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.

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

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

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

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.

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

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.

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

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

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

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

with

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

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

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

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

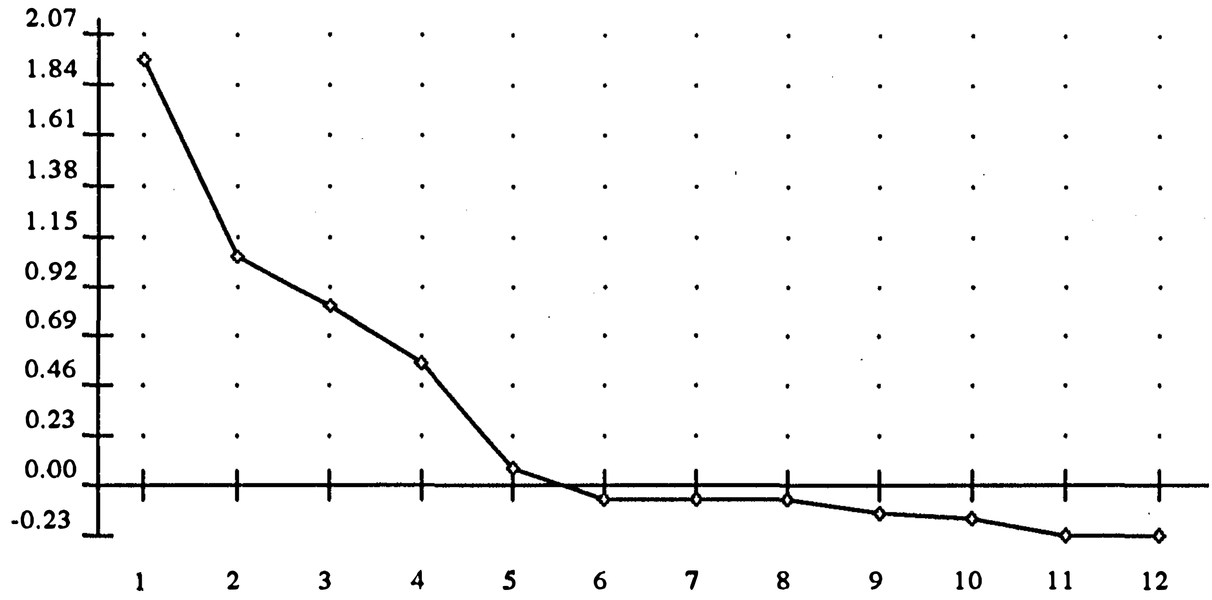
We have compared the multiple regression approach to the CNG approach using several data sets from the literature. Preliminary results indicate that the multiple regression approach usually agrees with a visual scree test and often provides a better solution than the CNG method.

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).

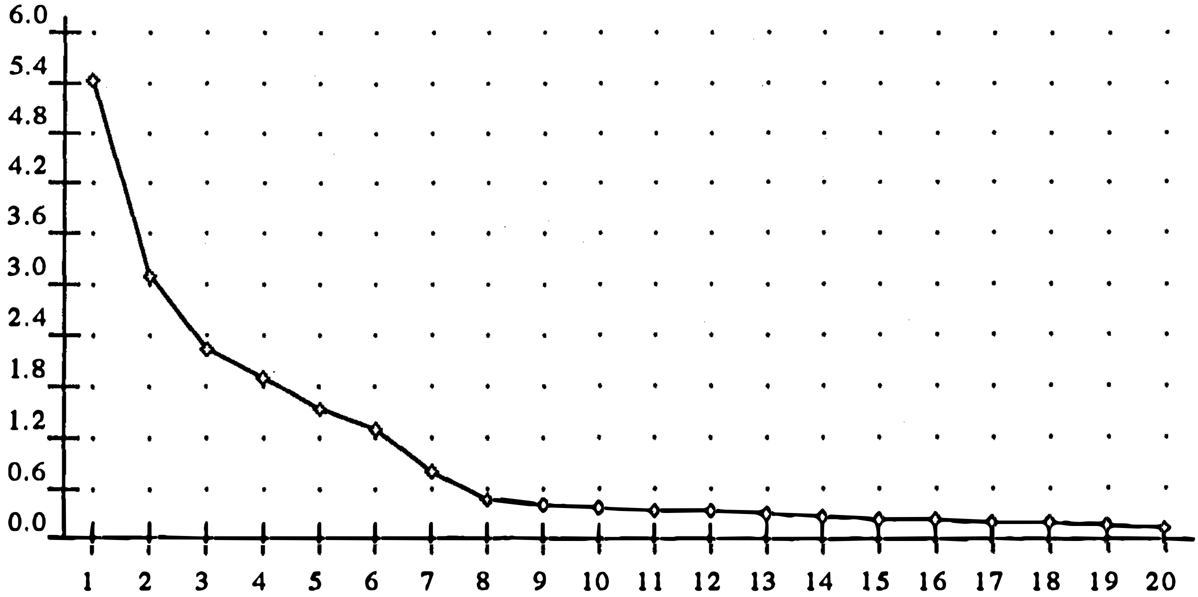


The second example is taken from Tucker, Koopman, & Linn (1969, p. 442). The plot of the eigenvalues is given in Figure 2 and the results from the multiple regression approach and the CNG approach are listed in Table 2. The data set was meant to have seven factors. The CNG approach yielded three factors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

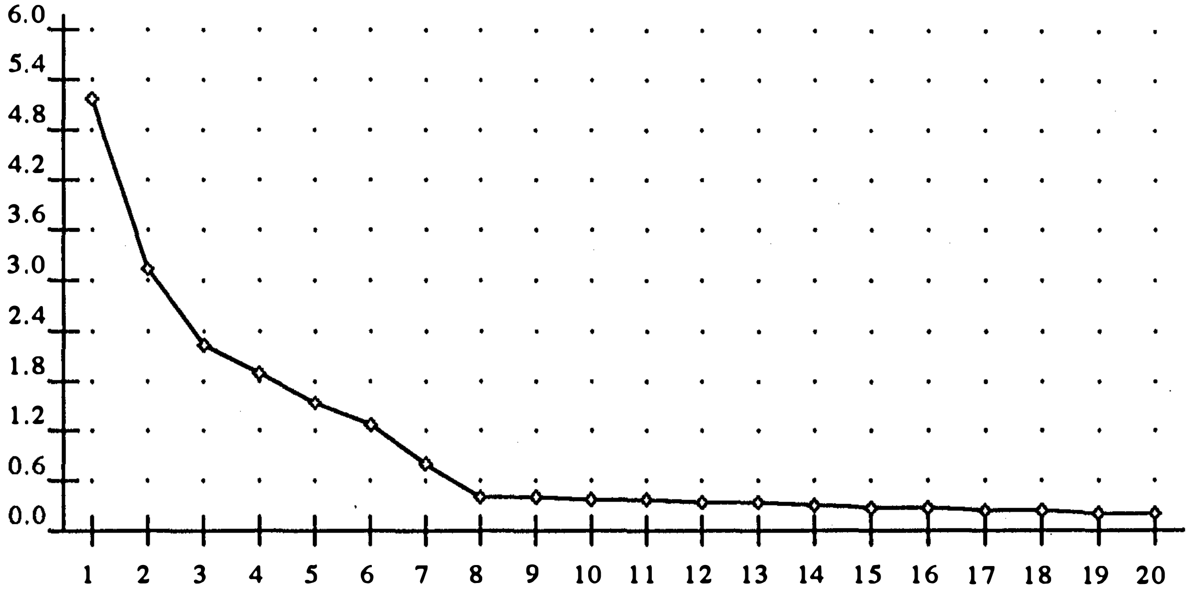


The third example was also taken from Tucker, Koopman, & Linn (1969, p. 442). This data set was intended to have seven factors and a visual inspection of the scree plot in Figure 3 suggests that there are seven factors. The analyses presented in Table 3 indicate that the CNG approach yielded only three 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.