$a_4 = a_5 = a_6$$

Model 7 obtains:

$$Y_4 = a_0U + a_1A_1 + a_2A_2 + E_7 \qquad [7]$$

where:

> A_j contains a one if the corresponding element in Y_4 was made by a subject of Treatment level j (j = 1, 2), zero otherwise.

Comparing Model 7 to Model 6 via the F ratio yields a value of F of 0.1679 with $df_1 = 4$, $df_2 = 18$. The values of F, df_1, and df_2 obtained with the MLR procedure are identical to those obtained with traditional computing formulas.

When the Groups within Treatments source of variation is not significant at a fairly high significance level (e.g., $\alpha = .25$) and if the experimenter has no a priori reason to suspect that groups within treatments should be a significant source of variation, then some authors suggest a pooling of error terms for the evaluation of the treatment variable (Kirk, 1968; Myers, 1972). One effect of pooling is to increase the number of degrees of freedom in the denominator of the F test; this may increase the power of the test.

The MLR procedure for pooling for the hierarchical design under consideration is as follows: Model 7 is considered to be the appropriate full model. The restriction on Model 7 is $a_1 = a_2$. Imposing this restriction on Model 7 yields the unit vector model, and the associated value of F in the comparison of the two models is 18.046 with $df_1 = 1$, $df_2 = 22$.

What has been presented represents another area of equivalence of MLR and traditional statistical analysis. Formulation of these and other designs using MLR procedures should help emphasize what starting conditions (full models) are assumed and what restrictions on these starting conditions are sufficient to produce F ratios and df's that are the same as those obtained using traditional computing formulas.

References

Bottenberg, R. A., & Ward, J. H. Applied multiple linear regression. Lackland Air Force Base, Texas: Aerospace Medical Division, AD 413128, 1963.

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

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

Kelley, F. J., Beggs, D. L., McNeil, K. A., Eichelberger, T., & Lyon, J. Research design in the behavioral sciences: Multiple regression approach. Carbondale, Illinois: Southern Illinois University Press, 1969.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks/Cole, 1968.

Myers, J. L. Fundamentals of experimental design. (2nd ed.) Boston: Allyn and Bacon, 1972.

Williams, J. D. A regression approach to experimental design. Journal of Experimental Education, 1970, 39(1), 83-90.

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 4, No. 3 1973

AN EMPIRICAL COMPARISON OF RESIDUAL GAIN ANALYSIS AND THE ANALYSIS OF COVARIANCE

Ronald C. Buzahora and John D. Williams

The University of North Dakota

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

Two similar techniques for statistical analysis, residual gain analysis and the analysis of covariance, have been mentioned by different researchers as useful measures of change. Perhaps because of their similarity, applied researchers have sometimes mistakenly concluded that the two techniques were in fact identical. Jennings (1972) contrasted the linear models for the two different analyses; Williams and Others (1972) presented limited data for comparing two groups at each of six grade levels, using the two measures of change. In both of the mentioned presentations, a regression approach was used to clarify similarities and dissimilarities between the two approaches. The present research effort has been directed toward a more extensive comparison of the two techniques.

Method

The sample for this investigation was selected from students enrolled at Carl Ben Eielson and Nathan Twining Elementary Schools in Grand Forks, North Dakota, who participated in a project entitled Human Awareness Through Self Enhancing Education. These schools were located adjacent to the Grand Forks Air Base, and served the children of those military personnel

living on or near this establishment. Students in grades three through eight from each school participated in this study. They were tested initially in the Fall (October) and retested using the same instruments near the close of the school year (April). Only those students who participated in the study for the entire treatment period were included in the data analysis. The number of students involved were: Grade 3, N = 119; Grade 4, N = 73; Grade 5, N = 103; Grade 6, N = 58; Grade 7, N = 95; and Grade 8, N = 80. The instruments used in this study were: The Self Appraisal Inventory (SAI), the School Sentiment Index (SSI) and the Iowa Test of Basic Skills (ITBS). Separately at each grade level, subgroups were formed (High, Middle, Low) on the composite score on each of the three tests. Using the high, middle and low groups on the ITBS, each of the 15 subtests on the ITBS, the 5 subtests of the SSI and the 6 subtests of the SAI were analyzed using the analysis of covariance and the residual gain analysis using the pre-test on the corresponding subtest as the covariate. A similar procedure was used, finding the high, middle and low groups on the composite score for the SAI and comparing the two techniques on each subtest of the SAI and SSI. Finally, each subtest on the SSI was analyzed using the high, middle and low groupings on the SSI composite pre-test score.

At a given grade level, the analysis of covariance was accomplished by the use of two linear models:

[1] $Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + e_1$, (Model 1) and

[2] $Y = b_0 + b_1 X_1$. (Model 2)

where

Y = the score on a given post-test score,

b_0 = the regression coefficient for the particular model (b_0 for Models 1 and 2 are in general different),

X_1 = the pre-test score,

X_2 = 1 if a score is from a member of the high group; 0 otherwise,

X_3 = 1 if a score is from a member of the middle group; 0 otherwise,

$b_1 - b_3$ = are regression coefficients, and

e_1 and e_2 are the error in prediction respectively for Model 1 and Model 2.

The use of Model 1 allows the finding of R^2; this value is denoted here as R^2_1. Similarly, the second model yields R^2_2. The test of significance is given by

$$[3] \quad F = \frac{(R^2_1 - R^2_2)/2}{(1 - R^2_1)/df_w},$$

where df_w is the degrees of freedom for within for a particular analysis. The use of equation 3 assumes homogeneous slopes of the regression lines for each group; if it is desired to test this assumption, Jennings' article contains the necessary test.

The residual gain analysis can be accomplished with two linear models, the first of which is Model 2. The residual values from Model 2 become the criterion scores for the following model:

$$[4] \quad Y' = b_0 + b_2X_2 + b_3X_3 + e_3, \qquad \text{(Model 3)}$$

where

Y' = the residual from Model 2 for each score, and

e_3 = the error in prediction with Model 3.

This model also yields a squared multiple correlation coefficient R^2_3.

Results

Tables 1-3 report the summarized results of using the groupings made on the ITBS composite score for the subtests of the ITBS, the SAI and the SSI respectively. Tables 4-5 report the summarized results of using the groupings

made on the SAI composite score for the subtests of the SAI and SSI respectively. Table 6 reports the summarized results of using the groupings made on the SSI composite score for the subtests of the SSI. Table 7 summarizes the results of Tables 1-6. While the original analyses were made separately at each grade level, the tables combine the results from the six grades. Each table includes the number of significant results found by each method; also included is an indication of a comparison of $R_1^2 - R_2^2$ to R_3^2. In Table 4, $R_1^2 - R_2^2$ was equal to R_3^2 three times; in Table 5, they were equal twice.

TABLE 1

SUMMARY OF SIGNIFICANCE LEVELS FOR THE IOWA TEST OF BASIC SKILLS SUBTEST RESULTS--GROUPED (HIGH, MIDDLE AND LOW) ON THE IOWA TEST OF BASIC SKILLS COMPOSITIVE SCORE

Level	Analysis of Covariance	Residual Gain Analysis
Nonsignificant	22	87
Significant at .05 but not at .01	11	2
Significant at .01	57	1
Higher R^2 value ($R_1^2 - R_2^2$ to R_3^2)	75	15

TABLE 2

SUMMARY OF SIGNIFICANCE LEVELS FOR THE SELF APPRAISAL INVENTORY SUBTEST RESULTS--GROUPED (HIGH, MIDDLE AND LOW) ON THE IOWA TEST OF BASIC SKILLS COMPOSITIVE SCORE

Level	Analysis of Covariance	Residual Gain Analysis
Nonsignificant	20	29
Significant at .05 but not at .01	7	1
Significant at .01	3	0
Higher R^2 value (R^2_1 - R^2_2 to R^2_3)	20	10

TABLE 3

SUMMARY OF SIGNIFICANCE LEVELS FOR THE SCHOOL SENTIMENT INDEX SUBTEST RESULTS--GROUPED (HIGH, MIDDLE AND LOW) ON THE IOWA TEST OF BASIC SKILLS COMPOSITIVE SCORE

Level	Analysis of Covariance	Residual Gain Analysis
Nonsignificant	35	32
Significant at .05 but not at .01	1	4
Significant at .01	0	0
Higher R^2 value (R^2_1 - R^2_2 to R^2_3)	12	24

TABLE 4

SUMMARY OF SIGNIFICANCE LEVELS FOR THE SELF APPRAISAL INVENTORY SUBTEST RESULTS--GROUPED (HIGH, MIDDLE AND LOW) ON THE SELF APPRAISAL INVENTORY COMPOSITIVE SCORE

Level	Analysis of Covariance	Residual Gain Analysis
Nonsignificant	29	29
Significant at .05 but not at .01	0	1
Significant at .01	1	0
Higher R^2 value (R^2_1 - R^2_2 to R^2_3)	19	8

TABLE 5

SUMMARY OF SIGNIFICANCE LEVELS FOR THE SCHOOL SENTIMENT INDEX SUBTEST RESULTS--GROUPED (HIGH, MIDDLE AND LOW) ON THE SELF APPRAISAL INVENTORY COMPOSITIVE SCORE

Level	Analysis of Covariance	Residual Gain Analysis
Nonsignificant	32	33
Significant at .05 but not at .01	4	3
Significant at .01	0	0
Higher R^2 value ($R^2_1 - R^2_2$ to R^2_3)	7	27

TABLE 6

SUMMARY OF SIGNIFICANCE LEVELS FOR THE SCHOOL SENTIMENT INDEX SUBTEST RESULTS--GROUPED (HIGH, MIDDLE AND LOW) ON THE SCHOOL SENTIMENT INDEX COMPOSITIVE SCORE

Level	Analysis of Covariance	Residual Gain Analysis
Nonsignificant	30	34
Significant at .05 but not at .01	5	2
Significant at .01	1	0
Higher R^2 value ($R^2_1 - R^2_2$ to R^2_3)	21	15

TABLE 7

SUMMARY OF SIGNIFICANCE LEVELS FOR ANALYSIS OF COVARIANCE AND RESIDUAL GAIN ANALYSIS ON TABLES 1-6

Level	Analysis of Covariance	Residual Gain Analysis
Nonsignificant	168	244
Significant at .05 but not at .01	28	13
Significant at .01	62	1
Higher R^2 value ($R^2_1 - R^2_2$ to R^2_3)	154	99

Discussion

Only Tables 1 and 2 show any marked differences in significance levels, both showing a greater degree of significance with the analysis of covariance. On the other hand, when $R_1^2 - R_2^2$ is compared to R_3^2 (the amount of accounted variance by the group membership variables in each instance), Tables 3 and 5 favor the residual gain analysis, with the remaining tables favoring the analysis of covariance. The major difference in the success of the two approaches occurs in Table 1 in which the ITBS composite scores on the pre-test were used for classification into the high, middle and low groups for analyzing the 15 subtests of the ITBS on the post-test. The effect of this grouping is to create a treatments X levels design; for the covariance analysis, the level was being covaried, in the sense that the grouping was based on ITBS composite scores; for the residual gain analysis, a treatments X levels design was executed on the residual gain data. For the covariance analysis, the adjusted means generally followed the pattern where in the high group was highest, the middle group was in the middle, with the low group being the lowest. No general pattern seemed to emerge for the residual gain analysis. Interestingly, the pattern for the covariance analysis held <u>least</u> when the criterion was the composite score.

Coming to conclusions in the present analysis is, at the least, tenuous. Considering the conceptual and computational difficulties attendent to the residual gain analysis, there is a tendency to dismiss this technique in deference to the analysis of covariance, or perhaps some other measure of change. The results of Table 5 are strong enough to give the residual gain analysis at least one reprieve. While there are undoubtedly good reasons to choose the residual gain analysis over all other methodologies

for some applications, the situations that would best be analyzed by the residual gain method would seem to be severely limited.

References

Jennings, E. Linear models underlying the analysis of covariance, residual gain scores and raw gain scores. Paper presented at the annual convention of the American Educational Research Association, April 7, 1972.

Williams, J. D., Maresh, R. T., and Peebles, J. D. A comparison of raw gain scores, residual gain scores and the analysis of covariance with two modes of teaching reading. Paper presented at the annual convention of the American Educational Research Association, April 7, 1972.

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 4, No. 3 1973

JUDGMENT ANALYSIS AND PORNOGRAPHY

Judith A. Houston and Samuel R. Houston
University of Northern Colorado

ABSTRACT

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

INTRODUCTION

For nearly as long as man has existed there has been confusion and concern for what is pornographic or obscene. This confusion has greatly increased in recent years in the United States because of vague and poorly defined decisions handed down by the Supreme Court. To the layman these decisions often appear to be contradictory and confusing in nature.

A technique for both capturing and clustering raters' policies, Judgment Analysis (JAN), was suggested by Bottenberg and Christal (1961) and Christal (1963). JAN utilizes a multiple regression model and a hierarchical grouping technique which clusters raters on the basis of the homogeneity of their raw score regression vectors (their prediction equations).

The primary purpose of this study was to determine if a policy-capturing methodology (JAN) which has been successfully utilized in military and educational research could be adapted for use as a procedure in identifying pornographic material. This attempt to use JAN techniques was considered significant in that it represented a first application of this procedure to an area of investigation which appears of concern today.

METHOD

Sample: The Ss were 28 doctoral Psychology, Counseling and Guidance (PCG) majors at the University of Northern Colorado, 21 lawyers and 25 police officers, both from the city of Greeley, Colorado. These individuals were identified by use of a table of random numbers and were requested to serve as judges on a voluntary basis.

Instrument: The instrument contained an 11-item profile for each of 100 pictorial representations. The reason for using 11 items on each profile and having 100 profiles available for each judge stems from the research of Dudycha (1970) who found that it was difficult to cluster judges or raters when fewer than 100 observations exist. Further, Dudycha found that the clustering technique is far more effective when at least 10 predictor or profile items are used.

A sample of the profiles used in this study appears below. Each of the 100 profiles was generated by using a table of random numbers. This guaranteed maximum variability in the distribution of profile descriptors, which is essential for the valid use of JAN. In addition, the use of the table of random numbers also guaranteed low intercorrelations between the profile variables. This feature was important for a clean interpretation of policy in the policy analysis stage of the study.

SAMPLE PROFILE OF PICTORIAL REPRESENTATION

Descriptors	Yes	No
1. Appeals to prurient interest		X
2. Goes beyond contemporary community standards		X
3. Has redeeming social importance	X	
4. Involves uncovered adult human genitalia		X
5. Displays human sexual intercourse		X
6. Displays masturbation		X
7. Displays anal intercourse	X	
8. Displays flagellation or torture	X	
9. Displays homosexuality		X
10. Involves interracial sexual relationships		X
11. Displays group sex activity		X

Judge's Rating (Circle only one) 1 2 3 4 5

Code

1. Should be banned from society
2. Allowed only for professional (medical and/or scientific) study
3. Available only to interested adults
4. Available to all adults
5. Unrestricted use in society

Procedure and Design: The investigators contacted a total of 74 individuals in the three groups. These individuals were requested to serve as judges for the study and each was presented with a set of 100 profiles descriptive of pictorial representations which were generated for this study by the investigators. Each judge was requested to rate each of the profiles on a five-point scale.

The JAN procedure was applied to the results of the ranking process to ascertain the feasibility of the process as well as to determine the number of policies which were present in each of the three groups.

The JAN procedure consisted of two basic stages. In the first stage, a least-squares solution of a multiple regression equation was computed for each judge, which predicted the criterion decision he had made. This determined how consistent the judge was in his use of specific variables in arriving at an over-all decision by giving his decision-making equation and the R^2, the square of the multiple correlation

coefficient from his multiple regression analysis. Those judges whose R (multiple R value or predictive efficiency value which is a measure of the judge's intrarater consistency) value in their multiple linear regression equation exceeded.70 were deemed to be making consistent policy and were included in the JAN analysis.

A hierarchical grouping of those judges who were identified as expressing a consistent policy (R value in excess of .70) was undertaken in order to determine the number of policies which were present. The investigators inspected the stages of the grouping according to the Ward and Hook (1963) recommendation of looking for a break in the objective function (over-all R^2 in this case). A determination of whether one or more judgmental systems are present among the judges was made for each of the three groups of judges by looking for a drop of .05 or more in the R^2 between successive clusters of judges in the category of judges under consideration. This procedure enabled the investigators to specify the number of policies present in each group. The researchers used multiple linear regression according to Ward (1962) to determine the significance of the unique contribution of each predictor variable ($X_1, X_2, \ldots\ldots, X_{11}$) and other specific subsets of variables in order to explain the captured policies.

The unique contribution of a variable to prediction may be measured by the difference in the square of the multiple correlation coefficient, R^2, for the regression model in which all predictor variables are used and the R^2 for a regression model in which the variable under consideration has been removed. The first model is called the full model (FM), and the second is called the restricted model, (RM). The R^2 for the RM can never be larger than the R^2 for the FM. The difference between the two R^2s may be tested for statistical significance with the variance ratio test, or the investigator may set a level drop which she is willing to accept. The former course of action was chosen and the significance level was set at .05.

RESULTS

The means and standard deviations for the profiles are given in Table 1. A variable received a score of 1 if the characteristic was present and a 0 if the characteristic was absent.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF PREDICTOR (PROFILE) VARIABLES

(N = 100)

	Variable	Abbrev-iation	Mean	Standard Deviation
1.	Appeals to prurient interest	API	0.50	0.50
2.	Goes beyond contemporary community standards	BCS	0.54	0.50
3.	Has redeeming social importance	RSI	0.48	0.50
4.	Involves uncovered human adult genitalia	UAHG	0.49	0.50
5.	Displays human sexual intercourse	HSI	0.42	0.49
6.	Displays masturbation	DM	0.44	0.50
7.	Displays anal intercourse	DAI	0.56	0.50
8.	Displays flagellation or torture	DFT	0.47	0.50
9.	Displays homosexuality	DH	0.47	0.50
10.	Involves interracial sexual relationships	ISR	0.46	0.50
11.	Displays group sex activity	GSA	0.62	0..49

Table 2 presents the intercorrelations between predictor or profile variables. An inspection of the intercorrelations in Table 2 shows that the vast majority of them are quite low.

TABLE 2

INTERCORRELATIONS OF PREDICTOR (PROFILE) VARIABLES

Variable		1	2	3	4	5	6	7	8	9	10
1.	API										
2.	BCS	.00									
3.	RSI	.04	.08								
4.	UAHG	-.06	.02	.02							
5.	HSI	.04	.05	.07	-.15						
6.	DM	.08	.01	-.17	-.10	.27					
7.	DAI	-.08	.11	.13	-.06	.26	.18				
8.	DFT	.02	.11	-.19	-.12	-.16	.09	-.13			
9.	DH	.06	-.10	-.02	.08	.01	.09	.07	.03		
10.	ISR	-.08	.09	.08	-.06	.03	.07	-.11	.05	-.15	
11.	GSA	.00	.02	.05	-.02	.46	.24	.14	.00	.08	-.02

PCG Judges: Of the 28 PCG judges contacted, 26 of the judges were used in the analysis. Two had to be eliminated because of incomplete responses. Nine of the PCG students had very specific policies and were not included in the JAN procedure. These nine judges included two judges who checked column 5 for each of the 100 profiles. This policy represented the position that the identified profiles should have unrestricted use in

society. One judge checked column 4 on each of the 100 profiles--this policy represented the position that the profile materials should be available to all adults. Finally, 6 judges checked column 3 which was a statement that the profile materials should be made available only to interested adults.

The remaining 17 judges who had expressed complex policies (i.e., they checked more than one column) were included in the JAN procedure to determine how many different policies were present. In this JAN analysis, nine judges had to be eliminated because their multiple R value was less than .7. This low R value suggested that the judges were not expressing a consistent policy.

In Table 3 are presented the R^2 and R values for the eight PCG judges who were retained in the final JAN analysis. The values of R range from a high of .9646 for Judge 5 to a low of .7586 for Judge 6.

TABLE 3

R^2 AND R VALUES FOR PCG JUDGES EXPRESSING COMPLEX POLICIES

Judge	R^2 value	Multiple R value
1.	.9116	.9547
2.	.5889	.7673
3.	.7642	.8741
4.	.6736	.8207
5.	.9305	.9646
6.	.5755	.7586
7.	.7501	.8660
8.	.6712	.8192

Table 4 gives the stages for the JAN procedure for the group of eight judges and the R^2 value associated with each stage. For example, in Stage 1 the eight judges remain ungrouped and the overall R^2 value for this system is .8416. The first two judges who are grouped together are Judges 2 and 5. They have been identified as having the most homogeneous policy between two judges in the system. The overall R^2 dropped only .0088 in going from Stage 1 to Stage 2. In Stage 6, three separate judgmental groups of judges have been formed. Judges 1,2,5,6,7,8 form one group; Judge 3 forms a second group; and Judge 4 forms the third group.

At Stage 5, the R^2 has dropped .0588 from Stage 4. Thus, by the criterion adopted in this study, Stage 4 gives the clustering of judges into five policy systems which are characteristic of the eight judges. Judges 1, 6, and 7 form the first system; Judges 2 and 5 form the second system; Judge 3 forms the third system; Judge 4 forms the fourth system; Judge 8 forms the fifth system.

TABLE 4

STAGES OF THE JAN PROCEDURE FOR EIGHT PCG JUDGES

Stage	Judges	R^2	Successive Drop in R^2	Collective Drop in R^2
1	1,2,3,4,5,6,7,8	.8416	--------	--------
2	1,(2,5),3,4,6,7,8	.8328	.0088	.0088
3	(1,7),(2,5),3,4,6,8	.8033	.0295	.0383
4	(1,6,7),(2,5,)3,4,8	.7701	.0332	.0715
5	(1,2,5,6,7)3,4,8	.7113	.0588	.1303
6	(1,2,5,6,7,8)3,4	.6279	.0834	.2137
7	(1,2,3,5,6,7,8)4	.4861	.1418	.3555
8	(1,2,3,4,5,6,7,8)	.1610	.3251	.6806

For an explanation of the five PCG judgmental policies, a subjective analysis of the predictor variables was made. It was conjectured that they formed a hierarchical pattern as displayed in Table 5. The groupings identified in this table were used in Table 6 in which the unique contribution of specific subset of variables was tested for statistical significance in explaining individual policies. The five PCG policies displayed in Table 6 were determined by the following grouping of judges: Policy I--Judges 1,6, and 7; Policy II--Judges 2 and 7; Policy III--Judge 3; Policy IV--Judge 4; and Policy V--Judge 8.

TABLE 5

SUBJECTIVE HIERARCHY OF PREDICTOR (PROFILE) VARIABLES

Impact Variables:	Appeals to prurient interest	(1)
	Goes beyond contemporary community standards	(2)
	Has redeeming social importance	(3)
Content Variables:		
	Natural:	
	Involves uncovered human genitalia	(4)
	Displays human sexual intercourse	(5)
	Displays masturbation	(6)
	Involves interracial sexual relationships	(10)
	Unnatural:	
	Displays anal intercourse	(7)
	Displays flagellation or torture	(8)
	Displays homosexuality	(9)
	Displays group sex activity	(11)

TABLE 6

FIVE PCG POLICY TABLE

Policy	I	II	III	IV	V
Judges	1,6,7	2,5	3	4	8
	Unique Contribution	Unique Contribution	Unique Contribution	Unique Contribution	Unique Contribution
Impact Variables (1-3)	.0499*	.5078*	.0107	.5837*	.0617*
Variable 1 API	.0487*	.0004	.0067	.0082	.0004
Variable 2 BCS	.0000	.0033	.0018	.3061*	.0050
Variable 3 RSI	.0003	.4934*	.0031	.3063*	.0576*
Content Variables (4-11)	.3288*	.0706*	.7210*	.0197	.4940*
Natural Content(4,5,6,10)	.0327*	.0200*	.0090	.0065	.0128
Variable 4 UAH G	.0161*	.0028	.0003	.0017	.0049
Variable 5 HSI	.0038	.0007	.0008	.0035	.0031
Variable 6 DM	.0082*	.0042	.0071	.0024	.0062
Variable 10 ISR	.0046	.0106*	.0005	.0000	.0006
Unnatural Content (7,8,9,11)	.2928*	.0411*	.5907*	.0160	.4247*
Variable 7 DAI	.0152*	.0221*	.5848*	.0006	.0053
Variable 8 DFT	.2706*	.0197*	.0000	.0117	.4150*
Variable 11 GSA	.0014	.0021	.0001	.0038	.0000

*Significant beyond the .05 level

Policy I was concerned with both the subset of impact variables and the subset of content variables. Specifically, Variable 1--Appeals to purient interest,

Variable 4--Involves uncovered human genitalia, Variable 6--Displays masturbation, Variable 7--Displays anal intercourse, Variable 8--Displays flagellation or torture, and Variable 9--Displays homosexuality were important considerations for Judges 1,6,7, who expressed Policy I.

Policy II again was concerned with both the subset of impact variables and the subset of content. The specific variables involved or attended to by Judges 2 and 5 who authored Policy II were different from those identified in Policy I. The specific significant variables identified in Policy II include Variable 3--Has redeeming social importance, Variable 10--Involves interracial sexual relationships, Variable 7--Displays anal intercourse, and Variable 8--Displays flagellation or torture.

The third policy expressed by PCG doctoral students, Policy III, was determined by a single judge (Judge 3). Policy III was not concerned with the impact variables but rather focused on the subset of content variables. The one specific variable which was significant in this policy was Variable 7--Displays anal intercourse.

Policy IV also was determined by a single judge--Judge 4. This policy was concerned with the subset of impact variables and not the subset of content variables. Specifically, Variable 2--Goes beyond contemporary community standards, and Variable 3--Has redeeming social importance, were important considerations for this judge.

The final and fifth policy which was captured by the JAN procedure, Policy V, was determined by a single judge (Judge 8) as was the case for Policy III and Policy IV. The policy identified both impact and content variables as important considerations. The specific variables attended to in this policy include Variable 3--Has redeeming social importance, and Variable 8--Displays flagellation or torture.

Lawyer Judges: Of the 21 lawyers in the city of Greeley, Colorado who volunteered to participate, 19 were used in the analysis. One judge had to be eliminated because the secretary volunteered that she had filled out the instrument instead of the lawyer. Still another was not usable because of incomplete items on the instrument. Of the 19 usable cases, nine of the lawyer judges had very specific policies and hence were not included in the JAN procedure. Four lawyers stated that their policy relative to the profiles was that all the profile materials should be made available only to interested adults, while three lawyers felt that the profile materials should be made available to all adults. Only two of the lawyers checked column five which indicated that all the materials should have unrestricted use in society.

The remaining ten lawyer judges who expressed complex policies were included in the JAN procedure. In this JAN anlaysis, six of the lawyer judges had to be eliminated because their multiple R value was less than .7.

In Table 7 are presented the R^2 and the R values for the four lawyer judges who were retained in the final JAN analysis. The values of R range from a high of .8425 for Judge 3 to a low of .7061 for Judge 4.

TABLE 7

R^2 AND R VALUES FOR LAWYER JUDGES EXPRESSING COMPLEX POLICIES

Judge	R^2 value	Multiple R value
1.	.6284	.7927
2.	.5628	.7502
3.	.7098	.8425
4.	.4986	.7061

Table 8 gives the stages for the JAN procedure for the group of four lawyer judges and the R^2 value associated with each stage. By the criterion adopted in this study, there are four different policies being expressed by the four lawyer judges as the R^2 drop from Stage 1 to Stage 2 was .0547 which exceeded the .05 criterion.

TABLE 8
STAGES OF THE JAN PROCEDURE FOR FOUR LAWYER JUDGES

Stage	Judges	R^2	Successive Drop in R^2	Collective Drop in R^2
1	1,2,3,4	.8205	-------	-------
2	(1,4)2,3	.7658	.0547	.0547
3	(1,3,4),2	.6242	.1416	.1963
4	(1,2,3,4)	.1439	.4803	.6766

For an explanation of the four captured lawyer policies, refer to Table 9. The unique contributions of specific variables are tested in an attempt to explicate the different policies.

Policy I--expressed by Judge 1--focused on the subset of content variables while ignoring the subset of impact variables. The specific content variables which were significant in Policy I include Variable 6--Displays masturbation, and Variable 8--Displays flagellation or torture. On the other hand, Policy II--determined by Judge 2--was concerned with both subsets of impact and content variables. Important variables for this policy are Variable 1--Appeals to prurient interest, Variable 3--Has redeeming social importance, Variable 5--Displays human sexual intercourse, Variable 10--Involves interracial sexual relationships, and Variable 8--Displays flagellation or torture. The third lawyer judge policy, Policy III, was again determined by a single judge (Judge 3). This policy included both subset of impact and content variables as significant. Specifically, Variable 1--Appeals to prurient interest, Variable 2--Goes beyond contemporary community standards, Variable 4--Involves uncovered adult human genitalia, Variable 7--Displays anal intercourse, and Variable 8--Displays flagellation or torture, were identified in the policy. The last policy captured, Policy IV, involved a single judge. Both impact and content variables were identified as important in this policy but the specific variables identified were different from the other three policies. The significant specific variables are Variable 1--Appeals to prurient interest, Variable 2--Goes beyond contemporary community standards, and Variable 8--Displays flagellation or torture.

TABLE 9
FOUR LAWYER POLICY TABLE

Policy	I	II	III	IV
	Unique Contribution	Unique Contribution	Unique Contribution	Unique Contribution
Judge	1	2	3	4
Impact Variable (1-3)	.0136	.2071*	.0561*	.2628*
Var. 1 API	.0025	.0277*	.0205*	.1389*
Var. 2 BCS	.0051	.0026	.0245*	.1182*
Var. 3 RSI	.0073	.1895*	.0119	.0014
Content Variables (4-11)	.6149*	.2437*	.6894*	.1724*
Natural Content Variables (4,5,6,10)	.0780*	.0780*	.0552*	.0197
Var. 4 UAHG	.0000	.0076	.0290*	.0107
Var. 5 HSI	.0008	.0200*	.0071	.0031
Var. 6 DM	.0727*	.0092	.0082	.0060
Var. 10 ISR	.0010	.0324*	.0118	.0013
Unnatural Content Variables (7,8,9,11)	.4697*	.1695*	.6023*	.1420*
Var. 7 DAI	.0059	.0000	.4624*	.0178
Var. 8 DFT	.4266*	.1579*	.1880*	.1244*
Var. 9 DH	.0127	.0063	.0009	.0007
Var. 11 GSA	.0127	.0014	.0101	.0015

*Significant beyond the 0.05 level

Police Judges: While 25 police officers were contacted only 15 volunteered to participate in the study. One return was not usable as the respondent failed to complete all the items on the instrument. Four of the members of the Greeley Policy Department who participated in the study had a very specific policy. One felt that all the profile materials should have unrestricted use in society; another believed that the materials should be available to all adults; two officers expressed the position that the profile materials should be allowed only for professional study. The remaining 10 police judges who had expressed complex policies were included in the JAN procedure. In this JAN analysis, five of the police judges had to be eliminated because their multiple R values were less than .7.

In Table 10 are presented the R^2 and R values for the five police judges who were retained in the final JAN analysis. The values of R range from a high of .9484 for Judge 3 to a low of .7407 for Judge 1.

TABLE 10
R^2 AND R VALUES FOR POLICE JUDGES EXPRESSING COMPLEX POLICIES

Judge	R^2 value	Multiple R value
1.	.5487	.7407
2.	.7928	.8904
3.	.8965	.9484
4.	.8575	.9260
5.	.6774	.8230

Table 11 gives the stages for the JAN procedure for the group of five police judges and the R^2 value associated with each stage. There are four different policies present as the R^2 drop from Stage 2 to Stage 3 was .0508 which exceeded the .05 cutoff criterion.

TABLE 11
STAGES OF THE JAN PROCEDURE FOR FIVE POLICE JUDGES

Stage	Judges	R^2	Successive Drop in R^2	Collective Drop in R^2
1	1,2,3,4,5	.8628	------	------
2	1,(2,4),3,5	.8300	.0328	.0328
3	(1,2,4),3,5	.7792	.0508	.0836
4	(1,2,4(,(3,5)	.6272	.1520	.2356
5	(1,2,3,4,5)	.1084	.5188	.7544

Table 12 indicates the four policies captured by the JAN technique.

TABLE 12
FOUR POLICE POLICY TABLE

Policy	I	II	III	IV
Judge	2,4	1	3	5
	Unique Contribution	Unique Contribution	Unique Contribution	Unique Contribution
Impact Variables (1-3)	.1252*	.0196	.8399*	.0308*
Var. 1 API	.0000	.0123	.0027	.0045
Var. 2 BCS	.0041	.0025	.8255*	.0224*
Var. 3 RSI	.1229*	.0057	.0004	.0051
Content Variables (4-11)	.2115*	.5006*	.0090	.6752*
Natural Content Variables (4,5,6,10)	.0075	.0566*	.0078	.0244
Var. 4 UAHG	.0053	.0040	.0020	.0227*
Var. 5 HSI	.0001	.0020	.0038	.0010
Var. 6 DM	.0021	.0465*	.0001	.0001
Var. 10 ISR	.0002	.0056	.0002	.0021
Unnatural Content Variables (7,8,9,11)	.1982*	.3550*	.0029	.6387*
Var. 7 DAI	.0000	.0461*	.0005	.3749*
Var. 8 DFT	.1821*	.0485*	.0001	.3269*
Var. 9 DH	.0002	.2435*	.0004	.0029
Var. 11 GSA	.0051	.0011	.0017	.0000

*Significant beyond 0.05 level

Policy I--formulated by Judges 2 and 4--emphasizes both the impact and content subset of variables. The important specific variables include Variable 3--Has redeeming social importance and Variable 8--Displays flagellation or torture. Policy II was determined by Judge 1 and it focused on the content variables while ignoring the subset of impact variables. Variable 6--Displays masturbation, Variable 7--Displays anal intercourse, Variable 8--Displays flagellation or torture, and Variable 9--Displays

homosexuality, were important contributors to this policy. The third policy captured from the police judges, Policy III, was expressed by Judge 3. This policy ignored the content variables and emphasized the subset of impact variables with specific emphasis on Variable 2--Goes beyond contemporary community standards. The final policy--Policy IV(expressed by Judge 5)--identified both the impact and content subsets as important. Specifically, Variable 2--Goes beyond contemporary community standards, Variable 4--Involves uncovered adult human genitalia, Variable 7--Displays anal intercourse, and Variable 8--Displays flagellation or torture, are important contributors to this policy.

Based upon the analyses of the data for the three groups the following results are reported:

1. JAN techniques were successful in capturing 39 out of 59 usable policies (66 percent effective) present in the three groups of judges.

2. The PCG judges were found to be clustered in eight judgmental systems. Three of the policy systems were specific categorical policies while five were more complex.

3. The lawyer judges were found to be clustered in seven judgmental systems. Three of the policy systems were specific categorical policies while four were more complex.

4. The police judges were found to be clustered in seven judgmental systems. Three of the police systems were specific categorical policies while four were more complex.

5. The complex lawyer-judge policies were more concerned with impact characteristics of the material than were the policies expressed by the PCG judges and the police judges.

6. The complex PCG-judge policies were more concerned with whether the material had redeeming social importance than were the lawyers and the police.

7. Variable 8--Displays flagellation or torture--was an important content variable in the majority of complex policies regardless of the group source.

8. Variable 11--Displays group sex activity--was not an important content variable in any of the complex policies captured from the three groups.

9. Variable 4--Involves uncovered adult human genetalia, Variable 5--Displays human sexual intercourse, Variable 6--Displays masturbation, Variable 9--Displays homosexuality, and Variable 10--Involves interracial sexual relationships,were less frequently identified as important in the complex policies which were captured.

SUMMARY

The JAN technique was successfully utilized in capturing the policies (specific and complex) of 39 out of 59 judges selected from doctoral PCG students, lawyers and police officers from the city of Greeley. The technique was even more effective in explaining the different policies. The variety of policies present suggest that this is not a simple problem area with a simple solution readily available.

REFERENCES

Bottenberg, R.A., and Christal, R.E. An Iterative Technique for Clustering Criteria Which Retains Optimum Predictive Efficiency. Lackland Air Force Base, Texas: Personnel Laboratory, Wright Air Development Division, March 1961. (WADD-TN-61-30, ASTIA Document AS-261 615)

Christal, R.E. JAN: A Technique for Analyzing Individual and Group Judgment. Lackland Air Force Base, Aerospace Medical Division, February, 1963. (PRL-TDR-63-3, ASTIA Document AD-403 813)

_____,"JAN: A Technique for Analyzing Individual and Group Judgment." The Journal of Experimental Education, 1968, 36, 24-27

Dudycha, A.L. "A Monte Carlo Evaluation of JAN: A Technique for Capturing the Clustering Raters' Policies." Organizational Behavior and Human Performance, 1970, 5, 501-516.

Ward, J.H., Jr. "Multiple Linear Regression Models." In H. Borko (ed.), Computer Applications in the Behavior Sciences. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1962.

_____, and Hook, M.E. "A Hierarchical Grouping Procedure Applied to a Problem of Grouping Profiles." Unpublished Manuscript, Personnel Laboratory, Lackland Air Force Base, Texas, 1961. (ASD, AFSC, Report No. ASD-TN-61-55)

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 4, No. 3 1973

IDENTIFYING FACULTY POLICIES OF TEACHING EFFECTIVENESS

Samuel R. Houston
University of Northern Colorado

ABSTRACT

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

INTRODUCTION

A merit system has recently been introduced at the University of Northern Colorado in which faculty members are evaluated in terms of their teaching, advising of students, service to the institution, and professional activities.

The primary purpose of this study was to determine if Judgment Analysis (JAN) could identify what policy (or policies) were present in the faculty members of College of Education relative to the importance of the four areas or categories of evaluation. JAN, as a technique for both capturing and clustering raters' policies, was suggested by Bottenberg and Christal (1961) and Christal (1963). JAN utilizes a multiple regression model and a hierarchical grouping technique which clusters raters on the basis of the homogeneity of their raw score regression vectors (their prediction equations).

METHOD

Sample: The Ss consisted of 150 faculty members in the College of Education at the University of Northern Colorado. These faculty members were asked to serve as judges on a voluntary basis. Of the 150 faculty members who were contacted during the Summer Quarter, 1973, 57 completed the instrument and returned it before the spcified deadline.

Instrument: The instrument consisted of 60 hypothetical profiles of university faculty members. Each faculty member was described in terms of his classroom teaching, advising of students, service to the university, and professional activities. Each of the dimensions was scored as outstanding (O), above average (AA), satisfactory (S), or unsatisfactory (U). A copy of the instrument appears in Table 1.

The profile scores for the 60 hypothetical faculty members were determined by the use of a table of random numbers. This procedure guaranteed maximum variability for each of the profile variables and minimum correlation between the variables. High intercorrelations between variables (multicollinearity) makes the interpretation of policy and policy differences especially difficult. For that reason, the table of random numbers was used in the construction of the instrument.

TABLE 1

PROFILES OF SIXTY HYPOTHETICAL FACULTY MEMBERS

Profile Number	Classroom Teaching	Student Advising	Institutional Service	Professional Activities	Outstanding	Above Average	Satisfactory	Unsatisfactory	Profile Number	Classroom Teaching	Student Advising	Institutional Service	Professional Activities	Outstanding	Above Average	Satisfactory	Unsatisfactory
1.	S	U	U	S					31.	S	S	AA	S				
2.	S	S	U	S					32.	O	S	O	O				
3.	O	U	S	O					33.	O	O	O	AA				
4.	AA	S	S	AA					34.	S	S	S	AA				
5.	O	U	U	AA					35.	O	O	AA	S				
6.	U	AA	U	AA					36.	U	AA	U	AA				
7.	O	O	U	AA					37.	AA	U	AA	U				
8.	U	AA	S	S					38.	U	S	O	O				
9.	AA	S	S	AA					39.	O	S	U	AA				
10.	AA	S	O	O					40.	S	AA	U	U				
11.	AA	AA	AA	O					41.	U	AA	U	U				
12.	S	AA	S	O					42.	O	U	AA	AA				
13.	S	S	O	AA					43.	U	AA	O	O				
14.	AA	AA	S	O					44.	AA	U	U	U				
15.	S	AA	S	O					45.	O	O	S	O				
16.	O	AA	O	O					46.	AA	S	O	AA				
17.	U	AA	O	S					47.	AA	S	U	O				
18.	U	S	U	AA					48.	U	U	S	O				
19.	U	AA	O	S					49.	O	U	S	AA				
20.	O	AA	U	S					50.	AA	U	AA	O				
21.	O	U	U	S					51.	U	U	AA	S				
22.	AA	S	O	AA					52.	U	AA	U	U				
23.	AA	U	S	U					53.	U	AA	O	U				
24.	U	U	S	U					54.	S	AA	O	U				
25.	AA	O	S	S					55.	S	AA	U	AA				
26.	AA	S	S	AA					56.	AA	S	AA	O				
27.	U	S	U	U					57.	AA	AA	O	O				
28.	AA	O	S	S					58.	S	O	S	AA				
29.	O	AA	U	U					59.	AA	AA	S	S				
30.	U	U	AA	O					60.	S	U	S	O				

Procedure and Design: A total of 150 College of Education faculty members at UNC were contacted by campus mail. Fifty-seven volunteered to serve as judges and were requested to assign an overall rating to each of the 60 hypothetical faculty members by checking one of the four categories after each of the profiles.

The JAN procedure was applied to the results of the ranking process in order to determine the number of policies which were present. Judgment Analysis is not a procedure for either validating the dimensions used (variables) or methods of obtaining information on the dimensions (profile scores). It is assumed that the variables used have some validity and that there is an objective way available in which the variables can be quantified.

RESULTS

The means and standard deviations for the profiles are given in Table 2. A variable received a score of 4 if the dimension was scored as outstanding, 3 if the dimension was scored as above average, 2 if the dimension was scored as satisfactory, and 1 if the dimension was scored as unsatisfactory. The independent or predictor variables used include Variable 1--Teaching, Variable 2--Advising of Students, Variable 3--Service to University, and Variable 4--Professional Activities.

TABLE 2

MEANS AND STANDARD DEVIATIONS (N=60)

Variable	Mean	Standard Deviation
1 (Teaching	1.50	1.12
2 (Advising)	2.33	0.98
3 (Service)	2.32	1.13
4 (Professional)	2.72	1.08

Table 3 presents the intercorrelations between the independent variables. An inspection of Table 3 shows that the correlations are quite low. None of the correlation coefficients were significant (N=60) at the 0.01 level. Thus, the objective of generating variables with low intercorrelations was realized by using the table of random numbers.

TABLE 3

Correlation Matrix For Profile Scores

Variable	1	2	3	4
1	1.00	.02	-.02	.20
2	.02	1.00	.06	-.04
3	-.02	.06	1.00	.26
4	.20	-.04	.26	1.00

Of the 57 faculty members who participated in the study, six expressed a specific policy (Policy I) which can be interpreted as a policy which is opposed to the evaluation system. These six judges checked a single column (five checked the column headed "satisfactory" while one judge checked the column headed "unsatisfactory") which is probably the only way to participate in the study and still express opposition to the concept of evaluation for merit pay.

Two additional policies (Policy II and Policy III) were expressed by the remaining fifty-one judges. The existence of two policies was made by inspecting the various stages of the hierarchical grouping of judges according to the suggestion of Ward and Hook (1963) who recommend looking for a break in the objective function (sequential R^2 differences). The two systems R^2 (2 policies) value is .76 while the one system R^2 (1 policy) value is .73. This .03 drop is statistically significant (F-test) and represents a drop three times as great (.01) as the previous drop.

Table 4 presents the validity coefficients for each of the fifty-one judges who expressed a complex policy. To determine the specific policy for each judge one can look across the four validity coefficients for the judge in question. Each of these validities represents a correlation of the independent variable with the overall ranking of the sixty hypothetical faculty members. Judge 1, for example, has determined Variable 1 (Teaching) to be important (significant at .01 level) in his policy while Variables 2,3, and 4 are not significant contributors.

For a clearer representation of Policies II and III which were captured by the JAN process, refer to Table 5 and Table 6. Table 5 presents the important variables (significant at .01 level) for judges who were identified as expressing Policy II. The important variables for the judges expressing Policy III are presented in Table 6.

Policy II was expressed by 27 judges (Table 5). This policy places heavy emphasis on Variable 1 (Teaching) and Variable 4(Professional Activities) and no emphasis on Variable 3 (Service). Variable 2 is considered important by some of the judges in this policy. On the other hand, Policy III (Table 6), also emphasized the importance of Variable 1 (Teaching) and Variable 4(Professional Activities) while placing some emphasis on Variable 3 (Service) which was ignored in Policy II. Variable 2 receives some support in Policy III. Usually the judges in Policy III were attending to either Variable 2 or 3 but not both.

SUMMARY

The JAN technique was successfully utilized in capturing the policies (specific and complex) of fifty-seven faculty members in the College of Education at UNC who volunteered to serve as judges. Three policies were captured. Policy I which was expressed by six judges can be interpreted as a policy which is opposed to the evaluation system for merit pay for faculty members. Policy II (expressed by twenty-seven judges) placed heavy emphasis on Variable 1 (Teaching) and Variable 4 (Professional Activities) and no emphasis on Variable 3 (Service). Variable 2 (Advising of Students) was considered to be important by some of policy makers in Policy II. The final policy, Policy III, also emphasized Variables 1 and 4. However, the policy makers of Policy III were also attending to Variable 3 which was ignored in Policy II. Variable 2(Advising of Students) was considered important by some of the decision makers in Policy III. In general, Variables 1 and 4 were important variables in both policies, while Variables 2 and 3 received less support.

TABLE 4

TABLE OF VALIDITY COEFFICIENTS

Judge	1	2	3	4
1	.94*	.19	.09	.32
2	.82*	.29	.38*	.32
3	.90*	.21	.13	.31
4	.64*	.32	.41*	.53*
5	.71*	.43*	.36*	.40*
6	.85*	.27	.18	.39*
7	.94*	.16	.11	.32
8	.87*	.13	.15	.30
9	.68*	.23	.14	.43*
10	.83*	.31	.19	.37*
11	.48*	.42*	.42*	.37*
12	.85*	.35*	.21	.36*
13	.73*	.22	.34*	.58*
14	.78*	.22	.35*	.36*
15	.83*	.14	.24	.54*
16	.75*	.25	.27	.38*
17	.93*	.11	.19	.26
18	.81*	.35*	.13	.31
19	.54*	.65*	.15	.29
20	.87*	.24	.17	.38*
21	.61*	.39*	.29	.29
22	.79*	.32	.22	.33*
23	.58*	.18	.31	.77*
24	.88*	.22	.22	.32
25	.74*	.21	.38*	.51*
26	.62*	.17	.48*	.57*
27	.68*	.35*	.22	.40*
28	.78*	.32	.21	.43*
29	.65*	.27	.35*	.60*
30	.64*	.17	.29	.75*
31	.71*	.38*	.38*	.45*
32	.76*	.38*	.23	.34*
33	.82*	.35*	.19	.35*
34	.66*	.34*	.42*	.52*
35	.83*	.38*	.15	.29
36	.72*	.50*	.07	.16
37	.76*	.30	.25	.39*
38	.65*	.26	.43*	.51*
39	.84*	.27	.27	.38*
40	.42*	.76*	.15	.21
41	.75*	.22	.33*	.52*
42	.87*	.22	.23	.42*
43	.71*	.37*	.34*	.35*
44	.71*	.44*	.22	.35*
45	.81*	.22	.23	.35*
46	.82*	.32	.18	.36*
47	.84*	.33*	.09	.34*
48	.77*	.43*	.20	.35*
49	.60*	.44*	.42*	.54*
50	.88*	.25	.16	.36*
51	.99*	.06	.00	.20

*Significant at .01 level (N=60)

TABLE 5
IMPORTANT VARIABLES IN POLICY II

Judge Number	Variable 1	2	3	4
1	*			
3	*			
6	*			*
7	*			
8	*			
9	*			*
12	*	*		*
16	*			*
17	*			
18	*	*		
19	*	*		
20	*			*
22	*			*
24	*			
27	*	*		*
32	*	*		*
33	*	*		*
36	*	*		
37	*			*
39	*			*
40	*	*		*
42	*			*
45	*			*
46	*			*
49	*	*	*	*
50	*			*
51	*			

TABLE 6
IMPORTANT VARIABLES IN POLICY III

Judge Number	Variable 1	2	3	4
2	*		*	
4	*		*	*
5	*	*	*	*
10	*			*
11	*	*	*	*
13	*		*	*
14	*		*	*
15	*			*
21	*	*		
23	*			*
25	*		*	*
26	*		*	*
28	*			*
29	*		*	*
30	*			*
31	*	*	*	*
34	*	*	*	*
35	*	*		
38	*		*	*
41	*		*	*
43	*	*	*	*
44	*	*		*
47	*	*		*
48	*	*		*

REFERENCES

Bottenberg, R.A., and Christal, R.E. An Iterative Technique for Clustering Criteria Which Retains Optimum Predictive Efficiency. Lackland Air Force Base, Texas: Personnel Laboratory, Wright Air Development Division, March 1961. (WADD-TN-61-30, ASTIA Document AS-261 615)

Christal, R.E. JAN: A Technique for Analyzing Individual and Group Judgment. Lackland Air Force Base, Aerospace Medical Division, February, 1963. (PRL-TDR-63-3, ASTIA Document AD-403 813)

Ward, J.H. Jr., and Hook, M.E. "A Hierarchical Grouping Procedure Applied to a Problem of Grouping Profiles." Unpublished Manuscript, Personnel Laboratory, Lackland Air Force Base, Texas, 1961. (ASD, AFSC, Report No. ASD-TN-61-55)

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 4, No. 3 1973

MULTIPLE LINEAR REGRESSION MODELS FOR ANALYSIS OF COVARIANCE INCLUDING TEST FOR HOMOGENEITY OF REGRESSION SLOPE

PART 1 – ONEWAY DESIGNS WITH ONE COVARIATE.

By M.A. BREBNER

UNIVERSITY OF CALGARY
CANADA

The objective of the paper is to describe the multiple linear regression models which correspond exactly in all respects to the classical analysis of covariance and tests for homogeneity of regression slope. The models described by Bottenberg and Ward (1) are certainly sound in the context they are given, but none of the models given correspond to the standard test for homogeneity of regression slope, though Bottenberg, Ward and Jennings have given the regression equations for this case (4). In my opinion the approach pioneered by Bottenberg and Ward gives a much clearer and more meaningful understanding of analysis of covariance than the classical method, and I hope this paper will extend the insight of education research workers on this matter.

Let $\underline{y}$ be the criterion vector, $\underline{z}$ the covariate vector, and $\underline{g}_1, \underline{g}_2, \underline{g}_3, \ldots \underline{g}_s$ be the categorical vectors for mutually exclusive groups 1,2, ... s. It is assumed that group r contains n_r subjects and that the total number of subjects N is given by

$$N = n_1 + n_2 + \ldots + n_s .$$

The elements y_i and z_i of the vectors $\underline{y}$ and $\underline{z}$ are the criterion and covariate scores for subject i and the corresponding elements in the vectors $\underline{g}_r$ are unity if subject i is in group r, and otherwise zero. As in Bottenberg and Ward (1) the product $\underline{g}_r\underline{z}$ is a vector $\underline{z}_r$ containing the covariate scores for group r in the appropriate positions and all other elements are zero. For example if

$$\underline{g}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{ and } \underline{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{bmatrix}, \text{ then } \underline{g}_2\underline{z} = \begin{bmatrix} 0 \\ z_2 \\ z_3 \\ 0 \\ 0 \end{bmatrix} = \underline{z}_r$$

The models to be used are now stated as follows:

Model 1 Full Model

$$\underline{y} = a_1\underline{g}_1 + a_2\underline{g}_2 + \ldots a_s\underline{g}_s + b_1\underline{g}_1\underline{z} + b_2\underline{g}_2\underline{z} + \ldots b_s\underline{g}_s\underline{z} + \underline{E}^{(1)}$$

Number of linearly independent vectors = 2s

Model 2 Restriction of Model 1 with $b_1 = b_2 = \ldots = b_s = b$

$$\underline{y} = a_1\underline{g}_1 + a_2\underline{g}_2 + \ldots + a_s\underline{g}_s + b\underline{z} + \underline{E}^{(2)}$$

Number of linearly independent vectors = s+1

Model 3 Restriction of model 1 with $a_1 = a_2 = \ldots = a_s = a$

$$\underline{y} = a\underline{u} + b_1\underline{g}_1\underline{z} + b_2\underline{g}_2\underline{z} + \ldots + b_s\underline{g}_s\underline{z} + \underline{E}^{(3)}$$

where $\underline{u} = \underline{g}_1 + \underline{g}_2 + \ldots + \underline{g}_s$ is the unit vector.

Number of linearly independent vectors = s+1

Model 4 Restriction of model 2 with $a_1 = a_2 = \ldots = a_s = a$
(Alternatively, though less meaningful, the restriction of model 1 with $a_i = a$ for i=1,2, ... ,s and $b_i = b$ for i=1,2, ... ,s)

$$\underline{y} = a\underline{u} + b\underline{z} + \underline{E}^{(4)}$$

Number of linearly independent vectors = 2

It should be pointed out that the least squares estimates of a_i and b_i will, in general, be different for each model.

For the classical test of homogeneity of regression slope the error sums of squares $\underline{E}^{(1)T}\ \underline{E}^{(1)}$ and $\underline{E}^{(2)T}\ \underline{E}^{(2)}$ for models 1 and 2 are compared. This corresponds to the within groups sum of squares for slopes, and the sum of squares assuming the regression slopes are equal for all groups. See for example figures 1(a) and 1(b).

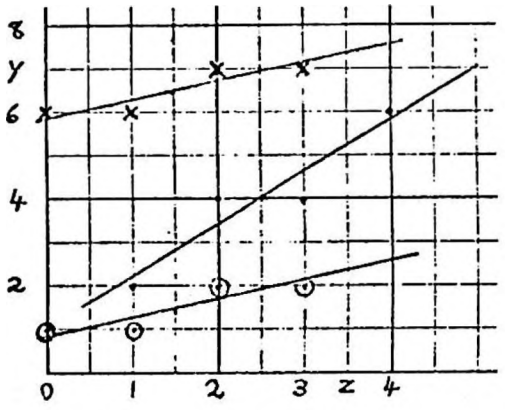


Fig. 1(a)

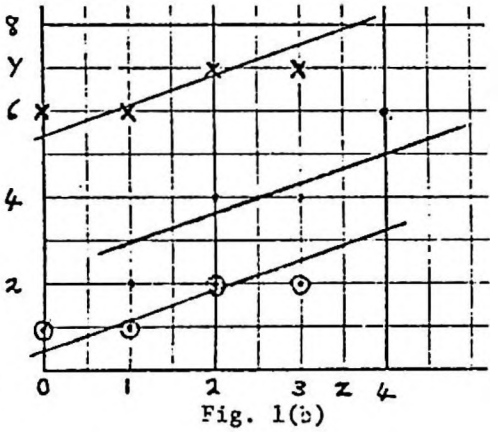


Fig. 1(b)

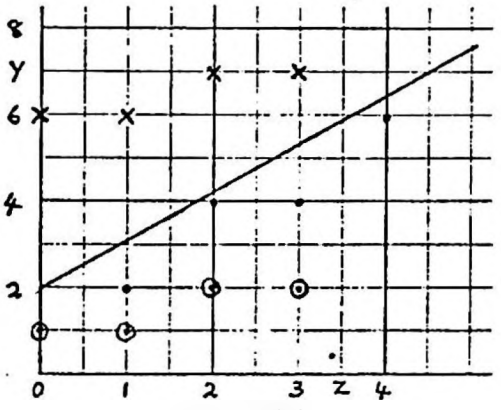


Fig. 1(c)

It will be proved later in the paper that the comparison of models 1 and 2 does yield exactly the same F ratio and degrees of freedom as the classical analysis for testing homogeneity of slope.

The F ratio to test the hypothesis is given by

$$F = \frac{(ESS_2 - ESS_1)/(2s - (s + 1))}{ESS_1/(N - 2s)}$$

and has $(s - 1)$ and $(N - 2s)$ degrees of freedom.

A similar comparison using models 2 and 4 yields the classical analysis of covariance result, the F statistics being computed as

$$F = \frac{(ESS_4 - ESS_2)/((s + 1)-2)}{ESS_2/(N -(s + 1))}$$

and having $(s - 1)$ and $(N - s - 1)$ degrees of freedom.

The following notation will be adopted for group sums and overall sums

$\sum^{(r)} z$ = sum of the covariate scores for group r = $\underline{g}_r^T \underline{z}$,

$\sum^{(r)} y$ = sum of the criterion scores for group r = $\underline{g}_r^T \underline{y}$,

$\sum^{(r)} zy$ = sum of the product of covariate and criterion scores for group r = $\underline{z}_r^T \underline{y}$,

$$\sum z = \sum_{i=1}^{N} z_i = \underline{u}^T \underline{z} ,$$

$$\sum y = \sum_{i=1}^{N} y_i = \underline{u}^T \underline{y} ,$$

$$\sum xy = \sum_{i=1}^{N} z_i y_i = \underline{z}^T \underline{y} \text{ and}$$

$$\sum z^2 = \sum_{i=1}^{N} z_i^2 = \underline{z}^T \underline{z} .$$

Let the N × s matrix

$$G = [\underline{g}_1 \ \underline{g}_2 \ \cdots \ \underline{g}_s]$$

and the N × s matrix

$$Z = [g_1 \underline{z} \ g_2 \underline{z} \ \cdots \ g_s \underline{z}] ,$$

then the regression models can be written in matrix notation as follows

Model 1

$$\underline{y} = [G\ Z]\begin{bmatrix}\underline{a}\\ \underline{b}\end{bmatrix} + \underline{E}^{(1)}$$

Model 2

$$\underline{y} = G\underline{a} + b\underline{z} + \underline{E}^{(2)}$$

Model 3

$$\underline{y} = a\underline{u} + Z\underline{b} + \underline{E}^{(3)}$$

Model 4

$$\underline{y} = a\underline{u} + b\underline{z} + \underline{E}^{(4)}$$

The least squares extimates for model 1 are obtained by solving for the elements of $\underline{a}$ and $\underline{b}$ the linear equations

$$\begin{bmatrix}G^T\\ Z^T\end{bmatrix}[G\ Z]\begin{bmatrix}\underline{a}\\ \underline{b}\end{bmatrix} = \begin{bmatrix}G^T\\ Z^T\end{bmatrix}\underline{y}$$

or $$\begin{bmatrix}G^TG & G^TZ\\ Z^TG & Z^TZ\end{bmatrix}\begin{bmatrix}\underline{a}\\ \underline{b}\end{bmatrix} = \begin{bmatrix}G^T\underline{y}\\ Z^T\underline{y}\end{bmatrix}$$

or $$\begin{bmatrix}D^{(1)} & D^{(2)}\\ D^{(2)} & D^{(3)}\end{bmatrix}\begin{bmatrix}\underline{a}\\ \underline{b}\end{bmatrix} = \underline{c}\ ,$$

where $D^{(1)}$, $D^{(2)}$ and $D^{(3)}$ are diagonal matrices.

The diagonal elements of $D^{(1)}$, $D^{(2)}$ and $D^{(3)}$ are given by

$$d_{rr}^{(1)} = n_r\ ,$$

$$d_{rr}^{(2)} = \sum^{(r)} z\ ,$$

and $$d_{rr}^{(3)} = \sum^{(r)} z^2\ .$$

The equations have the form

$$\left[\begin{array}{ccc:ccc} n_1 & & 0 & \sum^{(1)}z & & 0 \\ & n_2 & & & & \\ 0 & & n_s & 0 & & \sum^{(s)}z \\ \hdashline \sum^{(1)}z & & 0 & \sum^{(1)}z^2 & & 0 \\ & & & & & \\ 0 & & \sum^{(s)}z & 0 & & \sum^{(s)}z^2 \end{array}\right] \left[\begin{array}{c} \underline{a} \\ \hdashline \\ \underline{b} \end{array}\right] = \left[\begin{array}{c} \sum^{(1)}y \\ \vdots \\ \sum^{(s)}y \\ \hdashline \sum^{(1)}zy \\ \vdots \\ \sum^{(s)}zy \end{array}\right]$$

It is seen that the 2s linear equations are s pairs of uncoupled equations of the form

$$n_r a_r + (\sum^{(r)}z)b_r = \sum^{(r)}y \quad \text{---------- } (r)$$

and

$$(\sum^{(r)}z)a_r + (\sum^{(r)}z^2)b_r = \sum^{(r)}zy \quad \text{-------- } (s+r)$$

As noted these are the r^{th} and $(s + r)^{th}$ equations of the set, $r = 1,2,\ldots,s$.

Using appendix II

$$b_r = \frac{S^{(r)}_{zy}}{SS^{(r)}_z}$$

and

$$a_r = \bar{y}^{(r)} - \bar{z}^{(r)} b_r$$

Hence the regression equation for model 1 is

$$\underline{y} = \sum_{r=1}^{s} \left\{ (\bar{y}^{(r)} - \bar{z}^{(r)} b_r) \underline{g}_r + b_r \underline{z}_r \right\}$$

or

$$\underline{y} = \sum_{r=1}^{s} \left\{ \bar{y}^{(r)} + b_r(\underline{z} - \bar{z}^{(r)}) \right\} \underline{g}_r$$

This separates into the r independent regression equations

$$\underline{y}\underline{g}_r = \left\{\bar{y}^{(r)} + b_r(\underline{z} - \bar{z}^{(r)})\right\}\underline{g}_r \qquad r = 1, 2, \ldots, s$$

These s equations correspond to the equations (12.9) in Meyers (2), and the regression coefficients b_r correspond to $b_j^{(2)}$ given by equation (12.9').

The least squares equations for model 2 are

$$\begin{bmatrix} \underline{g}_1^T \\ \underline{g}_2^T \\ \\ \underline{g}_s^T \\ \underline{z}^T \end{bmatrix} \begin{bmatrix} \underline{g}_1 \, \underline{g}_2 \cdots \underline{g}_s \, \underline{z} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \\ a_s \\ b \end{bmatrix} = \begin{bmatrix} \underline{g}_1^T \\ \underline{g}_2^T \\ \\ \underline{g}_s^T \\ \underline{z} \end{bmatrix} \underline{y} \quad ,$$

or

$$\begin{vmatrix} n_1 & 0 & \cdot & \cdot & 0 & \sum^{(1)} z \\ 0 & n_2 & \cdot & \cdot & \cdot & \sum^{(2)} z \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & 0 & \\ 0 & \cdot & \cdot & 0 & n_s & \sum^{(s)} z \\ \sum^{(1)} z & \cdot & \cdot & & \sum^{(s)} z & \sum z^2 \end{vmatrix} \begin{vmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_s \\ b \end{vmatrix} = \begin{vmatrix} \sum^{(1)} y \\ \sum^{(2)} y \\ \cdot \\ \cdot \\ \cdot \\ \sum^{(s)} y \\ \sum zy \end{vmatrix}$$

The first s equations have the form

$$n_r a_r + (\sum^{(r)} z)b = \sum^{(r)} y$$

or

$$a_r = (\sum^{(r)} y)/n_r - ((\sum^{(r)} z)/n_r)b$$

$$= \bar{y}^{(r)} - \bar{z}^{(r)} b \qquad r = 1, 2, \ldots, s.$$

where $\bar{y}^{(r)}$ is the mean of the y scores for group r, etc.

The final equation is

$$(\sum^{(1)} z)a_1 + (\sum^{(2)} z)a_2 + \ldots + (\sum^{(s)} z)a_s + (\sum z^2)b = \sum zy$$

and substituting for each a_r gives the equation

$$\sum_{r=1}^{s} (n_r \bar{z}^{(r)})(\bar{y}^{(r)} - \bar{z}^{(r)} b) + (\sum z^2)b = \sum zy$$

which can be solved for b.

Hence

$$-\left(\sum_{r=1}^{s} n_r \bar{z}^{(r)} \bar{z}^{(r)}\right)b + (\sum z^2)b = \sum zy - \sum_{r=1}^{s} n_r \bar{z}^{(r)} \bar{y}^{(r)} ,$$

or
$$b = \frac{\sum_{r=1}^{s} \left\{ \left(\sum^{(r)} zy\right) - n_r \bar{z}^{(r)} \bar{y}^{(r)} \right\}}{\sum_{r=1}^{s} \left\{ \left(\sum^{(r)} z^2\right) - n_r \bar{z}^{(r)} \bar{z}^{(r)} \right\}} = \frac{\sum_{r=1}^{s} \left\{ SS_{zy}^{(r)} \right\}}{\sum_{r=1}^{s} \left\{ SS_z^{(r)} \right\}} = b_w \text{ say,}$$

following the notation of Appendix I.

Note the regression coefficient in this case is a pooled within groups estimate of the regression slope. This result gives the group r intercept a_r as

$$a_r = \bar{y}^{(r)} - \bar{z}^{(r)} b_w \qquad \text{where} \quad r = 1,2, \ldots,s.$$

Hence the regression equation for model 2 is

$$\underline{y} = \sum_{r=1}^{s} (\bar{y}^{(r)} - \bar{z}^{(r)} b_w)\underline{g}_r + b_w \underline{z}$$

or

$$\underline{y} = \sum_{r=1}^{s} \left\{ \bar{y}^{(r)} \underline{g}_r + b_w(\underline{z}_r - \bar{z}^{(r)} \underline{g}_r) \right\}$$

This is exactly the regression equation given in Meyer(2) as equation (12.10) and b_w corresponds to $b^{(3)}$ given in equation (12.10').

Hence we have shown that the regression equations for model 1 and 2 do correspond exactly to the classical regression equations with (1) different slopes for each group, and (2) a common slope for all groups respectively. The regression equation for model 4 clearly corresponds to the classical case when data from all groups is pooled to give a single regression equation (see fig. 1c). As the regression equations for models 1, 2 and 4 are the same as the classical case, the sums of squares ESS1 and ESS2 correspond exactly to the appropriate within groups sums of squares, and ESS4 will be the total sum of squares. By examining the second to fourth pages, we see the degrees of freedom computed by the regression approach also correspond exactly to the degrees of freedom for the classical analysis of covariance (see for example Meyer (2) pages 309 to 310 or Ferguson (3) pages 332 and 338). The comparison of models 2 and 4 corresponds to the second test given on page 68 of Bottenberg and Ward.

REFERENCES

1. Bottenberg, R. & Ward, J. H. Applied multiple linear regression. Technical Documentary Report PRL-TDR-63-6, 6570th Personnel Research Laboratory, Lackland Air Force Base, Texas, 1963.

2. Meyers, J. L. Fundamentals of Experimental Design. Allan and Bacon, Boston, 1966.

3. Ferguson, G. A. Statistical Analysis in Psychology and Education. McGraw Hill, 1966

4. Ward, J. H., Bottenberg, R. A., Jennings, E. and Koplyay, J. 1970 AERA Pression on Applied Linear Regression in Educational Research (unpublished)

APPENDIX I (used in Appendix II)

Let $S_{xy} = \sum_{i=1}^{m} (x_i - \bar{x})(y_i - \bar{y})$, where $\bar{x} = (\sum_{i=1}^{m} x_i)/m$ and $\bar{y} = (\sum_{i=1}^{m} y_i)/m$,

then $m\,S_{xy} = m(\sum_{i=1}^{m} x_i y_i) - (\sum_{i=1}^{m} x_i)(\sum_{i=1}^{m} y_i)$,

Note (a) $m\,SS_x = m\,S_{xx} = m(\sum_{i=1}^{m} x_i^2) - (\sum_{i=1}^{m} x_i)^2$,

and (b) $S_{xy} = (\sum_{i=1}^{m} x_i y_i) - \frac{1}{m}(\sum_{i=1}^{m} x_i)(\sum_{i=1}^{m} y_i)$.

Proof

$$S_{xy} = \sum_{i=1}^{m} (x_i - \bar{x})(y_i - \bar{y}) = (\sum_{i=1}^{m} x_i y_i) - \bar{x}(\sum_{i=1}^{m} y_i) - \bar{y}(\sum_{i=1}^{m} x_i) + (\sum_{i=1}^{m} \bar{x}\bar{y})$$

$$= (\sum_{i=1}^{m} x_i y_i) - \bar{x}(m\bar{y}) - \bar{y}(m\bar{x}) + m(\bar{x}\bar{y})$$

$$= (\sum_{i=1}^{m} x_i y_i) - m\bar{x}\bar{y} = (\sum_{i=1}^{m} x_i y_i) - \frac{1}{m}(\sum_{i=1}^{m} x_i)(\sum_{i=1}^{m} y_i),$$

using the definitions of $\bar{x}$ and $\bar{y}$.

Multiply by m to give the required equation

$$m\,S_{xy} = m(\sum_{i=1}^{m} x_i y_i) - (\sum_{i=1}^{m} x_i)(\sum_{i=1}^{m} y_i).$$

APPENDIX II

Consider the pair of equations

$$Aa + Bb = D \quad \text{——} \quad \text{(i)}$$

$$Ba + Cb = E \quad \text{——} \quad \text{(ii)}$$

where $A = m$, $B = \sum_{i=1}^{m} z_i$, $C = \sum_{i=1}^{m} z_i^2$, $D = \sum_{i=1}^{m} y_i$ and $E = \sum_{i=1}^{m} z_i y_i$.

To find b combine the two equations as follows

equation (ii) $-\ \frac{B}{A}$ x equation (i)

$$Cb - \frac{B^2}{A} b = E - \frac{BD}{A}$$

$$\text{or } b = \frac{AE - BD}{AC - B^2} = \frac{m\left(\sum_{i=1}^{m} z_i y_i\right) - \left(\sum_{i=1}^{m} z_i\right)\left(\sum_{i=1}^{m} y_i\right)}{m\left(\sum_{i=1}^{m} z_i^2\right) - \left(\sum_{i=1}^{m} z_i\right)^2}$$

$$= \frac{m\, S_{zy}}{m\, SS_z} = \frac{S_{zy}}{SS_z} \qquad \text{using Appendix I}$$

Similarly

$$a = \frac{CD - BE}{AC - B^2} = \frac{\left(\sum_{i=1}^{m} z_i^2\right)\left(\sum_{i=1}^{m} y_i\right) - \left(\sum_{i=1}^{m} z_i\right)\left(\sum_{i=1}^{m} z_i y_i\right)}{m\, SS_z}$$

$$= \frac{\left(\sum_{i=1}^{m} z_i^2\right)\left(\sum_{i=1}^{m} y_i\right) - \frac{1}{m}\left(\sum_{i=1}^{m} z_i\right)^2\left(\sum_{i=1}^{m} y_i\right) + \frac{1}{m}\left(\sum_{i=1}^{m} z_i\right)^2\left(\sum_{i=1}^{m} y_i\right) - \left(\sum_{i=1}^{m} z_i\right)\left(\sum_{i=1}^{m} z_i y_i\right)}{m\, SS_z}$$

$$= \frac{\left(\sum_{i=1}^{m} y_i\right)\left\{\left(\sum_{i=1}^{m} z_i^2\right) - \frac{1}{m}\left(\sum_{i=1}^{m} z_i\right)^2\right\} + \left(\sum_{i=1}^{m} z_i\right)\left\{\frac{1}{m}\left(\sum_{i=1}^{m} z_i\right)\left(\sum_{i=1}^{m} y_i\right) - \left(\sum_{i=1}^{m} z_i y_i\right)\right\}}{m\, SS_z}$$

$$= \frac{m\bar{y}\{SS_z\} + m\bar{z}\{-S_{zy}\}}{m\ SS_z} = \bar{y} - \bar{z}\left(\frac{S_{zy}}{SS_z}\right) = \bar{y} - \bar{z}b$$

using appendix I and formula for b derived above.

Hence the predicted value ϑ_i is given by

$$\vartheta_i = (\bar{y} - \bar{z}b) + bz_i$$

or $$\vartheta_i - \bar{y} = b(z_i - \bar{z}).$$

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 4, No. 3 1973

Regression Effects when the Assumption of Rectilinearity is not Tenable

John T. Pohlmann
Southern Illinois University, Carbondale

and

Isadore Newman
University of Akron, Akron Ohio

ABSTRACT

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

Introductory texts in measurement and statistics typically present the topic of regression by providing examples of bivariate distributions in which extreme scorers on the independent variable (X) tend to score closer to the mean on the dependent variable (Y). In the limiting case, where there is no systematic relationship between X and Y, the predicted Y ($\hat{Y}$) values, for all values of X, equal the mean of Y. On the other hand, when the correlation between X and Y is perfect, the regression effect is, by definition, absent. This implies that knowledge of X allows perfect prediction of Y, and when both X and Y have been transformed to Z scores, $\hat{Z}_y = Z_x$ for all values of Z_x. Predicted Z_y scores, then, are assumed to take on absolute values that are either equal to or less than the Z_x value from which they were derived. This notion is presented graphically in Figure 1.

In this discussion, raw scores have been transformed to standard

scores since differences in means and/or variances between X and Y can camouflage the regression phenomenon.

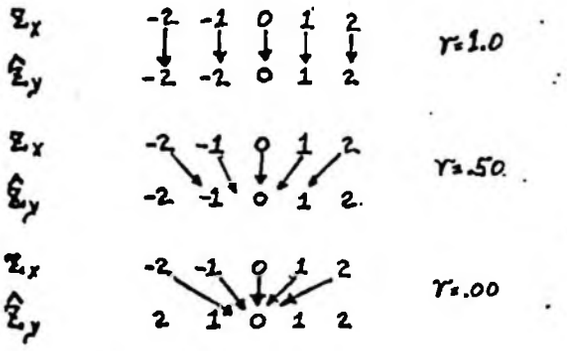


Figure 1

The projection of z_x to $\hat{z}_y$ under various degrees of association.

This conception of the regression phenomenon has implications in many areas where correlation is employed in analyzing and inferring from data. In covariance analyses, control is exerted on a covariate while comparing the performance of treatment groups against a criterion. The scores and group means are adjusted, in accordance with regression notions. When extreme groups in pretest - post-test designs are studied, the use of a control group is recommended so that change scores are not falsely attributed to intervening treatments, when the regression phenomenon is the real source of change. In repeated testing paradigms where change scores are the object of interest, expectations about changes by extreme scorers are often based on the regression phenomenon. A study reported by P. E. Vernon (7) provides an example of expectations based on the regression phenomenon failing to hold in a repeated testing paradigm. In that study changes in IQ were being studied as a function of repeated testing. It was expected that, upon repeated testing, high

scorers would score lower and low scorers would improve. Such expectations have been suggested by Campbell & Stanley (2), Kerlinger (5), Borg & Gall (1) and others. Vernon found, however, that high scorers scored even higher upon re-testing, gaining even more than low scorers.

When r_{xy} is not one, there is error about the least squares regression line relating X to Y. This error has two possible sources, (1) pure error, and (2) lack of fit error. Pure error is error that exists when the expected Z_y, given Z_x equals $\hat{Z}_y$ for all values of X in the population. Lack of fit error is characterized by expected Z_y values that are different from $\hat{Z}_y$ values over certain ranges of X. It should be stressed here that these departures must be significant at some specified alpha level before the presence of lack of fit error can be entertained when inspecting sample data.

The fitted regression line for r_{xy} has been calculated using a model that assumes the relationship between X and Y is linear. If the assumption of linearity is accurate, then error about the regression line is pure error. If, on the other hand, the assumption of linearity is not correct, error about the regression line is composed of both pure error and lack of fit error.

If a straight line is the best fitting regression line, and if r_{xy} is not unity, then regression to the mean of Y must occur. On the other hand, if the relationship between X and Y can best be represented by a curved line of any sort, the standard phenomenon of regression toward the mean will not result. For example, in Figure 2, data points have been drawn such that the relationship between Z_x and Z_y is clearly non-linear. Regression line I has been drawn through the

scatter plot using the least squares criterion, and a certain amount of error has been realized. The model that reflects line I in Figure 2 is as follows:

MODEL 1: $Y = a_o U + a_1 X_1 + E_1$

where:

Y is the dependent variable
U is a unit vector when multiplied by a_o yields the regression constant
a_1 is the regression weight

E_1 is the error vector

X_1 is the dependent variable

Admittedly, the magnitude of error in this model is small, but note what has occurred in our data. The subjects who scored 8 on X have an average score of 64 on Y. The Z_y value of 64 is greater than the Z_x value of 8. In other words, these subjects have "regressed" away from the mean of Y instead of regressing toward the mean of Y, as would be expected according to the regression phenomenon. The reason for this, of course, is that the relationship between X and Y is best represented by a regression model which allows for a quadratic relationship between X and Y. Model 2 is just such a model.

MODEL 2: $Y = a_o U + a_1 X_1 + a_2 X_1^2 + E_2$

where: Y is the criterion score

U is the unit vector

X_1 is the X value

X_1^2 is the X value squared

a_1, a_2, and a_o are the regression weights assigned to U, X, and X^2 respectively, and

E_2 is the vector of residuals

Since U and X will receive zero weights, a more parsimonious model would be Model 3.

MODEL 3: $Y = a_2X_1^2 + E$

where: Y is the criterion

a_2 is the weight for X_1^2

X_1^2 is the X score squared, and

E is the vector of residuals

Model 3, then, is the best model for representing the relationship between X and Y. Regression line II in Figure 2 reflects the use of Model 3 in fitting the data points.

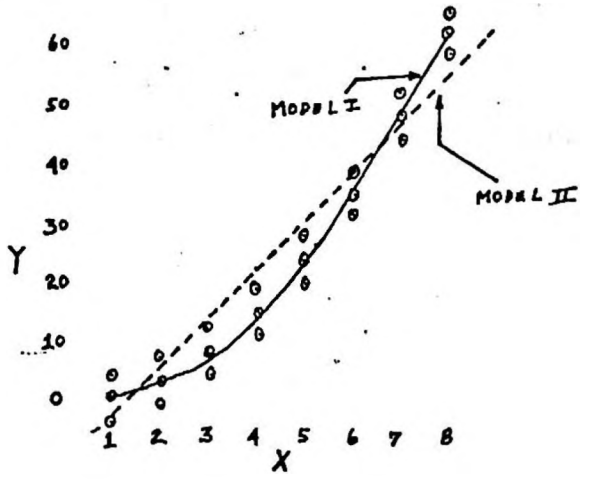


FIGURE 2

A COMPARISON OF LINEAR (MODEL 1) AND QUADRATIC (MODEL 2) REGRESSION LINES FITTED TO DATA REFLECTING A CURVED RELATIONSHIP BETWEEN X AND Y

Since the relationship between X and Y that is depicted in Figure 2 is curved, the assumption of rectilinearity is not correct. Consequently, expectations based upon the regression phenomenon would be in error and might mislead the researcher in interpreting his results. This example is particularly noteworthy, since the linear model reflected in MODEL 1 did an excellent job of accounting for variance in Y, but MODEL 3 did a much better job.

The fact that the scores did not distribute themselves in the bivariate distribution as the regression phenomenon predicted should have served as a cue for the researcher to attempt a non-linear fit to the data. By determining the kind of curvalinear relationship that existed, inferences might be made as to the underlying causes for that relationship.

Methods for Detecting Curvature

Various procedures exist for screening paired observations for possible non-linear trends. Oftentimes inspection of the scatter plot is sufficient to note the presence of curvature in the relationship. Inspection methods, though, lack certainty when there are few data points or if the correlation is rather small. The following methods are offered as suggestions to assist the researcher in determining if a non-linear trend is present in his data.

METHOD I: If there are only a few observations, fewer than 25 or so, the observations on X can be divided at each quintile and the mean Y values can be calculated for each group on X. The plot of these means could then be inspected for curvature. The researcher should recognize that for a set of data this small only distinctly non-linear trends will approach significance. If a researcher suspects, as a result of this inspection, that a non-linear trend is present he may want to test the significance of that trend.

METHOD II: A researcher could also calculate the correlation ratio or Eta^2 for his set of data.[1] Eta^2 could then be compared to r^2 to test if a significant non-linear trend exists in the data. An F-ratio could be generated as follows:

$$F = \frac{(Eta^2_{xy} - r^2_{xy}) / df_1}{(1.0 - Eta^2) / df_2}$$

where: Eta^2 is the Eta^2 for our data

r^2_{xy} is the Pearson r for the same data

df_1 is the number of linearly independent variables or scores on X used to calculate Eta^2, minus 2 and,

df_2 is the number of paired observations minus the number of linearly independent variables or scores used to calculate Eta^2.

If this F-ratio is found to be significant, a researcher may entertain the assumption that a non-linear trend is present in his data, but due to the nature of Eta^2, he cannot, from these analyses, describe the type of relationship that exists. Eta^2 could be significantly greater than r^2 because of any departure from linearity.

METHOD III: The most direct method of ascertaining not only if the relationship between X and Y is significantly non-linear, but also describing the nature of the relationship is a curve fitting approach employing multiple regression methods. Various functional relationships can be tested for and the partial regression weights may be used to describe the nature of the relationship. The use of a multiple regression approach which allows the researcher to try out various data transformations and the possible linear combinations among them is particularly useful.[2]

---

[1]Edwards (3) presents the method for calculating Eta^2 pp 136-139

CONCLUSION

One of the underlying assumptions necessary for the regression phenomenon to be accurate is that there is a linear relationship between X and Y. If the relationship is linear, all error will result in extreme scorers on X scoring closer to the mean of Y. But if the relationship between X and Y is non-linear, the regression phenomenon will not represent the true state of affairs in our data. Indeed, if researchers note that expectations based on the regression phenomenon are not supported by their data, it would behove them to seek out the nature of the curved relationship that exists. In fact, the nature of the curvature may serve as a clue for the researcher in determining why the regression phenomenon was insufficient for explaining his data.

---

[2]See Kelly, Beggs, McNeil, Eichelberger, and Lyon (1969) for a detailed discussion of a multiple regression approach such as this, specifically chapters 6, 7, & 8.

REFERENCES

1. Borg, W.R.; Gall, M.D., Educational Research: An Introduction, David McKay Co., Inc., New York, 1971, 533 pp.

2. Campbell, D.T.; Stanley, J.C., Experimental & Quasi-Experimental Design for Research, Rand McNally & Co., Chicago, 1969, 84 pp.

3. Edwards, A.L., Statistical Methods, Holt, Rinehart & Winston, Inc., New York, 1967, 462 pp.

4. Kelly, F.J.; Beggs, D.L.; McNeil, K.A.; Eichelberger, T; Lyon, J., Research Design in Behavioral Sciences: Multiple Regression Approach, Southern Illinois University Press, Carbondale, Illinois, 1969, 353 pp.

5. Kerlinger, F.N., Foundations of Behavioral Research, Holt, Rinehart & Winston, Inc., New York, 1965, 739 pp.

6. McNeil, K.; Spaner, S.D., "Brief Report: Highly Correlated Predictor Variables in Multiple Regression Models," Multivariate Behavioral Research, Vol. 6, pp. 117-125, 1971.

7. Vernon, P.E., "Practive & Coaching Effects in Intelligence Testing," The Educational Forum, pp. 273-274, 1954.

MULTIPLE LINEAR REGRESSION VIEWPOINTS
Vol. 4, No. 3 1973

A GENERAL LEAST SQUARES APPROACH TO THE ANALYSIS OF REPEATED MEASURES DESIGNS

George H. Olson

Dallas Independent School District

Abstract

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

The analysis of designs in which repeated observations are taken on each subject (split-plot designs, partially hierarchical designs, repeated-measures designs, etc.) sometimes cause problems for the researcher who desires a general least squares solution. A least squares analysis of such designs is particularly important when the researcher has either designed his experiment such that the levels of his "between-groups" factor contain sample sizes proportional to various strata in his population or has lost subjects for reasons related to his experimental treatment. In either of these cases a least squares solution, rather than an unweighted means solution, should be used.[1] Although most intermediate-level textbooks on analytical procedures for experimental designs present the necessary equations for obtaining least squares solutions to repeated-measures designs (e.g. Kirk, 1968, p. 280; Winer, 1962, p. 600), they do not discuss the practicalities of obtaining such a solution via a general linear hypothesis or multiple regression computer program. On the other hand, intermediate-level textbooks which treat regression analysis and the general linear model (e.g. Draper & Smith, 1966; Li, 1964) have not considered repeated-measures designs.[2]

The purpose of the present paper is to illustrate two approaches a researcher may take to obtain a general least squares solution to repeated-measures designs. These approaches assume that the researcher has at his disposal a general linear hypothesis, or regression analysis, computer program.

The first approach, appropriate when the total number of subjects is small, is a straight-forward extension of the typical regression approach to the analysis of variance. The second approach, appropriate when the total number of subjects is large, requires some minor computation in addition to that usually provided by most regression programs.

Approach 1: Small N

For the discussion in this section, let Table 1 represent a general 2 × 4 treatment by (repeated) measures design. In addition, let the notation, Y_{ijk}, represent the j'th measure (B_j) being taken on the k'th subject (S_k) in the i'th treatment group (A_i). For the example illustrated in Table 1, there are J (=4) measures with K_i (=3, 2) subjects in each of the I (=2) treatment groups. When the total sample size, N ($=\Sigma K_i$), is relatively small (a judgement left to the researcher), a least squares solution is easily obtained in a straightforward manner.

- - - - - - - - - - - - - - - - - -

TABLE 1

somewhere in here

- - - - - - - - - - - - - - - - - -

The analysis proceeds by treating the data as having been obtained from a nested design, with one observation per cell, in which "subjects" constitutes the nested factor. A so-called structural model for this design may be written as

$$Y_{ijk} = \mu + \alpha_i + \gamma_{k(i)} \beta_j + \alpha\beta_{ij} + \beta\gamma_{jk(i)} + \varepsilon_{ijk} \qquad (1)$$

where

(i) μ represents the grand mean effect,

(ii) α_i represents the treatment effect,

(iii) $\gamma_{k(i)}$ represents the subject-nested-within-groups effect (i.e. the "between-groups" error term),

(iv) β_j represents the measures effect,

(v) $\alpha\beta_{ij}$ represents the treatment by measures interaction,

(vi) $\beta\gamma_{jk(i)}$ represents the subjects by measures-nested-within-groups effect (i.e. the "within groups" error term), and

(vii) ε_{ijk} represents unaccounted-for sources of within-cell variation.

A general regression analysis using this model would require a minimum of NJ analysis of variance (ANOVA) design variables for a complete analysis, including tests of interactions. Since the design for which this model is appropriate is assumed to have only one observation per cell, there can be no within-cell variation. Therefore, the sums of squares due to ε_{ijk} (i.e., $SS_{\varepsilon_{ijk}}$) is zero. For this reason, the alternative model,

$$Y = \mu + \alpha_i + \gamma_{k(i)} + \beta_j + \alpha\beta_{ij} + \varepsilon^{(1)}_{ijk} \qquad (2)$$

where $\varepsilon^{(1)}_{ijk} = \beta\gamma_{jk(i)} + \varepsilon_{ijk}$, may be used advantageously since it requires only $I(J - 1) + N$ ANOVA design variables for a complete analysis. Writing (2) as a general linear model for the data in Table 1 yields

$$Y = uX_1 + a_1X_2 + c_{1(1)}X_3 + c_{2(1)}X_4 + c_{1(2)}X_5 + b_1X_6 + b_2X_7 + b_3X_8 + ab_1X_9 + ab_2X_{10} + ab_3X_{11} + e^{(1)} \qquad (3)$$

where the terms in (3) are more completely described in Figure 1. The model

- - - - - - - - - - - - - - - - - -

Figure 1

somewhere in here

- - - - - - - - - - - - - - - - - -

given in (2 or 3) may then be used as the "full model" in completing the regression analysis of the design given in Table 1. If the structural analogue

for each of the appropriate "reduced models" is written as follows:

$$Y_{ijk} = \mu + \gamma_{k(i)} + \beta_j + \alpha\beta_{ij} + \varepsilon^{(2)}_{ijk} \text{ (where } \varepsilon^{(2)}_{ijk} = \alpha_i + \beta\gamma_{jk(i)});$$

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon^{(3)}_{ijk} \text{ (where } \varepsilon^{(3)}_{ijk} = \gamma_{k(i)} + \beta\gamma_{jk(i)});$$

$$Y_{ijk} = \mu + \alpha_i + \gamma_{k(i)} + \alpha\beta_{ij} + \varepsilon^{(4)}_{ijk} \text{ (where } \varepsilon^{(4)}_{ijk} = \beta_j + \beta\gamma_{jk(i)});$$

$$Y_{ijk} = \mu + \alpha_i + \gamma_{k(i)} + \beta_j + \varepsilon^{(5)}_{ijk} \text{ (where } \varepsilon^{(5)}_{ijk} = \alpha\beta_{ij} + \beta\gamma_{jk(i)});$$

then the final steps in the analysis involve computing estimates of the terms shown algebraically in Table 2. The analysis of variance of the data in Table 1 was computed by using the computer program, BMD05V (Dixon, 1967). The results of this analysis are given in Table 3.

- - - - - - - - - - - - - - - - - -

TABLE 2

somewhere in here

- - - - - - - - - - - - - - - - - -

- - - - - - - - - - - - - - - - - -

TABLE 3

somewhere in here

- - - - - - - - - - - - - - - - - -

Approach 2: Large N

It should be obvious by now that the approach described above, in which subjects are coded as a main effects factor, can quickly become tedious as N increases. For N subjects, I treatments, and J repeated observations on each subject, a minimum of I(J - 1) + N ANOVA design variables would be needed to code the analysis for the model given in (2). The approach described in this section, however, requires fewer ANOVA design variables with the subsequent

advantages of requiring less coding effort and less computer storage.

Basically, this approach takes advantage of the similarities between the analysis of variance of crossed factorial designs and repeated-measures designs. If, for the moment, the data in Table 1 are considered as having been obtained according to a fully-crossed factorial design, the structural model for an observation at random could be written as

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon^{(6)}_{ijk}. \qquad (4)$$

By contrasting the sums of squares due to the components in (4) with those in (2), as has been done in Table 4, it can be seen that the sums of squares due to the components α_i, β_j, and $\alpha\beta_{ij}$ are identical.

- - - - - - - - - - - - - - - - - -

TABLE 4

somewhere in here

- - - - - - - - - - - - - - - - - -

Therefore, as could easily be demonstrated, $SS_{\varepsilon}^{(6)} = SS_{\gamma_{k(i)}} + SS_{\beta\alpha_{jk(i)}}$. In other words, the error sum of squares obtained by using a model for a fully crossed design is equal to the sum of the between-and within-group error sums of squares obtained by applying a repeated-measures model to the same data.

A general linear model, corresponding to the structural model given in (4) may be written as

$$Y = uX_1 + a_1X_2 + b_1X_3 + b_2X_4 + b_3X_5 + ab_1X_6 + ab_2X_7 + ab_3X_8 + e^{(6)}. \qquad (5)$$

(see Figure 2 for a more complete description of these terms).

Figure 2

somewhere in here

Using (5) as the "full model," the sums of squares due to effects are found by first computing the reduced models:

$$Y_{ijk} = \mu + \beta_j + \alpha\beta_{ij} + \varepsilon^{(7)}_{ijk} \text{ (where } \varepsilon^{(7)}_{ijk} = \varepsilon^{(6)}_{ijk} + \alpha_i);$$

$$Y_{ijk} = \mu + \alpha_i + \alpha\beta_{ij} + \varepsilon^{(8)}_{ijk} \text{ (where } \varepsilon^{(8)}_{ijk} = \varepsilon^{(6)}_{ijk} + \beta_j);$$

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon^{(9)}_{ijk} \text{ (where } \varepsilon^{(9)}_{ijk} = \varepsilon^{(6)}_{ijk} + \alpha\beta_{ij}).$$

The appropriate component sums of squares are then obtained as follows:

$$SS_\alpha = SS_{\varepsilon(7)} - SS_{\varepsilon(6)},$$

$$SS_\beta = SS_{\varepsilon(8)} - SS_{\varepsilon(6)},$$

$$SS_{\alpha\beta} = SS_{\varepsilon(9)} - SS_{\varepsilon(6)}.$$

There remains only the problem of computing the between- and within-group error terms ($SS_{\gamma_{k(i)}}$ and $SS_{\beta\gamma_{jk(i)}}$, respectively). From Table 3 the sum of squares due to the combined between sources of variance is given by

$$SS_{between} = J\Sigma_i\Sigma_k^{K_i}(\bar{Y}_{i\cdot k} - \bar{Y}_{\cdots})^2$$

$$= \Sigma_i\Sigma_k^{K_i}Y_{i\cdot k} \,/\, J - Y^2_{\cdots} \,/\, J\Sigma_i K_i.$$

Given this term, it is a simple matter of subtraction to find the sum of squares due to error between subjects. Hence,

$$SS_{\gamma_{k(i)}} = SS_{between} - SS_\alpha.$$

A convenient way of obtaining $SS_{between}$ is to compute the error sum of squares associated with the model

$$Y_{i \cdot k} = \mu + \varepsilon^{(10)} \tag{6}$$

In terms of general linear model, (6) may be written as (see Firgure 3)

$$Y = uX + e^{(10)}. \tag{7}$$

- - - - - - - - - - - - - - - - - -

Figure 3

somewhere in here

- - - - - - - - - - - - - - - - - -

It is a simple matter to show that this model leads to the following value for $SS_{\varepsilon}^{(10)}$:

$$SS_{\varepsilon(10)} = J^2 \Sigma_i \Sigma_k^{K_i} (\bar{Y}_{i \cdot k} - \bar{Y}_{...})^2$$

$$= \Sigma_i \Sigma_k^{K_i} Y_{i \cdot k}^2 - Y_{...}^2 / \Sigma K_i.$$

Note that this term is exactly equal to $J \times SS_{between}$. Therefore, the between error term may be obtained from

$$SS_{\gamma_{k(i)}} = SS_{\varepsilon(10)} / J - SS_{\alpha}.$$

The remaining sum of squares component, $SS_{\beta\gamma_{jk(i)}}$, is also obtained by subtraction,

$$SS_{\beta\alpha_{jk(i)}} = SS_{\varepsilon(6)} - SS_{\gamma_{k(i)}}.$$

The data in Table 1 were analyzed according to the techniques presented in this section. The program, BMD05V, was used first to compute the error sum of squares associated with the model given in (7). This gave $SS_{\varepsilon}^{(10)} = 50.000$.

The model given in (4) was run next. This resulted in the following error sums of squares for the full and reduced models: $SS_{\varepsilon(6)} = 24.833$, $SS_{\varepsilon(7)} = 30.042$, $SS_{\varepsilon(8)} = 125.992$, $SS_{\varepsilon(9)} = 25.252$. From these, the necessary treatment and measures sums of squares were computed: $SS_{\alpha} = 5.208$, $SS_{\beta} = 101.158$, $SS_{\alpha\beta} = .758$. Finally, the two error terms were computed: $SS_{\gamma_{k(i)}} = SS_{\varepsilon(10)} / 4 - SS_{\alpha} = 7.292$; $SS_{\beta\gamma_{jk(i)}} = SS_{\varepsilon(6)} - SS_{\gamma_{k(i)}} = 17.541$. All terms are, within rounding error, indentical to those given in Table 3.

References

Dixon, W. J. BMD Biomedical Computer Programs. Los Angeles: University of of California Press, 1967.

Draper, N., and Smith, H. Applied Regression Analysis. New York: Wiley, 1966.

Kirk, R. E. Experimental Design: Procedures for the Behavioral Sciences. Belmont, California: Brooks/Cole, 1968.

Li, J. C. R. Statistical Inference II. Ann Arbor, Michigan: Edwards Bros., 1964.

Winer, B. J. Statistical Principals in Experimental Design. New York: McGraw-Hill, 1962.

Winer, B. J. Statistical Principals in Experimental Design. New York: McGraw-Hill, 1971.

Footnotes

[1]The problem of whether to use a multivariate or univariate analysis is not relevant to this paper. Throughout this paper, the assumptions underlying the univariate analysis (e.g. homogeniety of within variance and covariance) are assumed. Concise treatments of these assumptions as well as additional references may be found in most intermediate level textbooks on statistics and experimental design.

[2]Although Li did treat split-plot designs in the later chapters of his volume, he did not show how these designs could be analyzed within a general regression framework.

Author

Olson, George H. Address: Dallas Independent School District, 3700 Ross Ave., Dallas, Texas, 75204. Title: Consultant--Research and Evaluation. Degrees: B.A., American International College; M.A., University of New Hampshire; Ph.D., Florida State University. Specialization: Educational Psychology, multivariate data analysis, evaluation design.

TABLE 1

2 × 4, Treatment × Measures Design

Treatments	Subjects	Measures			
		B_1	B_2	B_3	B_4
A_1	S_1	2	3	8	10
	S_2	7	5	7	9
	S_3	3	4	7	9
A_2	S_1	2	2	5	9
	S_2	3	5	7	8

TABLE 2

Sums of Squares Components for the

Repeated Measures Design Given in Table 1

Source of Variation	d. f.	Sums of Squares
Between Subjects	N - 1	
Treatment	I - 1	$SS_{\varepsilon(2)} - SS_{\varepsilon(1)}$
Subj. within Treatment	N - I	$SS_{\varepsilon(3)} - SS_{\varepsilon(1)}$
Within Subjects	N(J - 1)	
Measures	J - 1	$SS_{\varepsilon(4)} - SS_{\varepsilon(1)}$
Treatment × Measures	(I - 1)(J - 1)	$SS_{\varepsilon(5)} - SS_{\varepsilon(1)}$
Subj. × Measures w. Treatment	(N - I)(J - 1)	$SS_{\varepsilon(1)}$

TABLE 3

Least Squares Analysis of Variance for the Design Given in Table 1

Source of Variation	d. f.	Sums of Squares
Between Subjects	4	
Treatment	1	5.208
Subj. w. Treatment	3	7.292
Within Subjects	15	
Measures	3	101.158
Treat. × Measures	3	.758
Subj. × Measures w. Treat.	9	17.542

TABLE 4

Similarities Between Sums of Squares Components for Repeated Measures and Crossed Analysis of Variance

Repeated Measures Design	Crossed Design
$SS_{between} = J\Sigma_i\Sigma_k^{K_i}(\overline{Y}_{i\cdot k} - \overline{Y}_{\cdots})^2$	$SS_{between} = J\Sigma_i\Sigma_k^{K_i}(\overline{Y}_{i\cdot k} - \overline{Y}_{\cdots})^2$
$SS_{\alpha} = J\Sigma_i K_i(\overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdots})^2$	$SS_{\alpha} = J\Sigma_i K_i(\overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdots})^2$
$SS_{\beta} = \Sigma_i K_i\Sigma_j(\overline{Y}_{\cdot j\cdot} - \overline{Y}_{\cdots})^2$	$SS_{\beta} = \Sigma_i K_i\Sigma_j(\overline{Y}_{\cdot j\cdot} - \overline{Y}_{\cdots})^2$
$SS_{\alpha\beta} = \Sigma_i K_i\Sigma_j(\overline{Y}_{ij\cdot} - \overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdot j\cdot} + \overline{Y}_{\cdots})^2$	$SS_{\alpha\beta} = \Sigma_i K_i\Sigma_j(\overline{Y}_{ij\cdot} - \overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdot j\cdot} + \overline{Y}_{\cdots})^2$
$SS_{within} = \Sigma_i\Sigma_j\Sigma_k^{K_i}(Y_{ijk} - \overline{Y}_{ij\cdot})^2$	$SS_{within} = \Sigma_i\Sigma_j\Sigma_k^{K_i}(Y_{ijk} - \overline{Y}_{ij\cdot})^2$
$SS_{\gamma_{k(i)}} = J\Sigma_i\Sigma_k^{K_i}(\overline{Y}_{i\cdot k} - \overline{Y}_{i\cdot\cdot})^2$	
$SS_{\beta\gamma_{jk(i)}} = \Sigma_i\Sigma_j\Sigma_k^{K_i}(Y_{ijk} - \overline{Y}_{i\cdot k} - \overline{Y}_{ij\cdot} - \overline{Y}_{i\cdot\cdot})^2$	
$SS_{total} = \Sigma_i\Sigma_j\Sigma_k^{K_i}(Y_{ijk} - \overline{Y}_{\cdots})^2$	$SS_{total} = \Sigma_i\Sigma_j\Sigma_k^{K_i}(Y_{ijk} - \overline{Y}_{\cdots})^2$

$$
\begin{array}{c}
Y \\
\begin{bmatrix}
Y_{111} \\ Y_{112} \\ Y_{113} \\ Y_{121} \\ Y_{122} \\ Y_{123} \\ Y_{131} \\ Y_{132} \\ Y_{133} \\ Y_{141} \\ Y_{142} \\ Y_{143} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \\ Y_{231} \\ Y_{232} \\ Y_{241} \\ Y_{242}
\end{bmatrix}
\end{array}
=
\begin{array}{c}
\begin{matrix} X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 & X_8 & X_9 & X_{10} & X_{11} \end{matrix} \\
\begin{bmatrix}
1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & -1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & -1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\
1 & 1 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\
1 & 1 & 0 & -1 & 0 & -1 & 0 & 0 & -1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\
1 & 1 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\
1 & 1 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & -1 \\
1 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 \\
1 & 1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & -1 \\
1 & -1 & 0 & 0 & 1 & 1 & 1 & 1 & -1 & -1 & -1 \\
1 & -1 & 0 & 0 & -1 & 1 & 1 & 1 & -1 & -1 & -1 \\
1 & -1 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & 0 & 0 \\
1 & -1 & 0 & 0 & -1 & -1 & 0 & 0 & 1 & 0 & 0 \\
1 & -1 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 1 & 0 \\
1 & -1 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 1 & 0 \\
1 & -1 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \\
1 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 1
\end{bmatrix}
\end{array}
\times
\begin{bmatrix}
u \\ a_1 \\ c_{1(1)} \\ c_{2(1)} \\ c_{1(2)} \\ b_1 \\ b_2 \\ b_3 \\ ab_1 \\ ab_2 \\ ab_3
\end{bmatrix}
+
\begin{array}{c} e \\
\begin{bmatrix}
e^{(1)}_{111} \\ e^{(2)}_{112} \\ \vdots \\ e^{(1)}_{242}
\end{bmatrix}
\end{array}
$$

$$
\begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{113} \\ Y_{121} \\ Y_{122} \\ Y_{123} \\ Y_{131} \\ Y_{132} \\ Y_{133} \\ Y_{141} \\ Y_{142} \\ Y_{143} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \\ Y_{231} \\ Y_{232} \\ Y_{241} \\ Y_{242} \end{bmatrix}
=
\begin{bmatrix}
X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 & X_8 \\
1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 \\
1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 \\
1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 \\
1 & -1 & 1 & 0 & 0 & -1 & 0 & 0 \\
1 & -1 & 1 & 0 & 0 & -1 & 0 & 0 \\
1 & -1 & 0 & 1 & 0 & 0 & -1 & 0 \\
1 & -1 & 0 & 1 & 0 & 0 & -1 & 0 \\
1 & -1 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & -1 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & -1 & -1 & -1 & -1 & 1 & 1 & -1 \\
1 & -1 & -1 & -1 & -1 & 1 & 1 & -1
\end{bmatrix}
\times
\begin{bmatrix} u \\ a_1 \\ b_1 \\ b_2 \\ b_3 \\ ab_1 \\ ab_2 \\ ab_3 \end{bmatrix}
+
\begin{matrix} e \\ \begin{bmatrix} e^{(6)}_{111} \\ e^{(6)}_{112} \\ \vdots \\ e^{(6)}_{242} \end{bmatrix} \end{matrix}
$$

$$
\begin{matrix} Y \\ \begin{bmatrix} Y_{1\cdot1} \\ Y_{1\cdot2} \\ Y_{1\cdot3} \\ Y_{2\cdot1} \\ Y_{2\cdot2} \end{bmatrix} \end{matrix}
=
\begin{matrix} X \\ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{matrix}
\times u +
\begin{matrix} e \\ \begin{bmatrix} e_{1\cdot1}^{(10)} \\ e_{1\cdot2}^{(10)} \\ e_{1\cdot3}^{(10)} \\ e_{2\cdot1}^{(10)} \\ e_{2\cdot2}^{(10)} \end{bmatrix} \end{matrix}
$$

List of Captions

Figure 1. One of many possible matrix representations of the model given in equation (3).

Figure 2. One of many possible matrix representations of the model given in equation (5).

Figure 3. Matrix representation of the model given in equation (7).

It is the policy of the SIG=Multiple Linear Regression and of Viewpoints to consider for publication articles dealing with the theory and the application of Multiple Linear Regression. Manuscripts should be submitted to the Editor as an original, double-spaced typed copy. A cost of $1 per page should be sent with the submitted paper. Reprints are available to the authors from the Editor. Reprints should be ordered at the time the paper is submitted and 20 reprints will cost $.50 per page of manuscript. Prices may be adjusted as necessary in the future.

Judy T. McNeil, Chairman
SIG=Multiple Linear Regression
Isadore Newman, Editor
Multiple Linear Regression Viewpoints

The University of Akron
Akron, Ohio 44325

BOOKS - SPECIAL 4th CLASS RATE

TABLE OF CONTENTS