Multiple Linear Regression Viewpoints

A Publication sponsored by the
American Educational Research Association's
Special Interest Group on Multiple Linear Regression

MLRV

Volume 20 • Number 1 • Spring 1993

Table of Contents

Editor's Comment

Welcome to the new format of *Multiple Linear Regression Viewpoints (MLRV)*. Starting with this issue, the journal will accommodate three sections: General, Teaching, and SIG/Multiple Linear Regression News. The general section will contain research articles intended for a broad audience of readers interested in theoretical and applied aspects of multiple linear regression models and their extensions. The collection of papers on teaching will include articles that focus on effective teaching of multiple regression or related concepts. Finally, current news items from the AERA Special Interest Group on Multiple Linear Regression (SIG/MLR) will be published in the news section, keeping members and other interested individuals abreast of SIG happenings. For subscription and potential author information, consult the footnote below and the revised Information for Contributors on the inside backcover.

The current issue of *MLRV* features theoretical papers on (a) the possible estimation of statistical power for an ANOVA design from reliability, sample and effect size, and significance level, (b) an application of multiple regression in determining the number of important factors in an exploratory factor analysis, and (c) the use of multiple regression in determining the effects of truncation on the gamma distribution. In the teaching section, possible applications of the analysis of covariance and ordinal and criterion scaling in regression are discussed. In preparation for the 1993 AERA meeting in Atlanta, Georgia, April 12-16, 1993, the SIG/MLR program is listed in the news section.

This issue of *MLRV* is the first published under my editorship. For that reason, allow me to add a personal note. I thank Isadore Newman, past Editor of *MLRV*, Bruce Thompson, 1992/93 SIG/Multiple Linear Regression Chair, and other SIG/MLR members for supporting my nomination and appointment as Editor. Also, without financial support from The University of Toledo, *MLRV* in its new format could not have been published.

The past year was filled with many positive learning experiences, some frustrations, and several 25-hour days. Initial ideas for improving the journal that I presented during the 1992 SIG/MLR business meeting could not all be implemented at once. I hope and anticipate to do so over the next several issues. I invite your suggestions for expanding and improving the journal. Only with your active involvement will we continue to publish, improve, and expand the readership of *Multiple Linear Regression Viewpoints*.

Multiple Linear Regression Viewpoints (ISSN 0195-7171) is published by the AERA Special Interest Group on Multiple Linear Regression (SIG/MLR) through The University of Toledo. Subscription and SIG membership information can be obtained from Steven Spaner, Executive Secretary, SIG/MLR, School of Education, University of Missouri, St. Louis, MO 63121. *MLRV* abstracts appear in CIJE, the ERIC system, and microform copies are available from University Microfilms International, 300 North Zeeb Road, Ann Arbor, MI 48106. *MLRV* is listed in the *EBSCO Librarians Handbook*. Second-class postage paid at Toledo, OH. POSTMASTER: Send address changes to Ralph Mueller, Educational Research, College of Education and Allied Professions, The University of Toledo, Toledo, OH 43606.

Equations Which Include the Reliability of the Dependent Variable for Estimating the Power of a Two Group ANOVA Design

Dale Shaw
University of Northern Colorado

Rebecca Kopriva
California State University, Fresno

Susan Tracz
California State University, Fresno

Regression equations were obtained relating the power of the completely randomized, fixed-effects, one-factor ANOVA for two groups to four independent variables: effect size, alpha level, sample size, and the reliability with which the dependent variable in the ANOVA design is measured. One equation was singled out for discussion due to the ease with which power estimates could be calculated using it and their degree of accuracy. The effect of reliability on the power estimates, as well as the nature of the relationship of power to the independent variables, was also discussed.

The purpose of this study was to obtain regression equations for estimating the power of the completely randomized, fixed-effects, one-factor Analysis of Variance (ANOVA) for two groups from four independent variables: effect size, alpha level, sample size, and the reliability of the dependent measure. Of particular interest was the degree and nature of the contribution of the reliability variable (the reliability with which the dependent variable in the ANOVA design is measured) to estimating a design's power.

Background

Hopkins, Coulter, and Hopkins (1981) and Cohen (1988) presented tables for estimating the power of one-factor ANOVAs for various sample sizes, alpha levels, and effect sizes. These power estimates are not derived from an equation which relates these variables; rather, they are obtained by a series of steps which require a critical F-value to be obtained from a central F-distribution for a specified alpha level, and the calculation of a non-centrality parameter to determine which non-central F-distribution is then to be used to determine the power estimate. This process does not reduce to a single equation, nor does it include the effect of the reliability of the dependent measure on power.

Cleary and Linn (1969) demonstrated that reliability does indeed have a direct effect on the power of a one-factor ANOVA. Sutcliffe (1980) confirmed this and went on to show that the effect was both direct and monotonic. Although Nicewander and Price (1978; 1983), Overall and Woodward (1975; 1976), and Zimmerman and Williams (1986) have addressed the extent of the combined effect of reliability, sample size, and alpha level on power for a few specific values of these variables, no functional relationships have been presented which can be expressed in equation form. Kopriva and Shaw (1991) extended the work of Hopkins, Coulter, and Hopkins (1981) by deriving tables for estimating power which included among the predictor variables the dependent measure's reliability. Results of their work indicated that for certain combinations of values for effect size, alpha, and sample size, the effect of reliability on power was substantial.

Since the power estimates tabled by Hopkins, Coulter, and Hopkins (1981) and Kopriva and Shaw (1991) were derived from a series of steps and not from equations which permit their direct calculation, it seemed reasonable to determine if an equation could be derived from the tables which relates power functionally to the four independent variables: effect size, alpha, sample size, and reliability. Thus, regression equations were sought relating the power estimates (dependent variable) to the predictor variables above. Several such equations are presented in this paper.

Data

All data were obtained from the tables for estimating power for the completely randomized, fixed-effects, one-factor ANOVA for two groups which are presented in Hopkins, Coulter, and Hopkins (1981) and Kopriva and Shaw (1991). A portion of the data for α equal to .05 is presented in Table 1.

Table 1 Power Estimates for Two Groups (α = .05)

Effect Size	n per group	Reliability .10	.20	.30	.40	.50	.60	.70	.80	.90	1.0
	5	.02	.03	.03	.03	.03	.03	.03	.03	.03	.03
	10	.03	.03	.03	.03	.04	.04	.04	.04	.04	.04
	15	.03	.03	.04	.04	.04	.04	.04	.04	.04	.04
.10	25	.03	.04	.04	.04	.04	.04	.04	.05	.05	.05
	50	.04	.04	.05	.05	.05	.06	.06	.07	.07	.07
	100	.04	.05	.06	.07	.07	.08	.08	.10	.11	.11
	200	.04	.07	.08	.09	.11	.12	.13	.15	.16	.17
	400	.07	.10	.12	.14	.16	.19	.22	.24	.28	.30
	5	.03	.04	.04	.04	.04	.05	.05	.05	.05	.05
	10	.04	.04	.05	.05	.06	.06	.06	.07	.07	.08
	15	.04	.05	.05	.06	.07	.07	.08	.09	.09	.10
.25	25	.04	.06	.07	.08	.09	.10	.11	.12	.13	.13
	50	.06	.08	.10	.12	.14	.16	.18	.20	.22	.24
	100	.08	.13	.16	.20	.24	.27	.32	.35	.38	.41
	200	.12	.21	.27	.35	.42	.49	.54	.61	.65	.71
	400	.19	.35	.49	.62	.72	.78	.84	.88	.93	.93
	5	.04	.05	.06	.06	.07	.08	.08	.09	.10	.10
	10	.05	.07	.08	.10	.12	.13	.14	.16	.17	.18
	15	.06	.08	.11	.13	.15	.17	.19	.22	.24	.26
.50	25	.08	.12	.15	.19	.23	.27	.31	.34	.38	.41
	50	.12	.20	.28	.34	.42	.48	.54	.60	.65	.70
	100	.19	.36	.49	.60	.70	.80	.84	.88	.92	.94
	200	.35	.61	.78	.89	.94	.97	.99	.99	.99+	.99+
	400	.61	.89	.97	.99	.99+	.99+	.99+	.99+	.99+	.99+
	5	.07	.09	.12	.14	.17	.19	.21	.24	.26	.28
	10	.10	.15	.21	.27	.32	.37	.42	.48	.51	.56
	15	.13	.22	.30	.38	.46	.53	.60	.66	.71	.75
1.0	25	.20	.34	.48	.59	.69	.76	.83	.87	.91	.94
	50	.35	.60	.77	.88	.94	.97	.98	.99	.99+	.99+
	100	.61	.88	.97	.99+	.99+	.99+	.99+	.99+	.99+	.99+
	200	.90	.99	.99+	.99+	.99+	.99+	.99+	.99+	.99+	.99+
	400	.99	.99+	.99+	.99+	.99+	.99+	.99+	.99+	.99+	.99+

The observations of the power estimates given all combinations of eight sample sizes (ranging from 5 to 400), four alpha levels (.01, .025, .05, and .10), four effect sizes (.1σ, .25σ, .5σ, and 1σ), and ten reliabilities (.1, .2, ..., 1.0) served as the 1280 data points in the study.

Model Development and Preliminary Results

Variables in the models and discussion below are identified as follows:

1. Dependent variable: power (P).

2. Independent variables: sample size (n), level of significance (α), effect size which is the expected difference in means expressed in standard deviation units (d), and reliability estimate (r) for the dependent measure.

Initial inspection of the data set in Table 1 indicated a five-dimensional "surface" with substantial curvature in certain areas and little or no curvature in others. When any of the independent variables were increased in value either singly or in combination with others, P was increased and ultimately became asymptotic to the hyperplane P = 1. If such a surface could be determined, the fit would be essentially without error because all data would lie within the surface, not around or near it.

As a starting point, P was regressed onto the four predictor variables using the model below, a hyperplane with no curvature or warp.

$$P = -.274 + .001n + .170\alpha + .641d + .282r$$
$$R^2 = .698, \quad SE = .197 \qquad [1]$$

The SE of .197 for this model may be interpreted as there being approximately an average of .197 error made in estimating P for all 1280 data points. To have a model which would produce estimates of P within .02 or even .05 might be useful, but a model with an average error of .197 is not, since P itself ranges from 0

to 1. Equation 1 was not expected to fit the data well because it did not provide for curvature.

Adding all two-way interaction terms and squares to the model identified by Equation 1 in an effort to account for curvature or warp, brought the R^2 to .766. Numerous nonlinear transformations of variables in this expanded model were tried. Logarithmic, exponential, and square root transformations faired no better than polynomial fits. Not only were these models more difficult to use for computation and more difficult to interpret, they also did not capture the asymptotic nature of the surface's curvature to the hyperplane P = 1.

In an attempt to improve the fit substantially, separate models were sought for various values of reliability and effect size. Typical of the nature and complexity of the fits possible for r = 1.0 are the three models below, which are the one-term, two-term, and three-term models which produced the highest R^2 values from all possible regressions onto the predictor set consisting of d, α, n, their squares, and all possible two or more-way interactions of these six variables.

$$P = .017 + .036nd^2$$
$$R^2 = .865, \quad SE = .110 \qquad [2]$$

$$P = -.042 + .037nd^2 + .242\alpha$$
$$R^2 = .949, \quad SE = .052 \qquad [3]$$

$$P = -.029 + .035nd^2 + .194\alpha + .076nd\alpha$$
$$R^2 = .957, \quad SE = .049 \qquad [4]$$

The improvement in the R-squares here is dramatic. Adding additional terms produced R-squares in excess of .99 after 18 terms. The substantial improvement in R^2 was due to the inclusion of terms allowing complex interactions (such as nd^2) and the deletion of observations having power values greater than .95 or less than .20. Deleting these observations permitted the complex polynomial interaction terms to fit the surface without being constrained in the areas where the surface flattened out near P = 0 and P = 1.

At this stage in the study, a decision was made regarding the nature and types of additional models to try. Although models were being identified which demonstrated improved fits, they were computationally difficult to use. In addition, they were not likely to zero in on the true nature of the relationship of power to the independent variables; rather, they were simply surfaces that approximated the true relationship. Since the tables already existed, models providing simple computational estimates were sought which one might easily remember rather than having to carry or refer to tables.

The data set was modified to include only the 384 observations where P was greater than or equal to .20 and less than or equal to .95. With this modified data set, logarithmic and exponential transformations still faired no better than polynomial fits; however, square root transformations did offer enhanced prediction. Thus, models were created using the four independent variables, their squares and square roots, and all two or more-way interaction terms incorporating these twelve variables. Using this set of predictor variables, all possible regressions were obtained.

Results And Discussion

Of all the models produced by all possible regressions, the model below (Equation 5) was judged to have the best balance between (a) the accuracy of the power estimates and (b) the simplicity of computation.

$$P = .236d\sqrt{nr} + 3.178\alpha - .372$$
$$R^2 = .971, \quad SE = .040 \qquad [5]$$

This model produced the highest R^2 value of all possible two-term regression models. The R^2 of .971 for this model was substantially higher than the R^2 of .787 for the best one-term model. Furthermore, of the best three-term models which had R-squares in excess of .971, none exceeded .981 and all were considerably more complex to use for computations.

Equation 5 may be simplified somewhat by specifying values for α. For example, for the 106 observations where α = .05, Equation 6 provides extremely accurate power estimates and is more easily recalled.

$$P = .24d\sqrt{nr} - .18$$
$$R^2 = .997, \quad SE = .029 \qquad [6]$$

Power estimates produced by Equation 6 are on average within .029 of the actual power figures reported in the Kopriva and Shaw (1991) tables.

Equations 5 and 6 do have one unattractive feature. For large values of n or d they can produce power estimates which are 1.0 or greater. In this case, the user may think of the power as being greater than .99. Also, for very small values of n, d, or r, the equations can produce power estimates less than zero, in which case the power would be thought of as α.

The reliability of the dependent variable does have some effect on power. Equations 5 and 6 do reflect the direct and monotonic nature of the effect indicated by Cleary and Linn (1969) and Sutcliffe (1980). The extent of the effect might best be illustrated with an example from Table 1. If one ignores the reliability of the measuring device, the power estimates of Hopkins, Coulter, and Hopkins (1981) and Cohen (1988) are the same as those found in the last column of Table 1 where the reliability is 1.0. Thus if one is using an instrument with less than perfect reliability, the Hopkins, Coulter, and Hopkins (1981) and Cohen (1988) power estimates are inflated. Equations 5 and 6 indicate that these overestimates are high by a factor of $\sqrt{r}$.

References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cleary, T. A., & Linn, R. L. (1969). Error of measurement and the power of a statistical test. *British Journal of Mathematical and Statistical Psychology, 22*, 49-55.

Hopkins, K. D., Coutler, D. K., & Hopkins, B. R. (1981). Tables for quick power estimates when comparing means. *The Journal of Special Education, 15*, 389-394.

Kopriva, R. J., & Shaw, D. G. (1991). Power estimates: The effect of dependent variable reliability on the power of one-factor ANOVA's. *Educational & Psychological Measurement*, Autumn.

Nicewander, W. A., & Price, J. M. (1978). Dependent variable reliability and the power of statistical tests. *Psychological Bulletin, 85*, 405409.

Nicewander, W. A., & Price, J. M. (1983). Reliability of measurement and the power of statistical tests. *Psychological Bulletin, 95*, 524-533.

Overall, J. E., & Woodward, J. A. (1975). Unreliability of difference scores: A paradox for measurement of change. *Psychological Bulletin, 82*, 85-86.

Overall, J. E., & Woodward, J. A. (1976). Reassertion of the paradoxical power of tests of significance based on unreliable difference scores. *Psychological Bulletin, 83*, 776-777.

Sutcliffe, J. P. (1980). On the relationship of reliability to statistical power. *Psychological Bulletin, 88*, 509-515.

Williams, R. H., & Zimmerman, D. H. (1981). Error of measurement and statistical inference: Some anomalies. *Journal of Experimental Education, 49*, 71-73.

Zimmerman, D. W., & Williams, R. J. (1986). Note on the reliability of experimental measures and the power of significance tests. *Psychological Bulletin, 100*, 123-124.

Using Multiple Regression to Determine the Number of Factors to Retain in Factor Analysis

Keith Zoski
University of Toledo

Stephen Jurs
University of Toledo

A simple, analytical approach using multiple regression analysis is presented as a way to determine the number of factors to retain in a factor analysis. Two regression lines are found from the points in a scree plot and the number of retained factors is chosen at a point that maximally separates the two regression lines. Applications of the technique to data from the literature suggest that the results agree closely with solutions based on the somewhat subjective visual scree test and may be better than those from the analytical CNG method.

The number of factors to retain in a factor analysis has long been an important problem (Hakstian & Muller, 1973; Crawford, 1975; Horn & Engstrom, 1979; Hakstian, Rogers, & Cattell, 1982; Kano, 1990). This is critical because it demands a decision that affects the factor parameters and the interpretability of the factors (Lambert, Wildt, & Durand, 1990).

The most frequently used method for determining the number of factors is to select only those factors whose eigenvalues exceed 1.0 (Kaiser, 1970; Kaiser & Caffrey, 1965). Critics of this method (Gorsuch, 1983) are concerned that many times there is not a clear break among the eigenvalues at the 1.0 value and that underestimating or overestimating communalties would change the number of retained factors when the eigenvalues greater than 1.0 rule is used. Therefore, the selection or deletion of some factors may be a function of an arbitrary rule that is not sensitive to the nature or pattern of the data.

An approach that considers the relation of the eigenvalues to one another as well as their actual values is the scree test. Cattell (1966) first proposed the scree test to separate trivial from non-trivial factors. The procedure required one to plot the eigenvalues in decreasing order. The graph contained the values of the eigenvalues on the ordinate and the factors on the abscissa. A straight line could be drawn on the graph through the points associated with the smaller eigenvalues. The points near this line were judged trivial and the points above and to the left of the line were judged to be non-trivial (Cattell, 1978; Cattell & Vogelman, 1977; Cattell & Jaspers, 1967). Horn and Engstrom (1979) provided statistical support for the scree test.

Cattell and Vogelman (1977) and Cattell (1978) presented guidelines for this visual procedure. These guidelines, as summarized by Zoski and Jurs (1990), are:

1. Three sequential points form an undesirably low limit for drawing a scree.
2. The points on the part of the curve that one should consider scree should fit tightly.
3. The slope of the scree should not approach vertical. Instead, it should have an angle of 40° or less from the horizontal, that is, a slope of the tangent less than -.84.
4. In the case of multiple screes falling below 40°, the first scree on the left is the arbitrator.
5. Generally, a sharp, albeit sometimes small, break in the vertical level exists between the last point of the curve and the left-most point of the scree.

However, problems with this procedure can occur when there are multiple breaks in the eigenvalue curve, with several straight lines in the graph. It may be difficult to select as well as to justify one break over another (Gorsuch, 1983). Moreover, critics of visual approaches are concerned about researchers seeing what they want to see in the data unless they are constrained by a mechanical decision-making rule. This position is demeaning to the researchers and shifts the demand for objectivity over subjectivity to the final stages of research (decisions and conclusions) and ignores the more critical phase (research problem definition and variable selection). An analytical, programmable approach does have some appeal, if it provides results that are consistent with those obtained using the guidelines above. We propose that multiple regression techniques can be used to provide such a solution.

The Multiple Regression Approach

Gorsuch and Nelson (1981) developed an analytical

method for determining the number of factors to retain. The Cattell-Nelson-Gorsuch scree test requires one to compare the slope of the first three roots with the slope of the next three roots. Then the slope of roots 2, 3, and 4 is compared with the slope of roots 5, 6, and 7. This process continues so that all sets of three factors are compared. The number of factors is found where the difference between the slopes is greatest.

Because only three points are used to determine the slopes, the analysis is not based on as much information as is possible. Thus, we propose a somewhat different approach using multiple regression to accomplish the same thing; objective determination of the number of factors that is sensitive to the data.

The rationale for a regression approach is straightforward. It parallels the statistical work of Horn and Engstrom (1979) on Cattell's scree test using Bartlett's chi-square test (1950, 1951). The method used here provides virtually the same decision as the visual scree test but can be easily programmed. It uses a regression approach where the ordered eigenvalues are thought of as points in a scatterplot. One can then form two regression lines, one for the important factors and another for the scree or trivial factors. The decision about the number of factors to retain corresponds with the maximal differences between the two regression lines.

To use all the eigenvalues, form and compare these pairs of regression lines:

line 1	line 2
(points 1 through 3)	(points 4 through m)
line 3	line 4
(points 1 through 4)	(points 5 through m)
line 5	line 6
(points 1 through 5)	(points 6 through m)
•	•
•	•
•	•
line (m-2)	line (m-1)
(points 1, 2, ... (m-3))	(points (m-2), (m-1), m)

The slope of these regression lines will, of course, be negative and can be compared by the usual formulae (Howell, 1987, pp. 222, 239-240):

$$b = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - (\sum X)^2} \qquad [1]$$

$$t = \frac{b_1 - b_2}{s_{b_1 - b_2}} \qquad [2]$$

with

$$s_{b_1 - b_2} = \sqrt{\frac{s^2_{Y \cdot X_1}}{s^2_{X_1}(N_1 - 1)} + \frac{s^2_{Y \cdot X_2}}{s^2_{X_2}(N_2 - 1)}} \qquad [3]$$

and when homogeneity of error variances is assumed, we can pool:

$$s^2_{Y \cdot X} = \frac{(N_1 - 2)\left(s^2_{Y \cdot X_1}\right) + (N_2 - 2)\left(s^2_{Y \cdot X_2}\right)}{N_1 + N_2 - 4} \qquad [4]$$

The salient factors are those with eigenvalues in the odd numbered line of the line pair where the t-test is maximized (highest value). The even numbered line of the pair denotes the scree line. Some analysts may choose not to include the last factor. Note that neither the CNG nor the multiple regression approach would be appropriate when there are only one or two factors.

Examples

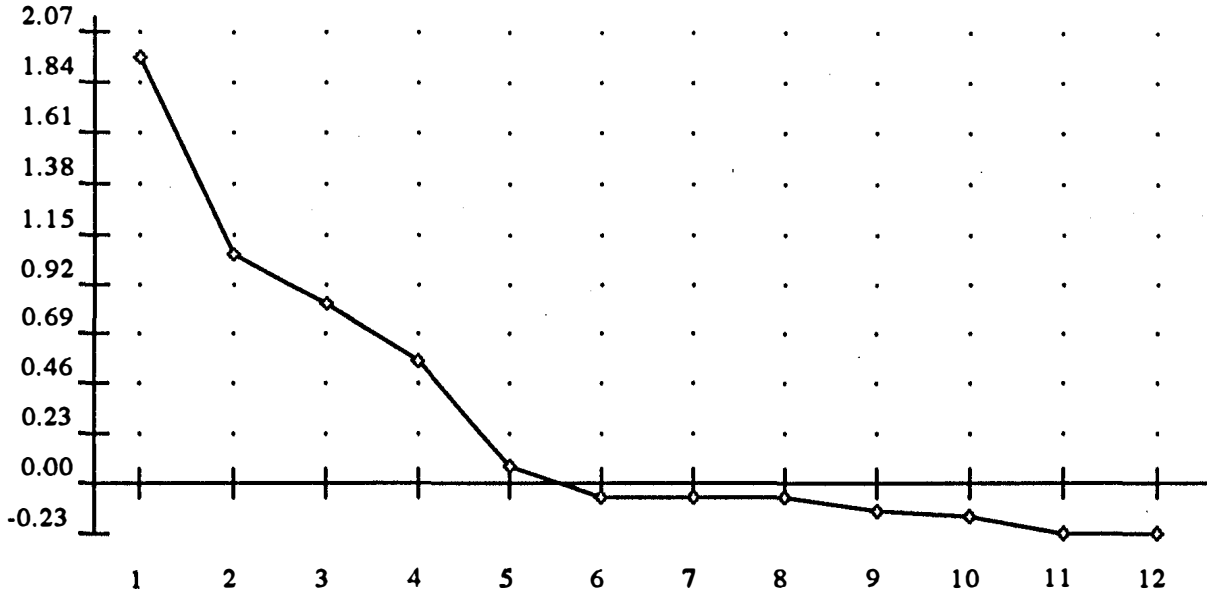
We have compared the multiple regression approach to the CNG approach using several data sets from the literature. Preliminary results indicate that the multiple regression approach usually agrees with a visual scree test and often provides a better solution than the CNG method.

Example 1 is based on eigenvalues taken from Cliff (1970). The eigenvalues are plotted in Figure 1. Table 1 contains the slopes of the regression lines and the t values for the multiple regression approach and the slopes and differences for the CNG approach (* indicates highest value). Note that in this case both procedures indicate that there are five factors and this agrees with a visual analysis of the plotted eigenvalues in Figure 1.

Table 1 Comparison of Multiple Regression and CNG Approaches: Example 1

	MR			CNG		
# of factors	slope 1	slope 2	t	slope 1	slope 2	difference
3	-.563	-.071	4.044	-.563	-.310	.253
4	-.441	.038	6.713	-.250	-.067	.183
5	-.426	-.032	8.448*	-.377	-.001	.376*
6	-.380	-.038	6.814	-.310	-.032	.278
7	-.323	-.042	3.752	-.067	-.043	.024
8	-.272	-.038	1.899	-.001	-.048	.047
9	-.234	-.040	0.890	-.032	-.040	.007

Figure 1 Scree Plot from Cliff (1970, p. 165, CS 600).

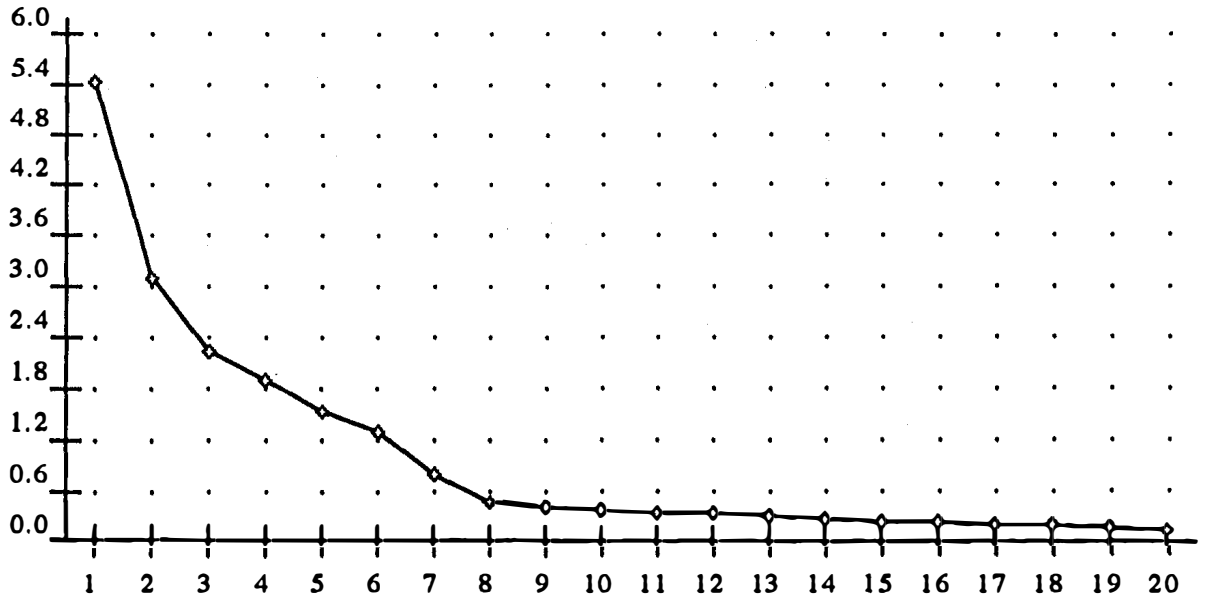


The second example is taken from Tucker, Koopman, & Linn (1969, p. 442). The plot of the eigenvalues is given in Figure 2 and the results from the multiple regression approach and the CNG approach are listed in Table 2. The data set was meant to have seven factors. The CNG approach yielded three factors and the multiple regression approach did yield the expected seven factors. Visual inspection of Figure 2 confirms that a seven factor solution is appropriate.

Table 2 Comparison of Multiple Regression and CNG Approaches: Example 2

	MR			CNG		
# of factors	slope 1	slope 2	t	slope 1	slope 2	difference
3	-1.595	-.084	6.346	-1.595	-.300	1.295*
4	-1.149	-.067	6.985	-.610	-.360	.250
5	-.904	-.051	7.327	-.360	-.415	.055
6	-.737	-.033	7.405	-.300	-.195	.105
7	-.651	-.023	7.665*	-.360	-.045	.315
8	-.590	-.022	7.563	-.415	-.030	.385
9	-.525	-.021	6.694	-.195	-.020	.175
10	-.465	-.021	5.507	-.045	-.025	.020
11	-.413	-.021	4.277	-.030	-.025	.005
12	-.367	-.020	3.176	-.020	-.025	.005
13	-.328	-.020	2.266	-.025	-.020	.005
14	-.295	-.019	1.555	-.025	-.020	.005
15	-.267	-.021	1.013	-.025	-.020	.005
16	-.243	-.020	.622	-.020	-.015	.005
17	-.222	-.025	.335	-.020	-.025	.005

Figure 2 Scree Plot from Tucker, Koopman and Linn (1969, p. 442, Middle 7)

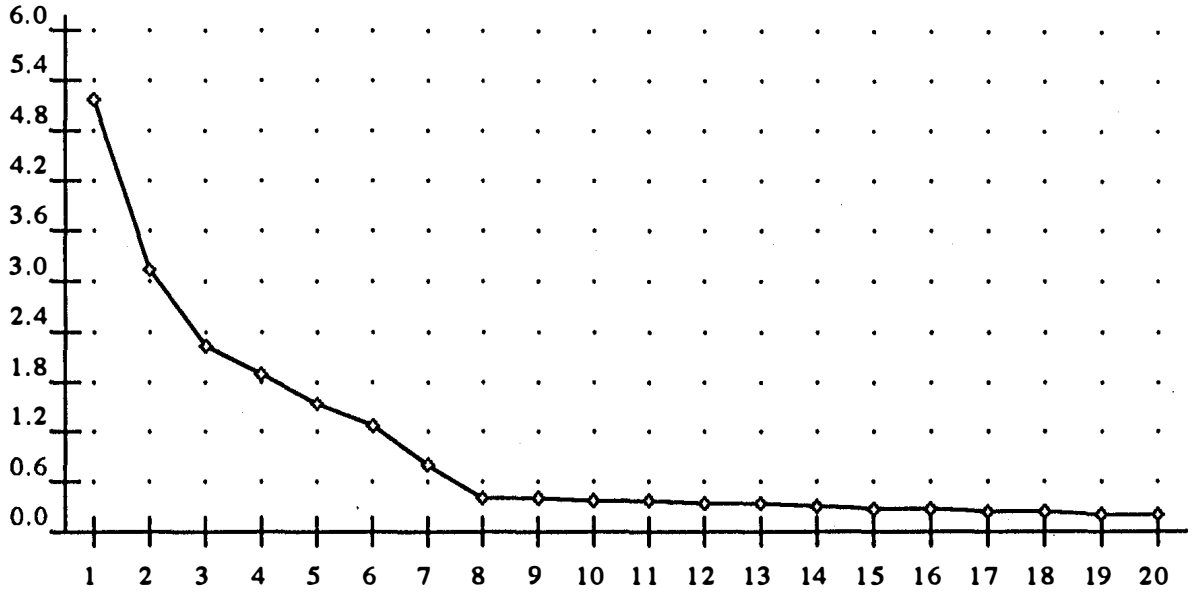


The third example was also taken from Tucker, Koopman, & Linn (1969, p. 442). This data set was intended to have seven factors and a visual inspection of the scree plot in Figure 3 suggests that there are seven factors. The analyses presented in Table 3 indicate that the CNG approach yielded only three factors and the multiple regression approach yielded eight factors. This example shows that results from the multiple regression approach may not always agree with results from a visual approach, but the technique seemed to work better than the CNG method for these data.

Table 3 Comparison of Multiple Regression and CNG Approaches: Example 3

			MR			CNG
# of factors	slope 1	slope 2	t	slope 1	slope 2	difference
3	-1.475	-.081	5.855	-1.475	-.315	1.160*
4	-1.071	-.063	6.818	-.610	-.365	.245
5	-.850	-.047	7.522	-.345	-.440	.095
6	-.702	-.029	7.944	-.315	-.210	.105
7	-.625	-.018	8.369	-.365	-.015	.350
8	-.574	-.018	8.443*	-.440	-.010	.430
9	-.513	-.018	7.401	-.210	-.010	.200
10	-.455	-.195	5.974	-.015	-.025	.010
11	-.403	-.019	4.554	-.010	-.025	.015
12	-.358	-.018	3.341	-.010	-.020	.010
13	-.320	-.178	2.365	-.025	-.015	.010
14	-.287	-.177	1.611	-.025	-.020	.005
15	-.260	-.018	1.047	-.020	-.020	.000
16	-.235	-.015	.647	-.015	-.020	.005
17	-.215	-.015	.356	-.020	-.015	.005

Figure 3 Scree Plot from Tucker, Koopman & Linn (1969, p. 442, Formal 7)



Conclusions

Multiple regression is a versatile set of techniques for which there are diverse applications. Our results indicate that multiple regression can be used successfully to determine how many factors to retain in a factor analysis. Preliminary analyses suggest that the results will usually agree with results from a visual scree test and the results often are better than those from alternative analytic techniques such as the CNG method. Further use of the multiple regression method will identify the strengths and limitations of this approach.

References

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Psychology, 3,* 77-85.

Bartlett, M. S. (1951). A further note on tests of significance in factor analysis. *British Journal of Psychology, 4,* 1-2.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1,* 245-276.

Cattell, R. B. (1978). *The scientific use of factor analysis in behavior and life sciences.* New York: Plenum.

Cattell, R. B. & Jaspers, J. (1967). A general plasmode No. 30-10-5-2 for factor analytic exercises and research. *Multivariate Behavioral Research Monographs, 2,* 1-212.

Cattell, R. B. & Vogelman, S. (1977). A comprehensive trial of the scree and KG criteria for determining the number of factors. *Multivariate Behavioral Research, 12,* 289-325.

Cliff, N. (1970). The relation between sample and population characteristic vectors. *Psychometrika, 35,* 163-178.

Crawford, C. B. (1975). Determining the number of interpretable factors. *Psychological Bulletin, 82,* 226-237.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Gorsuch, R. L. & Nelson, J. (1981). *CNG scree test: An objective procedure for determining the number of factors.* Paper presented at the annual meeting of the Society for Multivariate Experimental Psychology.

Hakstian, A. R., Rogers, W. T., & Cattell, R. B. (1982). The behavior of number-of-factors rules with simulated data. *Multivariate Behavioral Research, 17,* 193-219.

Hakstian, A. R. & Muller, V. J. (1973). *Users manual to accompany the Alberta General Factor Analysis Program,* Edmonton, Alberta: Division of Educational Research Services, University of Alberta.

Horn, J. L. & Engstrom, R. (1973). On the subjective character of the empirical base of the structure-of-intellect model. *Psychological Bulletin, 80,* 33-43.

Howell, D. C. (1987). *Statistical methods for psychology.* Boston: Duxbury Press.

Kaiser, H. F. (1970). A second generation Little Jiffy. *Psychometrika, 35,* 401-415.

Kaiser, H. F. & Caffrey, J. (1965). Alpha factor analysis. *Psychometrika, 30,* 1-14.

Kano, Y. (1990). Noniterative estimation and the choice of the number of factors in exploratory factor analysis. *Psychometrika, 55,* 277-291.

Lambert, Z. V., Wildt, A. R., & Durand, R. M. (1990). Assessing sampling variation relative to the number of factors criteria. *Educational and Psychological Measurement, 50,* 33-48.

Tucker, L. R., Koopman, R. F., & Linn, R. L. (1969) Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika, 34,* 421-459.

Zoski, K. & Jurs, S. (1990). Priority determination in surveys: An application of the scree test. *Evaluation Review, 14,* 214-219.

Multiple Linear Regression Techniques in Assessing the Effects of Truncation on the Gamma Distribution

Gannat Ramadan Mohamed El-Daian
Al-Azhar University for Girls, Egypt

Samuel R. Houston
University of Northern Colorado

Robert L. Heiny
University of Northern Colorado

This study examines, through Monte Carlo simulation, the effects of truncation on the gamma distribution (with one through four parameters). Specifically, three types of truncation (right, left, and double truncation) are considered. Computer facilities were used to generate 400 random samples from the gamma distribution with different parameter values for different sample sizes, shape, scale, location, power, and degree of truncation. Correlation and regression analysis demonstrated that the degree of left truncation has a significant correlation with the measures of central tendency for all distributions. The degree of right truncation had a significant relationship with the measures of deviation. On the other hand, the kind of truncation had a significant unique contribution for all models. However, the type of truncation had significant unique contributions for all models except two (scale and location models for gamma with three parameters).

The gamma distribution, or what may be called Pearson Type III of frequency curves, is one of the most important statistical distributions. It has been studied and investigated by many writers because of its application in different areas such as industrial engineering, physics, and quality control. For example, the gamma distribution can be considered as a description of duration variables such as the time taken for an instrument to be repaired, the time taken to get served at a store, etc.

If x is a continuous random variable having the gamma distribution, the generalized gamma density function is given by:

$$f(x; c, \beta, \lambda, \alpha) = \left(\beta\lambda^{-\alpha\beta}/\Gamma(\alpha)\right)(x-c)^{\alpha\beta-1}\exp\left(-((x-c)/\lambda)\beta\right) \qquad [1]$$

where $0 \le c < x < \infty$, $\beta > 0$, $\lambda > 0$, $\alpha > 0$, c is the location parameter, λ is the scale parameter, β is the power parameter, and α is the shape parameter.

The gamma distribution may have one, two, three, or four parameters. The probability density function for a gamma distribution with three parameters may be obtained from the generalized form (gamma with four parameters) by letting $\beta = 1.0$. The probability density function with two parameters can be obtained by using $c = 0.0$ and $\beta = 1.0$. The probability density function with one parameter can be obtained by letting $c = 0.0$ and $\beta = \lambda = 1.0$ in the generalized form.

In some situations, the complete range of the gamma distribution is not available to the researcher, in which case he or she works with a truncated distribution. The general form of the probability density function in this case becomes:

$$f(x|a < x < b) = \frac{f(x)}{\int_a^b f(x)dx} \qquad [2]$$

where $a < x < b$ and where the values of a and b are dependent on the type of truncation and truncation degree. Therefore, if the range of x in (2) is $[a,\infty)$, the distribution can be called a left truncated gamma distribution. However, if the range of x is (0,b], the distribution can be called a right truncated gamma distribution.

Many writers have discussed the subject of truncated distributions. Most of these studies are devoted to the exponential and normal distribution (Basu, 1964a; Yang & Sirvanci, 1977; Megahed, 1981; Depriest, 1983; Mittal, 1984; Mittal & Dahiya, 1987). Some writers have discussed the subject of the truncated gamma distribution. Most of these studies are devoted to the estimation of the parameters by the method of moments or by the method of maximum likelihood. Cohen (1950) used the method of moments while Des Raj (1953), Broeder (1955), and Chapman (1956) used the method of maximum likelihood.

In this paper, three kinds of truncation were considered: (a) degree of truncation is defined as size or degree of right, left, and double truncation ((tr), (tl) and (tl,tr)), (b) type of truncation is binary coded, where right truncation is (0,1), left truncation is (1,0) and

double truncation is (1,1), and (c) interaction is coded tl*tr(0,1) and tr*tl(1,0).

Truncation can be one-sided (right or left) or two-sided. The three main types of truncation are right truncation, left truncation, or double truncation. In each case the fraction of the distribution of the population that falls outside the truncation point or points is called the degree of truncation. In this study two different approaches were used to analyze the effects of truncation on the gamma distribution. First, correlation analysis was used to determine the effect of truncation (type and degree) on the measures of central tendency and/or variation for the gamma distributions. Second, the multiple linear regression approach was used to generate and analyze models to isolate unique contributions of the different kinds of truncation.

Briefly, the major purpose in using multiple linear regression was to determine the actual impact of truncation (type, degree, etc.) on the gamma distribution of one, two, three and four parameters. The measures of central tendency are changed as a result of the type of truncation and/or its degree. The regression technique assisted in defining the theoretical consequences of truncation in the gamma distribution.

Research Limitations

This study was carried out using empirical data generated by a Monte Carlo simulation. The data were used to investigate the effects of different kinds of truncation on the gamma distribution. Computer facilities at the University of Northern Colorado were used to generate 400 random samples from the gamma distribution with different parameter values. The generated sample sizes were 10, 20, 30, 50, and 100. The shape parameter (α = 0.5, 1.5, 2.0, 3.0, 5.0 and 10.0), the scale parameter (λ = 1.0, 1.5, 2.0, 3.0, 4.0, and 5.0), the location parameter (c = 0.0, 0.1, 0.2, 0.3, 0.5, 1.0, 2.0, 3.0, and 5.0), and the power parameter (β = 1.0, 2.0, 3.0, 4.0, and 5.0) were considered. The truncation degrees were t times α, where t = 0.1, 0.2, 0.5, 1.0, 2.0, and 3.0.

Generation Of Gamma Variables

Since the gamma forms are easily obtained from raw data, they can approximate a wide variety of functional shapes. It could play a major role in digital simulation studies. Various investigators have been concerned with generating gamma variables. Phillips and Beightler (1972) presented a technique for generating random gamma varieties depending on two parameters:

$$f(x) = (\lambda^{-\alpha}/\Gamma(\alpha))x^{\alpha-1}e^{-x/\lambda}, \quad 0 < x < \infty$$
$$f(x) = 0 \quad \text{elsewhere} . \qquad [3]$$

They made a comparison between the composition technique and the rejection technique. With regard to statistical goodness-of-fit based on limited experiments, the first three methods were capable of generating random gamma variables closely approximating the desired gamma density for values of $\alpha > 1$. For lower values of the scale parameter λ, their method was better than the others. With regard to computer generation times, the rejection method is recommended for values of $\alpha < 2.5$. Naylor's (1971) method did quite well for values of $\alpha > 1$; although it was an easy method to program, it did require more computer running time.

Whittaker (1974) introduced a method of generating random variables from uniform variables with a gamma or beta distribution having a non-integer shape parameter. Cheng (1977) described the rejection method for generating gamma varieties with shape parameter α where $\alpha > 1$. The scheme used to generate gamma varieties with one parameter is described below.

1. Set $a = (2\alpha - 1)^{-1/2}$, $b = \alpha - \log 4$ and $c' = \alpha + \alpha^{-1}$.
2. Generate a pair of uniform random numbers, U_1 and U_2.
3. Set $v = a\log(U_1/(1-U_1))$, $x = \alpha e^{v}$.
4. If $b + c'v - x \geq \log(U_1^2 U_2)$, accept x, otherwise go to Step 2.

This method was better than the previously published method (Phillips & Beightler, 1972; Whittaker, 1974) in terms of speed and program compactness.

Cheng and Feast (1979) introduced a simpler and faster method for generating a gamma variate by using the rejection technique. This method is suitable for all $\alpha > 1$ and it uses the ratio of uniform variable. It was noticed that some studies for generating gamma variable depend on specified conditions. Other methods need more time because the expected number of trials needed for each accepted variate could be more complicated than anticipated, and the complexity of the calculations required per trial are more extensive than anticipated. This is usually determined largely by the number of uniform random numbers needed and by the number of logarithmic or exponential function evaluations required.

Ripley's (1983) algorithm was developed on the basis of a recent study to generate gamma variables. It can be executed by using a small computer. However, it was found that the results were not suitable for most values of α and the generated variables did not follow the gamma distribution when $\alpha > 3$. The Ripley algorithms are called GKM1, GKM2, and GKM3. Algorithm GKM1 is described for $1 < \alpha < 4$. Algorithm GKM2 applies for $\alpha > 4$. More importantly, they are composite algorithm, called GKM3, which remains correct while covering all $\alpha > 1$.

Main Steps of Algorithm GKM1

1. Set $a = \alpha - 1$, $b = (\alpha - (6\alpha)^{-1}) / a$, $c' = 2 / a$ and $d = c' + 2$.
2. Generate independent U(0,1) variates U_1 and U_2.
3. Let $w = bU_1/U_2$. If $(c'U_2 - d + w + w^{-1}) \leq 0$ go to 4.
4. If $c' \log U_2 - \log w + w^{-1} \geq 0$ go to step 1.
5. Deliver $x = aw$.

Main Steps of Algorithm GKM2

1. Let a, b, c', d as in GKMI; $f = \sqrt{\alpha}$.
2. Generate independent U (0,1) variable U_1 and U. Set $U_2 = U_1 + f^{-1}(1\text{-}1.86U)$. Repeat this step unless $0 < U_2 < 1$ (the constant f must be less than $\left(1+\sqrt{2}/e\right)$.

Steps 3, 4, and 5 are as in GKM1.

This method is complementary to GKM1 in that it is slower than GKMI for a near 1, but it rapidly becomes faster as a increases. Since GKMI and GKM2 differ only in step 2, it is easy to combine them by use of a switch. The composite algorithm is called GKM3.

Main Steps of Algorithm GKM3

1. This is exactly the same as GKM2 except that step 2 is replaced by 2'.

2'. Use step 2 of GKMI or GKM2 according to whether α is less than or greater than a prescribed value α_0. The suggested value for α_0 is 2.5; this ensures that the speed of variate generation is substantially the same for all $\alpha > 1$.

Three algorithms are needed for the three separate programs; therefore, it is preferred that the three algorithms be combined to make a modified algorithm which will be suitable for all values of α. In this way, a program is developed which can be generalized for generating gamma random variables with three or four parameters for any value of α, and value of λ, ß, or c. The modified algorithm is introduced in the next section.

Research Design And General Procedures

The Faculty of Commerce Computer facilities at Al-Azhar University for Girls were used first to apply the Monte Carlo simulation. The Pseudo-Random Number Subroutine from the alphatronic Microcomputer System (BH/01824/e/k) was used next. Finally, the IBM 3081 G32 Computer System at the University of Northern Colorado was used.

Different random samples were used for generating gamma variables and different gamma distributions were generated with shape parameter (α = 0.5, 1.5, 2.0, 3.0, 5.0, and 10.0), scale parameter (λ = 1.0, 1.5, 2.0, 3.0, 4.0, and 5.0), location parameter (c = 1.0, 0.2, 0.3, 0.5, 1.0, 3.0, and 5.0), and power parameter (ß = 1.0, 2.0, 3.0, 4.0, and 5.0). BASIC programs which were designed for generating gamma varieties were dependent on the following modified algorithm.

The researchers developed an algorithm through the mixture of GKMI, GKM2, and GKM3. The algorithm was developed to be suitable for all values of the parameters of the three parameter gamma distribution. Supposing that ($\alpha < 1$) is a shape parameter, the algorithm follows these steps:

1. Let $a = \alpha - 1$, $b = (\alpha - (6\alpha)^{-1}) / a$, $c' = 2 / a$, $d = c' + 2$ and $f = \sqrt{\alpha}$.
2. Generate independent U(0,1) variables U_1 and U_2.
3. If ($\alpha < 4$) go to step 5.
4. Set $U_2 = U_1 f^{-1}(1 - 1.86 U_0)$. Repeat this step unless $0 < U_2 < 1$.
5. Let $w = b U_1/U_2$. If $(c'U_2 - d+w+w^{-1}) \leq 0$, go to step 7.
6. If $(c'\log U_2 - \log w + w - 1) \geq 0$ go to step 2.
7. Deliver $x = (aw)\lambda + c$.

Description of the Variables

The variables used for the analyses in this study are listed in Table 1.

Table 1 List of Variables

Parameters		
1	Shape Parameter	Y1
2	Scale Parameter	Y2
3	Location Parameter	Y3
4	Power Parameter	Y4
Measures of Central Tendency and Deviation		
5	Mean	$\bar{x}$
6	Variance	s^2
7	Skewness	Sk
8	Kurtosis	Ku
9	Mode	Mo
10	Median	Me
Sample Size		
11	Sample Size	n
The Kind of Truncation		
12	Degree lt truncation	tl.
13	Degree rt truncation	tr.
14	Left truncation	tl(1,0)
15	Right truncation	tr(0,1)
16	Interaction 1	tl*tr(0,1)
17	Interaction 2	tr*tl(1,0)
The Kind of Distribution		
18	Gamma one parameter	k1
19	Gamma two parameters	k2
20	Gamma three parameters	k3
21	Gamma four parameters	k4

Table 2 Relationships Between Central Tendency and Deviation Measures and Kind of Truncation for Gamma Distributions (N=250)

Measure	tl	tr	tl(1,0)	tr(0,1)	tl*tr(0,1)	tr*tl(1,0)
Mean						
One	.977*	-.011	.238*	-.347*	-.023	-.027
	(.0001)	(.857)	(.0001)	(.0001)	(.718)	(.677)
Two	.727*	-.439*	.199*	-.524*	-.122	-.297*
	(.0001)	(.0001)	(.002)	(.0001)	(.059)	(.0001)
Three	.139*	.306*	-.114	-.194*	.113	.089
	(.028)	(.0001)	(.072)	(.002)	(.075)	(.089)
Four	.403*	.314*	.078	-.269*	-.059	.171*
	(.0001)	(.0001)	(.219)	(.0001)	(.352)	(.007)
Mode						
One	.980*	-.030	.239*	-.345*	-.030	-.030
	(.0001)	(.641)	(.0001)	(.0001)	(.636)	(.539)
Two	.756*	-.423*	.223*	-.503*	-.103	-.285*
	(.0001)	(.0001)	(.0003)	(.0001)	(.109)	(.0001)
Three	.185*	.309*	-.062	-.199*	.150*	.123
	(.003)	(.0001)	(.329)	(.002)	(.017)	(.052)
Four	.412*	.302*	.081	-.278*	-.062	.159*
	(.000)	(.0001)	(.202)	(.0001)	(.325)	(.012)
Median						
One	.977*	-.011	.857*	-.347*	-.022	-.027
	(.0001)	(.858)	(.0001)	(.0001)	(.718)	(.676)
Two	.728*	-.438*	.200*	-.524*	.122	-.297*
	(.0001)	(.0001)	(.002)	(.0001)	(.054)	(.0001)
Three	.140*	.308*	-.114	-.192*	.114	.090
	(.027)	(.0001)	(.073)	(.002)	(.072)	(.155)
Four	.403*	.314*	.078	.269*	-.059	.171*
	(.0001)	(.0001)	(.221)	(.0001)	(.352)	(.007)
Variance						
One	.492*	.533*	.111	-.218*	.174*	.319*
	(.0001)	(.0001)	(.081)	(.0005)	(.006)	(.0001)
Two	.389*	-.419*	-.009	-.475*	-.189*	-.294*
	(.0001)	(.0001)	(.892)	(.0001)	(.003)	(.0001)
Three	-.136*	.145*	-.297	-.142*	-.120*	-.145*
	(.0001)	(.0001)	(.892)	(.0001)	(.003)	(.0001)
Four	.020	.193*	-.028	-.058	-.030	.106
	(.752)	(.002)	(.660)	(.363)	(.638)	(.095)
Skewness						
One	-.139*	-.222*	-.063	.009	-.022	-.127*
	(.028)	(.0004)	(.323)	(.891)	(.734)	(.044)
Two	.045	-.183*	.115	-.167*	.146*	.013
	(.476)	(.004)	(.071)	(.008)	(.021)	(.636)
Three	-.077	-.537*	.301*	-.347*	-.135*	-.202*
	(.228)	(.0001)	(.0001)	(.0001)	(.034)	(.001)
Four	.245*	.081	.104	-.158*	.344*	.186*
	(.0001)	(.204)	(.102)	(.012)	(.0001)	(.003)
Kurtosis						
One	-.138	-.200*	-.202*	.148*	-.038	.151*
	(.300)	(.002)	(.001)	(.020)	(.550)	(.017)
Two	-.015	.139*	.144*	.160*	.319*	.222*
	(.816)	(.028)	(.023)	(.011)	(.0001)	(.0004)
Three	-.054	.229*	.347*	.004	-.064	-.077
	(.397)	(.0003)	(.0001)	(.947)	(.315)	(.228)
Four	.227*	.213*	.052	-.041	.343*	.205*
	(.0003)	(.0007)	(.410)	(.524)	(.0001)	(.001)

Note. One, Two, Three, and Four are gamma distributions with one through four parameters; top numbers refer to Pearson correlation coefficients; numbers in parentheses refer to p-values; asterisks indicate a significant relationship ($p < .05$).

Correlation Analysis

Table 2 shows the correlation coefficients between the central tendency measures and the kind of truncation for the gamma distribution with one through four parameters. This table indicates that the degree of left truncation had the highest correlation with the measures of central tendency of the gamma distribution with one through four parameters.

The degree of right truncation had negative correlations with the measures of central tendency of the gamma distribution with one and two parameters. The degree of right truncation has a significant relationship with the measures of central tendency for gamma distributions except the gamma with one parameter.

Table 2 indicates the relationship between the type of truncation and the measures of central tendency. The relationships between the left truncation and the measures of central tendency were positive and significant for the gamma with one and two parameters. The table also indicates that right truncation had negative and significant correlations with measures of central tendency for gamma distributions with one through four parameters. Further, there was no evidence to support that the interaction tl*tr(0,1) had a significant correlation with the measures of central tendency for the gamma distributions except the mode of the three parameter gamma distribution. However, the interaction tr*tl(1,0) had significant correlations with measures of central tendency for the gamma with one and four parameters.

Also, Table 2 indicates that all types of truncation except the left truncation had a significant correlation with the variance of the gamma distributions with one through three parameters. The degree of left truncation has a significant correlation with the skewness of the gamma with one and four parameters, but there was no evidence to support that there were relationships with the skewness of the gamma with two and three parameters. The degree of right truncation had a significant correlation with the skewness of the gamma with one through three parameters.

Left truncation was not significantly correlated with the skewness of the gamma with one, two, and four parameters; however, it has a significant correlation with the skewness of the gamma with three parameters. The right truncation has a significant correlation with skewness in the different cases, except for the gamma with one parameter.

Degree of left truncation has significant correlations only with the kurtosis of the gamma with four parameters; however, the degree of right truncation has significant correlation with kurtosis of the gamma with one through four parameters. Left truncation has significant correlations with the kurtosis of the gamma with one and three parameters; however, right truncation has significant correlations with the gamma with one and two parameters. The interactions had significant correlation with the gamma with one and four parameters.

Table 3 presents the relationships between measures of central tendency, dispersion, distributional shape measures, and the kind of truncation. This table is based on all samples which are used for one through four parameter distributions, and is a summary of these relationships. It indicates that the relationships between the measures of central tendency and both the degree and type of truncation were significant.

There was no evidence that the degree of truncation was significantly related to the variance but the type of truncation was significantly related to the variance. The degree and the type of right truncation had a significant correlation with the measures of distributional shape, while there was no evidence to support that left truncation was significantly related to the measures of distributional shape.

Table 3 Relationships Between Distribution Measures and Kind of Truncation (N=1000)

Measure	tl	tr	tl(1,0)	tr(0,1)	tl*tr(0,1)	tr*tl(1,0)
Mean	.655	.081	.066	-.331	.032	-.009
	(.0001)	(.010)	(.036)	(.0001)	(.306)	(.774)
Mode	.697	.076	.105	-.329	.039	.001
	(.0001)	(.016)	(.0009)	(.0001)	(.217)	(.968)
Median	.655	.082	.067	-.330	.033	-.009
	(.0001)	(.010)	(.035)	(.0001)	(.301)	(.788)
Variance	.049	.034	-.190	-.183	-.064	-.111
	(.119)	(.277)	(.0001)	(.0001)	(.045)	(.0004)
Skewness	.005	-.267	.060	-.167	-.010	-.118
	(.870)	(.0001)	(.057)	(.0001)	(.762)	(.0002)
Kurtosis	-.044	-.135	.036	.072	.044	.071
	(.162)	(.0001)	(.254)	(.022)	(.153)	(.026)

Note. Top numbers refer to Pearson correlation coefficients; numbers in parentheses are p-values.

Table 4 Relationships Between Kind of Truncation and Number of Parameters as Coded by Binary Vectors (N=1000)

	Number of Parameters of Gamma Distribution			
Kind of Truncation	One	Two	Three	Four
tl	.088	.036	-.042	-.082
	(.005)	(.256)	(.181)	(.010)
tr	-.019	-.149	.088	.080
	(.547)	(.0001)	(.005)	(.011)
tl(1,0)	.077	-.030	-.169	.124
	(.014)	(.336)	(.0001)	(.0001)
tr(0,1)	.004	-.095	.047	.045
	(.912)	(.003)	(.140)	(.159)
ti*tr(0,1)	-.061	-.059	.152	-.033
	(.055)	(.060)	(.0001)	(.300)
tr*tl(1,0)	-.011	.094	-.026	.131
	(.731)	(.003)	(.415)	(.0001)

Note. Top numbers are Pearson correlation coefficients; numbers in parentheses are p-values.

Table 4 identifies the relationships between the kinds of truncation and the distributions. This table indicates that the gamma distribution with one parameter had significant relationships with the degree and type of left truncation. The gamma distribution with two parameters had significant relationships with the degree and type of right truncation and the interaction tr*tl(1,0). The gamma distribution with three parameters had significant relationships with the degree of right truncation, left truncation and the interaction (both left and right), left truncation and the interaction tr*tl(1,0).

Briefly, the relationships between truncation and the measures of gamma distributions were dependent on the kind of distribution, the type of truncation, and the distribution measures; whereas the relationships among the measures of central of tendency and the types or degrees of truncation were all significant.

Multiple Regression Analysis

Multiple regression analyses were performed to examine the relationships between each parameter of the gamma distribution (dependent variable) and the set of descriptive characteristics (independent variables) of the gamma distributions. Determination of which variable serves as criterion variable and which set of variables are predictor variables was dependent on the model that was considered.

In general, the set of descriptive characteristics of the gamma distributions includes the variables,

1. Measures of central tendency (mean, mode, and median) to be represented by X_1, X_5, and X_6, respectively.
2. Measures of deviation (variance, skewness, and kurtosis) represented by X_2, X_3, and X_4, respectively.
3. Sample size represented by X_7.
4. Degree of truncation (tr) and (tl), represented as X_8 and X_9, respectively.
5. Type of truncation tl(1,0) and tr(0,1), represented as X_{10} and X_{11}, respectively.
6. Interaction tl*tr(0,1) and tr*tl(1,0), represented as X_{12} and X_{13}, respectively.
7. Kind of gamma distribution with one through four parameters represented as X_{14}, X_{15}, X_{16}, and X_{17}, respectively. Each variable was binary coded.

The full model considers all seventeen variables (X_1, X_2, ... X_{17}) as predictors for the dependent variables (one of the four gamma distribution parameters).

Sample size should be related to the number of variables and should increase as the number of variables increases (Barcikowski & Stevens, 1975). One informal guide for a lower limit is that there should be 10 subjects for each variable. To insure sufficient sample size for a small set of variables, Thorndike (1978) offered the following rule to determine the sample size: $N \geq 10(p+c)+50$, where p is the number of independent variables (predictors) and c is the number of dependent variables (criteria). In this study, Thorndike's sample size requirement was satisfied in all regression models which were analyzed.

A hierarchical chart of reduced models was constructed (see Table 5). Each model was reduced from the full model and significant drops in R^2 indicated significant contributions to the dependent variable, Y, by omitted predictors are identified by asterisks using the .05 level of significance (see Table 6).

Also, comparisons between R^2 for the different models are summarized in Table 6 for the gamma with one through four parameters and the general model (which includes all the data of the gamma distribution with one through four parameters) for the shape of the parameter model. This table indicates that a sequential decrease in R^2 occurred from gamma with one through gamma with four parameters for the shape models (.952 to .733). R-squares for full models depended on the kind of distribution and N. Furthermore, the R^2 value

in the general case for the four parameter model dropped to .733 (N=1000) from .868 (N=250). Thus, it can be seen that both the sample size and number of parameters were factors which influence the R^2 value.

Table 5 List of Regression Analysis Reduced Models

#	The Reduced Model
1	Full model - measures of central tendency (MCT)
2	Full model - measures of deviation (M.Dev)
3	Full model - sample size (n)
4	Full model - kinds of truncation
5	Full model - mean
6	Full model - mode (mo)
7	Full model - median (me)
9	Full model - distributional shape measure (ds)
10	Full model - skewness (sk)
11	Full model - kurtosis (ku)
12	Full model - (tr,tl)
13	Full model - (tl(1,0), tr(0,1)
14	Full model - interactions (int)
15	Full model - tl.
16	Full model - tr
17	Full model - tl (1,0)
15	Full model - tr (0,1)
19	Full model - kinds of distribution (for the general model)

Table 6 R-Square Values for Full and Reduced Models in Regression Analysis for Gamma Distribution Where Y is the Shape Parameter

	Number of Parameters of Gamma Distribution				
Model	One	Two	Three	Four	General
FM	.952	.948	.886	.868	.733
FM-MCT	.948*	.511*	.864*	.576*	.580*
-Mean	.952	.948	.885	.868	.732
-Mo	.952	.948	.879*	.865*	.732
-Me	.952	.948	.885	.868	.732
FM-Dev.	.946*	.810*	.865*	.860*	.714*
-Var.	.952	.812*	.873*	.861*	.717*
FM-Ms.	.947*	.942*	.884	.867	.731
-Sk	.949*	.942*	.885	.867	.733
-Ku	.948*	.943*	.884*	.868	.733
FM-n	.952	.948	.882	.867	.732
FM-Truncation	.897*	.498*	.809*	.626*	.519*
-degrees	.947*	.645*	.857*	.693*	.613*
-tl.	.948*	.663*	.886	.724*	.667*
-tr.	.951*	.948	.858*	.862*	.705*
-types	.947*	.931*	.872*	.854*	.722*
-tl(1,0)	.951*	.948	.881*	.867	.722*
-tr(0,1)	.949*	.931*	.879*	.855*	.732
-interaction	.924*	.848	.885	.818*	.709*

Note: Asterisk indicates that the unique contribution for these variables was significant at .05 level.

Table 7 Significant Unique Contributions for each Kind of Truncation in Regression Analysis Models Kind of Truncation

Y	tl	tr	tl(1,0)	tr(0,1)	Interaction
Shape Parameters					
One	*	*	*	*	*
Two	*			*	*
Three		*	*	*	
Four	*	*		*	*
General	*	*	*		*
Scale Parameters					
Two	*	*	*	*	*
Three		*			
Four	*		*	*	*
General	*	*			
Location Parameters					
Three		*			
Four	*			*	*
General	*	*		*	*
Power Parameters					
Four			*	*	
General	*				

Note. Asterisk indicates that the unique contribution for these variables was significant at the .05 level.

Moreover, Table 6 indicates that the measures of central tendency, deviation, and the kinds of truncation (and especially the types of truncation) had significant unique contributions. Most kinds of truncation had significant unique contributions for all models. The variance had a significant, unique contribution for all distributions except for the gamma with one parameter. Moreover, skewness and kurtosis had a significant unique contribution for the gamma with one and two parameters.

Briefly, Table 7 shows the significant unique contributions of each kind of truncation in each model. It indicates that the degree of left truncation, had a significant unique contribution with all models except the models of the gamma with three parameters and the model of the power parameter in the gamma with four parameters. The degree of right truncation had significant unique contributions for all models except for the shape model for the gamma with two parameters, for location model for gamma with four parameters, and for the power models.

Left truncation had significant unique contributions with the shape models, except the general model, in the scale model for the gamma with two and three parameters, and in the power model for the gamma with four parameters. Right truncation had significant unique contributions with the shape models, except the general model, with the scale model for the gamma with two and four parameters model, with the location model for the gamma with four parameters, with the general gamma model, and with the power model for the gamma with four parameters.

The last kind of truncation (interaction) had significant unique contributions with the shape model for the gamma with two and four parameters, with the scale model for the gamma with two and four parameters, and with the location model for the gamma with four parameters.

Moreover, the types of truncation were considered as elements in the best proper subset for the different models. According to the backward elimination procedure for a dependent variable, Y, for each model, the best proper subset of the predictors for each model is given in Table 8.

In general, Table 8 indicates that it is possible that the best proper subset (using the backward elimination procedure) for the shape parameter included some type of truncation, for example, tl, tr, tl(1,0) and tl*tr(0,1) for the general model. For the shape parameter model of the gamma with four parameters, these same types of truncation patterns were included along with tr*tl(1,0). The best proper subset of independent variables for the scale parameter included tl, tr, and tl*tr(0,1) for the general model. For the shape parameter model of the gamma with four parameters, truncation variables including tl, tl(1,0), tr(0,1), and tl*tr(0,1) were identified as important. The best proper subset for the location parameter included tl, tr, tr(0,1), and tl*tr(0,1) for the general model. For the shape parameter model of the gamma with four parameters, it was found that tl, tr, tr(0,1), tl*tr(0,1), and tr*tl(1,0) were significant.

General Concluding Remarks

This study dealt with the effects of truncation on the family of gamma distributions. The goal was to determine whether the kind of truncation [tl, tr, tl(1,0), tr(0,1), tl*tr(0,1), and tr*tl(1,0)] had an influence on the

Table 8 The Best Proper Subset for Each Model in the Different Cases of Gamma Distributions

Model	General Model				Gamma with Four Parameters			
Subset/Y	Shape	Scale	Location	Power	Shape	Scale	Location	Power
1	mean	mean	Var	Var	Var	Var	mean	Sk
2	Var	Var	Sk	Sk	Mo	Mo	tl	Mo
3	tl	Sk	tl	Ku	tl	tl	tr	Me
4	tr	Ku	tr	Me	tr	tl(1,0)	tr(0,1)	
5	tl(1,0)	Me	tr(0,1)	tl	tr(0,1)	tr(0,1)	tl*tr(0,1)	
6	tl*tr(0,1)	tl	tl*tr(0,1)	tl*tr(0,1)	tl*tr(0,1)	tr*tl(1,0)	tr*tl(1,0)	
7	k1	tr	k1	k1	tr*tl(1,0)			
8	k2	tl*tr(0,1)		k2	k2			
9	k3	k1	k3	k3				
10		k2						
11		k3						
R^2	.729	.597	.875	.778	.865	.722	.845	.301

Model	Gamma with Three Parameters			Gamma with Two Parameters		Gamma with One Parameter
Subset/Y	Shape	Scale	Location	Shape	Scale	Shape
1	mean	var	mean	var	mean	mean
2	var	tl	var	sk	var	sk
3	mode	tl(1,0)	tl(1,0)	ku	ku	ku
4	n			Me	Me	tr
5	tr			tl	tl	tr(1,0)
6	tl(1,0)			tr	tr	tl*tr(0,1)
7	tr(0,1)			tr(0,1)	tr(0,1)	tr*tl(1,0)
8				tl*tr(0,1)	tl(1,0)	
R^2	.882	.589	.500	.948	.696	.948

parameters, central tendency, dispersion, and distributional shape measures of the gamma distributions.

From these analyses, it can be concluded that the truncation significantly affected the gamma distribution (its measures and its parameters). The effect was dependent on the kind of truncation (type and degree), the type of distribution, and the values of parameters.

The numerical analyses in this research presented abstract concepts about irregular relationships, but also introduced some details about the relationships between each kind of truncation and the parameters, and each kind of truncation and the most characteristic measures of gamma distributions. For example, the degree of left truncation had a significant correlation with the measures of central tendency for all kinds of distributions. On the other hand, the degree of right truncation had a significant relationship with the measures of central tendency and deviation for most kinds of gamma distributions.

In terms of multiple regression analysis, for all models, the kind of truncation had significant unique contributions. The degree of truncation made significant, unique contributions to all models, but the types of truncation had a significant unique contribution for all models except two (scale and location models for the gamma with three parameters). However, the unique contribution for each kind of truncation was dependent on the kind of truncation, the distribution, and the measure being considered.

The overall findings of this study were generally supportive of the findings of the reviewed research. Although this study was limited to the effects of truncation on the gamma distribution, it can be concluded that the findings apply also for Erlang, exponential, and chi-square distributions.

Suggestion for Further Studies

The focus of this study emphasizes the effects of truncation on the gamma distribution. Based upon the findings and conclusions described above, these recommendations are made for further investigation:

1. Results were dependent on specific values of the parameters. Therefore, for more generalizable results, different values of the parameters may be utilized in subsequent research.
2. Exploration should be continued in an effort to study the effect of truncation on other distributions such as beta, lognormal, Weibull, etc. Additional study could determine if the research findings for

different distributions are comparable for similar types of truncation. This could provide the researcher with information about the effects of truncation on different kinds of distributions. Also, research should be extended to include the study of effects of truncation on mixture distributions.

3. Moreover, the effects of truncation in both estimation and hypothesis testing when using the transformations mentioned by Mohamed (1981) should be examined.
4. The methods used in this study should be extended to examine the effects of inner truncation and partial truncation on the gamma distribution.
5. Other studies should be designed using other techniques such as canonical correlation or factor analysis to determine whether rotated and/or unrotated factor solutions are affected by the type and/or degree of truncation.

References

Barcikowski, R. S. & Stevens, J. P. (1975). A Monte-Carlo study of the stability of canonical correlation, canonical weights, and canonical variate-variable correlation. *Multivariate Behavioral Research, 10,* 353-364.

Basu, A. B. (1964a). Effect of truncation on a test for the scale parameter of the exponential distribution. *Annals of Mathmatical Statistics, 35,* 209.

Broeder, G. G. (1955). On parameter estimation for truncated pearson type III distribution. *Annals of Mathematical Statistics, 26,* 659-663.

Chapman, D. G. (1956). Estimating the parameters of truncated gamma distribution. *Annals of Mathematical Statistics, 27,* 498-506.

Cheng, R. C. H. (1977). The generation of gamma variables with non-integral shape parameter. *Applied Statistics, 26,* 71-75.

Cheng, R. C. H. & Feast, G. M. (1979). Some simple gamma variable generators. *Applied Statistics, 28,* 290-295.

Cohen, A. C. (1950). Estimating parameters of Pearson type III populations from truncated samples. *Journal of the American Statistical Association, 45,* 411-423.

Depriest, D. J. (1983). Using the singly truncated normal distribution to analyze satellite data. *Communication in Statistical Theory and Methods, 12,* 263-272.

Des Raj (1953). Estimation of the parameters of type III populations from truncated samples. *Journal of the American Statistical Association, 48,* 336-349.

Megahed, Abdel-Razek (1981). *A Monte-Carlo investigation of the normality assumption with truncated normal population.* Unpublished doctoral dissertation, University of Northern Colorado, Greeley, CO.

Mittal, M. M. (1984). *Estimating the parameters of truncated distributions.* Unpublished doctoral dissertation, Old Dominion University, Norfolk, VA.

Mittal, M. M. & Dahiya, R. C. (1987). Estimating the parameters of a doubly truncated normal distribution. *Communication in Statistics/ Simulations, 16,* 141-149.

Mohamed, G. R. (1981). *Statistical implications of using a Gaussianting transformation.* Unpublished master's thesis. Al-Azhar University for Girls, Egypt.

Naylor, T. (1971). *Computer simulation experiments with models of economic systems.* New York, Wiley.

Phillips, 0. T. & Beightler, C. S. (1972). Procedures for generating gamma varieties with non-integer parameter sets. *Journal of Statistical Computer Simulation, 1,* 197-208.

Ripley, B. D. (1983). Computer generation of random variable: A tutorial. *International Statistical Review, 51,* 301-319.

Thorndike, R. M. (1978). *Correlation procedures for research.* New York, Wiley.

Whittaker, J. (1974). Generating gamma and beta random variables with non-integral shape parameters. *Applied Statistics, 23,* 210.

Yang, G. & Sirvanci, M. (1977). Estimation of a time-truncated exponential parameter used in life testing. *Journal of the American Statistical Association, 72,* 444-447.

Some Nonstandard Applications of the Analysis of Covariance Model

John T. Pohlmann
Southern Illinois University
Carbondale

This paper illustrates two applications of the ANCOVA model under problematic conditions: Johnson-Neyman significance region analysis and the analysis of a regression discontinuity design. The differences between intact and randomized group designs are discussed in the ANCOVA context. The analyses are demonstrated using the SAS REG program.

The analysis of covariance (ANCOVA) is used when a dependent variable and an antecedent independent variable are measured in multiple groups. The antecedent variable, or covariate, measures a source of variation that is to be statistically controlled. In order to accommodate the covariate, the ANCOVA model posits structural assumptions about the relationship between the dependent variable and the covariate. The simplest ANCOVA model assumes a homogeneous linear relationship between the dependent variable and covariate in each of the design groups. When the structural assumptions of ANCOVA are met, the ANCOVA can yield more powerful tests of significance than the analysis of variance (ANOVA). The power advantage of ANCOVA derives from a reduction of the error variance due to the effects of the covariate.

A number of authors (Elashoff, 1969; Lord, 1969; Mueller, 1990) have cautioned that the ANCOVA, performed with intact groups, does not control for preexisting group differences with the same level of rigor as does a randomized design. These authors argue that statistical controls cannot be considered as equivalent substitutes for randomization. Randomization will equate design populations for differences on the covariate along with any other differences that might exist prior to the experiment. Whereas, statistical adjustments can only be applied to variables that have been measured and only over the ranges of the variables observed in the sample. Statistical controls are also highly dependent on the model's structural assumptions. One can only statistically control for the relationships allowed by the model.

The ANCOVA, however, continues to be used with intact groups because of its convenience. Statistical adjustment is often the only control mechanism available to a researcher, and the ANCOVA may be the best statistical treatment of the data. The ANCOVA can be used under the same conditions that would justify the use of a partial correlation coefficient. The major difference between the ANCOVA and a partial correlation analysis is that the ANCOVA model is used when the independent variable is categorical.

This paper will treat the ANCOVA as a multiple group regression model solution. It is assumed that a researcher has measured a dependent variable and a covariate in each of J groups. The dependent variable is regressed separately on the covariate for each of the J groups. The structural assumptions of linearity and homogeneity of regression are tested. Then, tests of hypotheses about expected values of the dependent variable are demonstrated. The tests illustrated in this paper will not be used to test for group differences generally, but will instead examine group differences on the dependent variable at specific point values of the covariate. A regression discontinuity design and Johnson-Neyman significance region analysis will be used to illustrate this approach. These applications were chosen because they present alternatives that a researcher may use when structural or design problems are encountered.

Linear Models for the ANCOVA

The simplest form of the ANCOVA assumes a homogeneous linear relationship between the dependent variable (Y) and the covariate (X) for each of J groups. The linear model for the simple ANCOVA is given by Winer (1971, p. 757) as:

$$Y_{ij} = \mu + \alpha_j + \beta_w\left(X_{ij} - \mu_x\right) + \varepsilon_{ij} \qquad [1]$$

Y_{ij} and X_{ij} are the measures on the dependent and concommitant variables for case i in group j. The expected values of Y and X are denoted by μ and μ_x respectively. β_w is the within groups' regression coefficient, and is assumed to be the common slope of the regressions of Y on X for all groups. Group j's deviation on

Y is represented by α_j, and is called the treatment effect for group j. The only random term in the model is ε_{ij}, the error term, and is assumed to be NID(0,σ^2). All other terms in the model are fixed. Essentially, model [1] fits J parallel regression lines, predicting Y from X for the J groups in the design.

For the purposes of this paper, model [1] will be reparameterized into the following form:

$$Y_{ij} = \alpha_j + \beta_w X_{ij} + \varepsilon_{ij} \qquad [2]$$

Model [2] expresses α_j as the Y-intercept and β_w as the common slope for all groups. Model [2] permits the regression lines to have different intercepts, but only one slope; the regression lines are assumed to be parallel. If heterogeneity of regression is detected, model [2] can be revised to reflect nonparallel regression lines by replacing β_w with β_j. Each group is then allowed to have its own slope parameter in addition to a unique intercept parameter. Rewriting [2] accordingly, yields:

$$Y_{ij} = \alpha_j + \beta_j X_{ij} + \varepsilon_{ij} \qquad [3]$$

In models [1] and [2] group differences on Y could be measured by differences in the α_j values. Since the regression lines were assumed to be parallel, the differences between the α_j could be generalized over the full range of the covariate. In model [3], however, the difference between any two groups on Y depends on the value of X. Specifically, when X = C, the difference between the expected Y values of groups k and l is developed as follows:

$$E(Y_k|X=C) = \alpha_k + \beta_k C,$$

$$E(Y_l|X=C) = \alpha_l + \beta_l C,$$

$$E(Y_k|X=C) - E(Y_l|X=C)$$

$$= (\alpha_k + \beta_k C) - (\alpha_l + \beta_l C) \qquad [4]$$

If one hypothesized that the E(Y) values were equal at C, this condition could be expressed as a statistical hypothesis:

$$H_o: (\alpha_k + \beta_k C) = (\alpha_l + \beta_l C) \qquad [5]$$

In regression parlance, expression [4] is the difference between the predicted Y values for populations k and l at the value C on X. Expression [5] will serve as the null hypothesis for the tests described in this paper.

Applications Of Model [3]

Two applications of model [3] will be presented in this paper: a regression discontinuity analysis (Campbell and Stanley, 1963, p. 61) and Johnson-Neyman significance regions (Pedhazur, 1982, p. 469-472). These applications are interesting since they both represent tests that are performed under what is traditionally thought to be an undesirable situation. The regression discontinuity design represents an extreme case of group differences on the covariate. Group differences on the covariate can confound treatment effects. The Johnson-Neyman technique applies when heterogeneous regressions are observed. Heterogeneous regressions preclude a straightforward analysis of the α_j values, since, as per model [3], the α_j values only assess differences on Y when X=0.

The regression discontinuity design is used to test for effects on a dependent variable when individuals are treated differently, depending on the value of the covariate. The Campbell and Stanley (1968) illustration presents a quasi-experimental design for determining if a scholarship award, given on the basis of performance on a selection test, positively influences academic achievement. The covariate is the selection test and the two design groups are students who received the award and students who did not receive an award. Achievement is regressed on the selection test separately for each group. Then, the difference between the predicted values at the award cut-off is tested to assess the effect of the award. A test of the hypothesis in formula [5] could be used to perform this analysis.

Figure 1 shows a situation where the regressions of achievement on the selection test are homogeneous. The diagonal lines in the figure represent the separate regression lines for the award and no award groups. The regression lines are represented in Figure 1 as being parallel. The use of model [3], however, does not require homogeneity of regression. Each group's regression line could have any equation, and hypothesis [5] is still testable.

The Johnson-Neyman significance region technique uses tests of hypotheses like expression [5] to define regions on the covariate where groups differ, or do not differ, significantly on Y. Figure 2 below illustrates a possible outcome of a Johnson-Neyman significant region analysis. Testing differences between predicted values on Y for groups 1 and 2 might show that for values of $X < X_1$, group 2's predicted values are significantly higher on Y than those of group 1. Between X_1 and X_2 there is no significant difference between the groups' regression lines. Finally, for $X > X_2$ there is a significant difference favoring group 1. These regions can be defined by testing hypotheses like expression [5] for a full range of values on X and then noting which regions permit a significant interpretation. It might be necessary to iterate on X for reasonably accurate values of X_1 and X_2.

Figure 1 Illustration of the Regression Discontinuity Design

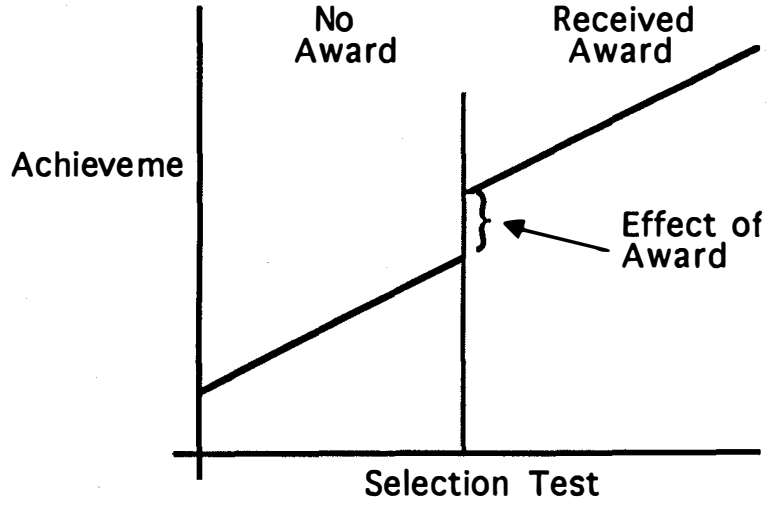
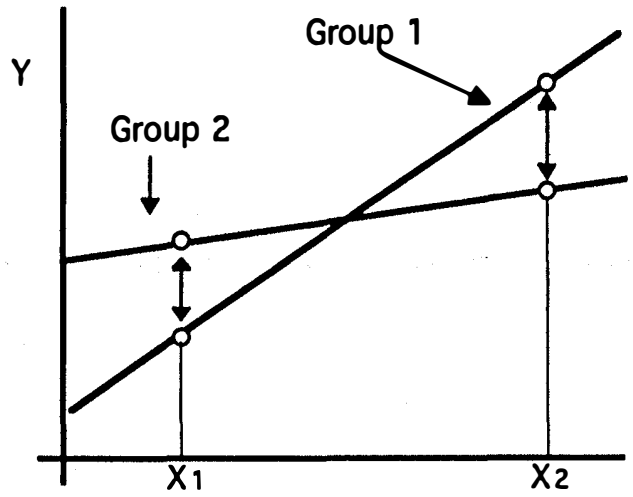


Figure 2 An Illustration of Johnson-Neyman Significance Regions



The two illustrations presented above have assumed a linear relationship between the covariate and the dependent variable. This assumption has been made to simplify the illustrations, not because it is an essential assumption for the ANCOVA. For any linear model application, the essential assumption is that the model be correctly specified.

Model [3] would be inappropriate if X and Y had a nonlinear relationship. Suppose the relationship between X and Y could be described by a quadratic model, then model [3] could be revised as follows:

$$Y_{ij} = \alpha_j + \beta_{j1}X_{ij} + \beta_{j2}X_{ij}^2 + \varepsilon_{ij} \qquad [6]$$

Model [6] allows each group to have its own quadratic regression line. For group j, the quadratic model is described by the regression constant α_j ; β_{j1}, the coefficient for X; and β_{j2}, the coefficient for X^2.

Using The SAS REG Program For Nonstandard ANCOVA Tests

The SAS REG program will be used to illustrate

the analyses. The SAS REG program is used for this purpose since it conveniently provides F-tests of any linear hypothesis for model parameters. PROC REG allows natural language expressions of linear hypotheses of the form H_0: Lß=c, where Lß=c represents a generalized system of linear combinations of the ß parameters (SAS Institute Inc., 1990, Chapter 6). Note that expression [5] is a linear combination of the α and ß parameters, and will serve as the basis for test requests in PROC REG.

Coding The Data For Analysis

Models [3] and [6] require that each of the J groups in the analysis have a separate model fitting the covariate to the dependent variable. Any computer model used for these analyses must posses this fundamental property. Now, there are an infinite number of ways to fit such models. For example, models [1] and [2] are isomorphic representations of the same structural model. In fact, any linear combination of the variables in these models will produce parameter estimates that will fit Y equally well. The approach shown here is not distinctive, in any important sense. Rather, its value lies in its simplicity.

Group Membership Coding

Binary variables will be used to code membership in the design groups. If a case is a member of group 1, then a binary variable, G1, will be coded 1 for that case. If a case is not a member of group 1, then that case's value on G1 will be 0. In a like manner each design group will be represented by a binary variable. For a J group design there will be J mutually exclusive and exhaustive binary variables; (G1, G2, ... , GJ)

Coding The Covariate

The covariate will be denoted as X in the following illustrations. In order to fit models like [3] and [6], X will be expanded to J variables. Cases in group 1 will have variable X1 coded with their value of X, while cases not in group 1 will have a 0 coded in X1. In a like manner X will be expressed as J variables (X1, X2, ... , XJ). In the case of model [6], X^2 will also be expanded the same way as X was expanded into (X1, X2, ... , XJ) in the preceding coding scheme. A case in group 1 will be coded in the variable XSQ1 with its value of X^2. A case that is not a member of group 1, will receive a code of 0 in XSQ1. X^2 will thereby be expanded to (XSQ1, XSQ2, ...,XSQJ) variables.

Table 1 below illustrates this coding scheme for a three group problem fit by model [3].

If the data set contains only Y, X and the GROUP variables given in Table 1, the following data transformations can be used to generate G1 through X3:

```
IF GROUP = 1 THEN G1 = 1; ELSE G1 = 0;
IF GROUP = 2 THEN G2 = 1; ELSE G2 = 0;
IF GROUP = 3 THEN G3 = 1; ELSE G3 = 0;
X1 = G1*X;
X2 = G2*X;
X3 = G3*X;
```

PROC REG Commands

The following commands illustrate how model (3) would be estimated with a data set like that in Table 1. The MODEL statement will use the NOINT option, which means that the program is not to estimate a common intercept for the entire sample. The coding of G1, G2 and G3 will permit separate intercepts to be estimated for each group.

```
PROC REG;
MODEL Y = G1 G2 G3 X1 X2 X3 / NOINT;
```

The parameter estimates for G1 to X3 in this SAS model statement are interpreted as follows:

G1 = Y intercept for group 1,
G2 = Y intercept for group 2,
G3 = Y intercept for group 3,
X1 = the slope for the regression line for group 1,
X2 = the slope for the regression line for group 2,
X3 = the slope for the regression line for group 3.

If one wanted to test the hypothesis that the expected values of Y when X = 50 in populations 1 and 2 were equal, the null hypothesis would become:

$$H_0: (\alpha_1 + \beta_1 50) = (\alpha_2 + \beta_2 50) \qquad [7]$$

Table 1 Coding for a Three Group ANCOVA Data Set in PROC REG

Y	X	GROUP	G1	G2	G3	X1	X2	X3
6	3	1	1	0	0	3	0	0
8	2	1	1	0	0	2	0	0
5	4	2	0	1	0	0	4	0
9	5	2	0	1	0	0	5	0
8	4	3	0	0	1	0	0	4
7	6	3	0	0	1	0	0	6

The following TEST request could be inserted after the model statement to produce an F-test of this hypothesis.

```
TEST G1 + 50*X1 = G2 +50*X2;
```

If groups 1 and 2 were the award and no award groups in the Campbell and Stanley regression discontinuity design, and if the value $X \geq 50$ qualifies one for an award, then the above statement would produce an appropriate test of the effect of the award on achievement.

The same type of test could be used for many values of X to locate regions on X for which there is a significant difference between the predicted Y values for groups 1 and 2. For example, if X was observed in the range (1,10), ten tests could be requested to locate the significance regions.

```
TEST G1 + 1*X1 = G2 + 1*X2;
TEST G1 + 2*X1 = G2 + 2*X2;
TEST G1 + 3*X1 = G2 + 3*X2;
TEST G1 + 4*X1 = G2 +4*X2;
TEST G1 + 5*X1 = G2 + 5*X2;
TEST G1 + 6*X1 = G2 + 6*X2;
TEST G1 + 7*X1 = G2 + 7*X2;
TEST G1 + 8*X1 = G2 + 8*X2;
TEST G1 + 9*X1 = G2 + 9*X2;
TEST G1 + 10*X1 = G2 +10*X2;
```

If the first three tests were significant and the last seven were not significant, the significance region would be $1 \leq X \leq 3$ and the nonsignificance region would be $4 \leq X \leq 10$. SASLOGs and listings with demonstrations of nonlinear ANCOVA extensions of these same tests can be obtained by writing the author.

Conclusion

The ANCOVA can provide a flexible approach to many analysis problems. Researchers are encouraged to use ANCOVA models that are structurally appropriate for their data and their research questions. This paper illustrated some simple tests of expected values that can be expressed as linear combinations of the model parameters. These simple tests were applied to the regression discontinuity design and the Johnson-Neyman significance region analysis. These applications were selected because they both are performed when a researcher encounters a problem with structural or design assumptions. The tests shown here illustrate how a researcher can articulate and test interesting hypotheses under these problematic conditions.

References

Campbell, D. T. and Stanley, J. C. (1968). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Elashoff, J. D. (1969). Analysis of covariance: A delicate instrument. *American Educational Research Journal, 6*, 383-401.

Lord, F. M. (1969). Statistical adjustments when comparing preexisting groups. *Psychological Bulletin, 72*, 336-337.

Mueller, R. O. (1990). Teaching ANCOVA: The importance of random assignment. *Multiple Linear Regression Viewpoints, 17* (2), 1-14.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research* (2nd Ed), New York: Holt, Rinehart and Winston.

SAS Institute Inc. (1990). *SAS/STAT user's guide* (Vol. 2). Cary, N.C.: SAS Institute Inc.

Winer, B. J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.

Teaching Ordinal and Criterion Scaling in Multiple Regression

Randall E. Schumacker
University of North Texas

John Delane Williams
University of North Dakota

This article illustrates the appropriate use of ordinal and criterion scaling techniques in multiple regression. Since multiple regression is a widely used data analytic technique, it is important to know how special coding is done to answer certain research questions. These coding techniques involve effect, characteristic, ordinal, or contrast coding of vectors for proper interpretation of statistical hypotheses.

The focus of this article is to demonstrate ordinal and criterion scaling techniques in multiple regression. Instructors might wish to include these topics when teaching multiple regression. Examples of each coding technique are presented to aid in understanding the approach and interpreting results.

Various authors have previously discussed specific coding strategies needed in multiple regression to answer certain research questions. For example, binary, or characteristic coding can be used to test research hypotheses involving group mean differences (Draper & Smith, 1966; Kerlinger & Pedhazur, 1973; Williams, 1974; McNeil, Kelly, & McNeil, 1975) or be used in multiple comparisons (Williams, 1974, 1976, 1980). Similarly, contrast coding can be used to investigate complex comparisons and other types of research questions (Lewis & Mouw, 1978). Effect, orthogonal, and polynomial coding techniques have also been elaborated (Cohen & Cohen, 1975; Pedhazur, 1982). Within orthogonal techniques, Helmert contrasts (see Bock, 1975) or polynomial regression can be completed. Newman (1988) presented several examples of how various coding strategies in multiple regression yield a t-test, analysis of variance, chi-square, discriminant, and other statistical results. Newman, Williams and Bobner (1982) had shown earlier that the Cochran Q test could be readily conceptualized into a regression format; they used a Monte Carlo study to show that the outcomes of using regression virtually coincided with the traditional Cochran Q analysis. Coding for two-way analysis of variance questions has also received close scrutiny (Bottenberg & Ward, 1963; Cohen, 1968; Overall & Spiegel, 1969; Overall, Spiegel, & Cohen, 1975; Speed & Hocking, 1976; Timm & Carlson, 1975; Ward & Jennings, 1973; Williams, 1972, 1977b).

Ordinal and criterion scaling techniques have not received as much attention in the research literature as binary, effect, orthogonal, and polynomial coding strategies. Their application to specific Likert (ordinal) scaled questionnaire data and repeated measures designs (criterion scaled), for example, have not been as well understood. Consequently, this paper presents ordinal and criterion scaling techniques in multiple regression.

Ordinal Scaling

Boyle (1970), Lyons and Carter (1971), and Lyons (1971) have elaborated on the use of ordinal scaling in multiple regression. Basically, ordinal scaling permits the interpretation of Likert (ordinal scaled) questionnaire data using multiple regression techniques. This approach defines the regression line between each ordinal point individually, disregarding the linear least squares rule applied to the entire set of data across the scale. The technique applies an eta-squared function and the relative contribution made by each segment of the ordinal variable; in essence, computing the slope of each regression line connecting the Y-means for successive categories of the ordinal scaled variable.

The cumulative nature of the coding in the regression equation is the basis for interpretation of the ordinal coefficients. Consequently, each successive beta weight represents the change in predicted Y from the previous category of the ordinal variable to the next. The ordinal approach doesn't force a uniform b_{yx} for the full range of values, but instead allows a separate prediction for each interval (a separate b_i for each segment between levels of the ordinal variable), and thus a maximum nonlinear prediction of Y given a specific category of the ordinal variable. The non-linear eta-squared value therefore will be equal to or greater than the linear least squares R-squared value. The ordinal interpretation is found in the b_i values themselves which are additive across categories of the ordinal variable.

Figure 1 Computer Program and Output: Ordinal Coding.

PROGRAM

```
TITLE     Regression analysis using ordinal coded variable
COMMENT   Interpretation of regression weights is additive
DATA LIST  FIXED RECORDS=1/ Y 1-2 SA 4 A 6 D 8
VARIABLE LABELS  Y 'score' SA 'strongly agree' A 'agree' D 'disagree'
BEGIN DATA
10      1      1      1
14      1      1      1
13      1      1      1
11      1      1      1
 9      0      1      1
11      0      1      1
12      0      1      1
12      0      1      1
 6      0      0      1
 9      0      0      1
10      0      0      1
11      0      0      1
 6      0      0      0
11      0      0      0
 7      0      0      0
 8      0      0      0
END DATA
REGRESSION VARIABLES = Y SA A D/
NOORIGIN/
DEPENDENT= Y/
METHOD= ENTER SA A D
FINISH
```

COMPUTER OUTPUT

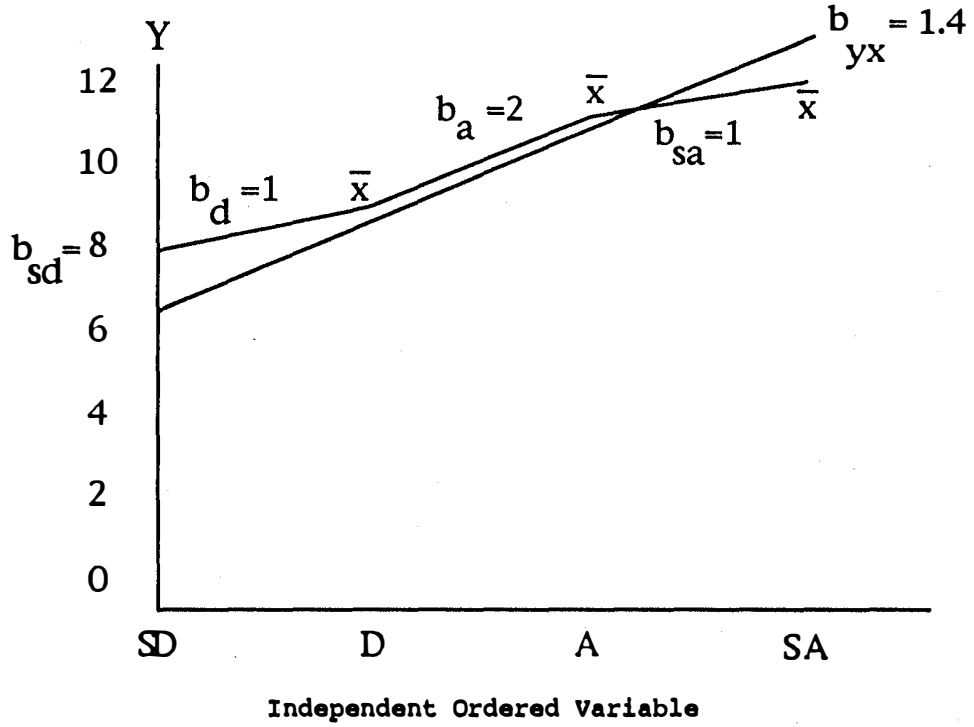
1. Analysis of Variance Summary Table

Source	df	Sum of Squares	Mean Square	F Value	Prob>F	R-Square
Model	3	40.00	13.33	3.63	.04	.47619
Error	12	44.00	3.67			
Total	15	84.00				

2. Parameter Estimates

Variable	df	Parameter Estimate	Standard Error	T for H0:	Prob > |T|
INTERCEPT (SD)	1	8.00	0.95	8.35	.0001
SA	1	1.00	1.35	0.74	.47
A	1	2.00	1.35	1.48	.16
D	1	1.00	1.35	0.74	.47

Figure 2 b_i Values for Ordinal Scaled Mean Compared to the Common Slope Value b_{yx}.



An example computer program with the required ordinal coding and the regression analysis output are in Figure 1. Corresponding parameter estimates are graphed in Figure 2, where the common slope, b_{yx}, is drawn relative to the successive b_i ordinal parameter estimate values. The dependent variable is score and the k-1 vectors represent responses to an ordinally scaled question. The subject responses are coded across k-1 vectors with the neutral or undecided response omitted. In the example, strongly disagree (SD), the first level of the ordinally scaled variable, has been omitted. Consequently, the value for b_{sd} is computed as the intercept term and reflects the starting value for interpreting change in the predicted Y values for successive ordinal categories. The ordinal coding causes the ordered variable means to range from lowest to highest.

Typically, the means for each ordinal category will deviate from linearity such that predicting Y would occur to a lesser extent than would be possible using a non-linear function with line segments between the ordinal categories. The inherent feature in interpreting ordinal coding is the cumulative or aggregate nature of responses across the categories. The predicted Y values (Y') are the respective sums of all the b_i values plus the intercept value (b_i for the omitted vector, e.g., $b_{sd} = 8$, and $b_d = 1$, $b_a = 2$, $b_{sa} = 1$ for each successive b_i). Each successive b_i reflects the change in the predicted Y value from the previous category. The parameter estimates are also additive to produce the next ordered variable group mean on Y, e.g., $\bar{X}_{sd} = 8$, $\bar{X}_d = 9$, $\bar{X}_a = 11$, $\bar{X}_{sa} = 12$. The important points for illustrative purposes is simply that ordinal coding does not compute a common slope (b_{yx}) across the full range of Y values, but instead permits separate prediction (separate b_i) between each successive ordinal category, and that a cumulative effect is apparent across levels of ordered categories.

How does ordinal scaling compare to other techniques? It might be seen as interesting that ordinal scaling yields similar results to binary coding the separate responses. The Likert response is binary coded for k-1 responses, (i.e., SA is binary coded as 1 if SA, 0 otherwise; A is binary coded as 1 if A, 0 otherwise; D is binary codes as 1 if D, 0 otherwise). Then, remarkably, if these three variables are used as predictors, an identical analysis of variance summary table to that in Figure 1 is formed. For both the ordinal scaling and the binary coding, $R^2 = .47619$. Performing

the binary coding rather than assuming interval level data was suggested by McNeil and Kelly (1970). The two coding approaches are similar because the process of creating the ordinal variables is very similar to testing for the departure from linearity as described by Bottenberg and Ward (1963).

The advantage of ordinal scaling is its intuitive appeal; it has a similarity to the unfolding technique described by Coombs (1964). To test for the departure from linearity, one additional vector (predictor) is needed; it is coded 4 for SA, 3 for A, 2 for D, and 1 for SD. The regression analysis yields a R^2 =.46667, b_{yx} = 1.4, and intercept = 6.5 (see Figure 2). The F-test for departure from linearity is given by:

$$F = \frac{\left(R^2_{FULL} - R^2_{RESTRICTED}\right)/(k-2)}{\left(1-R^2_{FULL}\right)/(N-k)} \qquad [1]$$

where R^2_{FULL} =.47619, calculated using either the ordinal scaling or the binary coding; where $R^2_{RESTRICTED}$ =.46667, calculated using the single predictor; k=4; and where N=16. Using these values, F = .213 which is not significant. This indicates that the ordered variable means do not significantly deviate from linearity.

To summarize, a variable's linear effect on a dependent variable can only be less than or equal to the nonlinear effect. A straight regression line can poorly describe the relationship between means to the extent that the means do not lie on a straight line (curvilinear path of categorical means). The test of departure from linearity will assess whether this exists. Moreover, to include ordinal variables in regression analysis is appealing, especially when one can discover which intervals, if any, contribute more or less and the degree of change from one category to another.

Criterion Scaling

Criterion scaling was first developed by Beaton (1969a; 1969b) to solve certain problems encountered in multiple regression. A basic problem occurred when using categorical variables because N - 1 vectors had to be created using dummy or effect coding. If the number of categorical independent predictors became large, then the number of vectors became overwhelming. A second problem pertained to variable selection methods where the categorical vectors might only be partially selected making interpretation difficult. Missing data on one of the predictor variables also presented a problem and usually meant exclusion from analysis even when a criterion value was present. These problems were resolved using the criterion scaling approach.

A categorical variable is criterion scaled when it is transformed into a single vector in which each individual score is replaced with the mean of the group to which the individual belongs (Pedhazur, 1982, p. 387). By criterion scaling a single categorical variable, the multiple regression analysis reduces to a bivariate regression analysis in which the dependent variable is regressed on the criterion scaled variable. This holds true regardless of the number of categories and for equal or unequal n's (Williams, 1977a). Comparing this to the traditional analysis of variance, this process removes within cell variability and leaves only variability due to group differences.

If the study involves several categorical variables, each variable can be criterion scaled separately and the criterion scaled variables used in the regression equation. Criterion scaling is therefore very useful when using a variable selection procedure to obtain the best set of independent predictors. For example, if five categorical variables resulted in 15-20 coded vectors, it may be difficult to have these sets of vectors added to or dropped from the equation as a set. (See however, Williams and Lindem, 1971 for a description of setwise regression). With criterion scaling, each categorical variable is represented by a single criterion scaled vector, so the problem is averted.

The degrees of freedom associated with criterion scaled variables however present a unique problem. Typically, the categorical variable is associated with k - 1 degrees of freedom. However, in criterion scaling, the variable will only have one degree of freedom reported by most computer programs. The actual degrees of freedom are k - 1. An example of criterion scaling is shown in Table 1. The three dummy coded vectors, D1, D2, and D3 are reduced to a single vector X1. This new vector is criterion scaled and contains the mean on Y for each respective dummy coded vector. The bivariate regression equation then becomes: $Y' = b_0 + b_2 X_2$, which would yield the same results as the dummy coded vectors in a regression analysis.

Table 1 Criterion Scaling Example

Score	Dummy Coded Vectors			Criterion Scale
Y	D1	D2	D3	X2
4	1	0	0	5
5	1	0	0	5
6	1	0	0	5
7	0	1	0	8
8	0	1	0	8
9	0	1	0	8
10	0	0	1	11
11	0	0	1	11
12	0	0	1	11

Criterion Scaling - Repeated Measures

Pedhazur (1977, 1982) and Williams (1977a) elaborated the usefulness of criterion scaling in treatment by subject repeated measures designs. This approach involves reducing the coding of N - 1 vectors to represent subjects into a single vector whereby each subject in a treatment group receives the sum of the criterion scores for that group. This single vector in a bivariate regression analysis yields the same R-squared value as does the N - 1 binary coded subject vectors. The traditional analysis proceeds with three linear models (Williams, 1974): (a) treatment effects, (b) subject effects, and (c) combined treatment and subject effects.

The three regression equations can be expressed as:

$$Y_{treat} = b_0 + b_2X2 + b_3X3 + e_1 \quad [2]$$

$$Y_{subj} = b_0 + b_4X4 + b_5X5 + \ldots b_{12}X12 + e_2 \quad [3]$$

$$Y_{comb} = b_0 + b_2X2 + b_3X3 + \ldots b_{12}X12 + e_3 \quad [4]$$

The criterion scaling approach also involves three linear models, however, the subject effects in equation [3] and the combined treatment and subject effects in equation [4] would be substantially reduced as follows:

$$Y_{subj} = b_0 + b_1X1 + e_2 \quad [5]$$

$$Y_{comb} = b_0 + b_1X1 + b_2X2 + b_3X3 + e_3 \quad [6]$$

The criterion scaling for this example treatment by subjects design is in Table 2. The variable, X_1, is the criterion scaled vector which reduces the N-1 subject vectors (X4 to X12) into a single vector for bivariate regression analysis to obtain the subjects effect in equation [5]. Score sums on Y rather than means were used in the criterion scaled vector, which is permissible.

Table 2 Criterion Scaling in Treatment by Subject Design[a]

Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
10	33	1	0	1	0	0	0	0	0	0	0	0
11	33	0	1	1	0	0	0	0	0	0	0	0
12	33	0	0	1	0	0	0	0	0	0	0	0
13	42	1	0	0	1	0	0	0	0	0	0	0
14	42	0	1	0	1	0	0	0	0	0	0	0
15	42	0	0	0	1	0	0	0	0	0	0	0
14	38	1	0	0	0	1	0	0	0	0	0	0
13	38	0	1	0	0	1	0	0	0	0	0	0
11	38	0	0	0	0	1	0	0	0	0	0	0
6	21	1	0	0	0	0	1	0	0	0	0	0
7	21	0	1	0	0	0	1	0	0	0	0	0
8	21	0	0	0	0	0	1	0	0	0	0	0
10	35	1	0	0	0	0	0	1	0	0	0	0
14	35	0	1	0	0	0	0	1	0	0	0	0
11	35	0	0	0	0	0	0	1	0	0	0	0
15	35	1	0	0	0	0	0	0	1	0	0	0
12	35	0	1	0	0	0	0	0	1	0	0	0
8	35	0	0	0	0	0	0	0	1	0	0	0
14	44	1	0	0	0	0	0	0	0	1	0	0
15	44	0	1	0	0	0	0	0	0	1	0	0
15	44	0	0	0	0	0	0	0	0	1	0	0
12	39	1	0	0	0	0	0	0	0	0	1	0
17	39	0	1	0	0	0	0	0	0	0	1	0
10	39	0	0	0	0	0	0	0	0	0	1	0
22	60	1	0	0	0	0	0	0	0	0	0	1
21	60	0	1	0	0	0	0	0	0	0	0	1
17	60	0	0	0	0	0	0	0	0	0	0	1
11	44	1	0	0	0	0	0	0	0	0	0	0
18	44	0	1	0	0	0	0	0	0	0	0	0
15	44	0	0	0	0	0	0	0	0	0	0	0

[a] Y (criterion), X1 (criterion scaled), X2 and X3 (Treatments) X4 to X12 are the N-1 subject coded vectors.

Table 3 Criterion Scaled Analysis of Variance Summary

Source	SS	df	MS	F	R^2
Treatments	21.67	2	10.83	2.22	.0532
Subjects	297.63	9			.7314
Residual	87.66	18	4.87		.2154
Total	406.96	29			1.0000

The results obtained are identical with the traditional N - 1 coded subject vectors (sum of squares and R-squared values), except the degrees of freedom are reported as N - 1, because X1 contains (N -1) linearly independent vectors. The researcher must correct the degrees of freedom when using general purpose multiple regression computer programs. The criterion scaled analysis of variance results are presented in Table 3. The combined subject and treatment effects yield R^2 = .7846.

If multiple comparisons were of interest, the tests of significance for comparing Group 1 to Group 3 and Group 2 to Group 3 would be derived from the reported t tests for the regression coefficients b_2 and b_3 respectively in equation [6] (output not shown). Because of the degrees of freedom issue already described, the reported t's would need to be adjusted by multiplying by:

$$\sqrt{\frac{N-S-g+1}{N-g-1}} \qquad [7]$$

where N is the number of observations (treatment by subject combinations), S is the number of subjects and g is the number of groups; an appropriate table, such as Dunnett's (1964), Dunn's (1961), or some other multiple comparison procedure being used now can be entered.

Summary

This article focused on presenting ordinal and criterion scaling techniques in multiple regression. The computer programs, coding, and output examples reflect the need to teach how special coding can be used to answer certain research questions. These coding techniques also highlight the need for proper interpretation of results.

Ordinal scaling techniques in multiple regression provide for the analysis and interpretation of ordinal variables such as Likert scaled questionnaire data. The regression weights provide a step interpretation between each point on the scale or the degree of change from one category to the next. A test for the departure from linearity can also be conducted.

Although only one example was presented, criterion scaling techniques can solve many of the problems encountered in multiple regression, namely, (a) the use of extensive categorical variables, (b) the selection of the best set of predictor variables (maximize R-squared), (c) coding of subjects in repeated measures designs, and (d) the handling of predictor variables with missing values. Criterion scaled vectors may contain either means or sums in the case of equal n's. The criterion scaling technique is applicable to both linear and non-linear regression line fitting (Hinkle, Wiersma, & Jurs, 1988, pp. 540-544). Continuous predictor variables can also be criterion scaled by dividing them into equal intervals with the scores in each interval coded the value of the mean on the criterion for that category. In the case of multiple predictors, each predictor is coded into a single vector, then all possible regression techniques can be applied. Remember, however, that the degrees of freedom must be adjusted to N-1 not the reported df = 1. Mixed regression models which combine categorical and continuous predictors are also possible (Gocka, 1973).

References

Beaton, A. E. (1969a). Scaling criterion of questionnaire items. *Socio-Economic Planning Sciences, 2,* 355-362.

Beaton, A. E. (1969b). Some mathematical and empirical properties of criterion scaled variables. In O. W. Mayeske (Ed.), *A study of our nation's schools.* Washington, D.C.: Office of Education.

Bock, R. D. (1975). *Multivariate Statistical Methods in Behavioral Research.* New York: McGraw-Hill.

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

Boyle, R. P. (1970). Path analysis and ordinal data. *American Journal of Sociology, 75,* 461-480.

Cohen, J. (1968). Multiple regression as a general data analytic system. *Psychological Bulletin, 81,* 335-343.

Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Coombs, C. H. (1964). *A theory of data.* New York: Wiley.

Draper, N. R., & Smith, H. (1966). *Applied regression analysis.* New York: Wiley.

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association, 56,* 52-64.

Dunnett, C. W. (1964). New tables for multiple comparisons with a control. *Biometrics, 20,* 482-491.

Gocka, E. F. (1973). Stepwise regression for mixed mode predictor variables. *Educational and Psychological Measurement, 33,* 319-325.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1988). *Applied*

statistics for the behavioral sciences. Boston: Houghton Mifflin.

Kerlinger, F. N., & Pedhazur, E. J. (1973). *Multiple regression in behavioral research*. New York: Holt, Rinehart, & Winston.

Lewis, E. L., & Mouw, J. T. (1978). *The use of contrast coefficients*. Carbondale, IL: Southern Illinois University Press.

Lyons, M. (1971). Techniques for using ordinal measures in regression and path analysis. In H. L. Costner (Ed.), *Sociological methodology* (pp. 147-171). San Francisco, CA: Josey-Bass.

Lyons, M., & Carter, T. M. (1971). Comment on Boyle's path analysis and ordinal data. *American Journal of Sociology, 76*, 1112-1132.

McNeil, K. A., & Kelly, F. J. (1970). Express functional relationships among data rather than assume "intervalness". *The Journal of Experimental Education, 39*, 43-48.

McNeil, K. A., Kelly, F. J., & McNeil, J. T. (1975). *Testing research hypotheses using multiple linear regression*. Carbondale, IL: Southern Illinois University Press.

Newman, I. (1988, October). *There is no such thing as multivariate analysis: all analyses are univariate!* Presidential address presented at the annual Midwestern Educational Research Association meeting, Chicago, IL.

Newman, I., Williams, J. D., & Bobner, R. (1982). Dependent binary data: A comparison of regression and chi-square solutions. *Proceedings of the American Statistical Association, Statistical Education*, 178-183.

Overall, J. E., & Spiegel, D. H. (1969). Concerning least squares analysis of experimental data. *Psychological Bulletin, 72*, 311-322.

Overall, J. E., Spiegel, D. H., & Cohen, J. (1975). Equivalence of orthogonal and nonorthogonal analysis of variance. *Psychological Bulletin, 82*, 182-186.

Pedhazur, E. J. (1977). Coding subjects in repeated measures designs. *Psychological Bulletin, 84*, 298-305.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research*. New York: CBS College Publishing.

Speed, F. A., & Hocking, R. R. (1976). The use of R () - notation with unbalanced data. *The American Statistician, 30*, 30-33.

Timm, N. H., & Carlson, J. E. (1975). Analysis of variance through full rank models. *Multivariate Behavioral Research: Monograph, 75*(1).

Ward, J. H., & Jennings, E. (1973). *Introduction to linear models*. Englewood Cliffs, NJ: Prentice Hall.

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

Williams, J. D. (1974). *Regression analysis in educational research*. New York: MSS Publishing.

Williams, J. D. (1976). Multiple comparisons by multiple linear regression. Monograph Series No. 2, *Multiple Linear Regression Viewpoints, 7* (l).

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

Williams, J. D. (1977b). Full rank and non-full rank models with contrast and binary coding systems for two-way disproportionate cell frequency analyses. *Multiple Linear Regression Viewpoints, 8*(1), 1-18.

Williams, J. D. (1980). Multiple comparisons in higher dimensional designs. Monograph Series No. 5, *Multiple Linear Regression Viewpoints, 10*(3).

Williams, J. D., & Lindem, A. C. (1971). Setwise regression analysis: A stepwise procedure for sets of variables. *Educational and Psychological Measurement, 31*, 747-748.

RELATED AERA SPECIAL INTEREST GROUPS

As a member of the SIG/MLR, the AERA Special Interest Groups listed below also might be of interest to you. The 1993 AERA Annual Meeting provides the opportunity to attend a session sponsored by one of these groups or to contact an officer to obtain additional information and a membership application.

SIG/Educational Statisticians
CHAIR: Schuyler Huck, University of Tennessee
SECRETARY: Lynne Edwards, University of Minnesota

SIG/Professors of Educational Research
CHAIR: William Schafer, University of Maryland
SECRETARY: Bruce Thompson, Texas A&M University

SIG/Structural Equation Modeling
CHAIR: Ralph Mueller, University of Toledo
SECRETARY: Jack Pollard, SWRL

1993 SIG/MLR Program at the AERA Annual Meeting

April 12th - 16th Atlanta, Georgia

The AERA Special Interest Group on Multiple Linear Regression will be sponsoring three sessions at this year's AERA Annual Meeting. Each of these sessions will no doubt be interesting and informative. Plan now to support the SIG/MLR and attend all three sessions.

Monday, April 12, 6:15 - 8:15 pm
Quebec Room, 2nd Level, Marriott

Business Meeting (Session 8.38)
CHAIR: Keith McNeil, New Mexico State University
EXECUTIVE SECRETARY: Steven Spaner, University of Missouri, St. Louis
HONOREE: Isadore Newman, University of Akron
Retiring Editor, *Multiple Linear Regression Viewpoints*

Tuesday, April 13, 8:15 - 10:15 am
Flemish Room, Terrace Level, Marriott

Applications of Multiple Linear Regression (Session 11.46)
(Co-sponsored with SIG/Structural Equation Modeling)
CHAIR/DISCUSSANT: John Pohlmann, Southern Illinois University

Type IV Errors in Path Analysis: Testing for Interaction
Isadore Newman, University of Akron & Gregory J. Marchant, Ball State University

Using Linear Regression to Determine the Number of Factors to Retain in Factor Analysis
Ralph Mueller, Stephen Jurs, & Keith Zoski, University of Toledo

Wednesday, April 14, 8:15 - 10:15 am
Yukon Room, 3rd Level, Marriott

Multiple Linear Regression Methodology (Session 25.51)
CHAIR: Steven Spaner, University of Missouri, St. Louis
DISCUSSANT: Susan Tracz, California State University, Fresno

Full Versus Restricted Model Testing in Discriminant Analysis Applied to Personnel Selection
Susan Whiting, Broward County Schools & John D. Morris, Florida Atlantic University

Treatments for Missing Data in Multiple Regression: Stochastic Imputation, Deterministic Imputation, and the Deletion Procedures
Lantry Brockmeier, Constance V. Hines, & Jeffrey D. Kromrey, University of South Florida

A Comparison of the Accuracy of Analytical and Empirical Estimates of Shrinkage in Multiple Regression
Jeffrey D. Kromrey & Constance V. Hines, University of South Florida

Information for Contributors (revised)

Multiple Linear Regression Viewpoints (MLRV), a publication sponsored by the American Educational Research Association's Special Interest Group on Multiple Linear Regression (SIG/MLR), is published biannually (Spring, Fall) to facilitate communication among professionals who focus their investigations on the theory, application, or teaching of multiple linear regression models or their extensions. Also, the journal accepts news items of interest to members of the SIG/MLR.

All manuscripts submitted to *MLRV* should conform to the language, style, and format of the *Publication Manual of the American Psychological Association* (3rd ed., 1983), available from Order Department, American Psychological Association, P.O. Box 2710, Hyattsville, MD 20784. Three copies of the manuscript, all double spaced (including equations, footnotes, quotes, and references) and accompanied by an abstract of 100 words or less, should be submitted to the editor at the address listed below. Mathematical symbols and Greek letters should be precise and clear, and should leave no question as to interpretation. All figures must be camera ready. Manuscripts that do not conform to the above specifications may be returned to the author for style changes before the review process will begin. A submitted manuscript will receive a blind review from at least two members of the editorial board (except occasional invited contributions, letters to the editor, editorials, or news items). Any author identifying information should appear on the title page only. Efforts will be made to keep the review process to a maximum of eight weeks. The final version of an accepted manuscript should be submitted on a 3.5" disk, preferably in Apple Macintosh Microsoft Word, Version 5.1, although other formats might be acceptable. The editor reserves the right to make minor changes to an accepted manuscript in order to facilitate a clear and coherent publication.

Potential authors are encouraged to contact the editor to discuss ideas for contributions or to informally determine whether manuscripts might be appropriate for publication in *MLRV*. The editor also welcomes suggestions for debates, theme issues, other innovative presentation formats, and general inquiries about the journal. SIG/MLR news items should be sent to the editor as soon as they become available.

Manuscripts and other correspondents with the editor should be addressed to:

Ralph O. Mueller, Editor, *MLRV*
Department of Educational Psychology,
Research and Social Foundations
College of Education and Allied Professions
The University of Toledo
Toledo, OH 43606

phone: (419) 537-4722
fax: (419) 537-7719
bitnet: FAC2464 @ UOFT01

The University of Toledo
College of Education and Allied Professions
Philip J. Rusche, Dean