Bruce
Here is our W.E.A.
article. Sorry
for the delay.
Shirley

THE LOG-LINEAR MODEL AND PATH ANALYSIS: TWO MULTIVARIATE

MODELING AND ANALYSIS TECHNIQUES FOR MICROCOMPUTERS

Shirley C. Anderson, Calif. State U., Northridge
Judith E. Hennessey, Calif. State U., Northridge
H. Bruce Lammers, Calif. State U., Northridge

ABSTRACT

A common data set was used to demonstrate two techniques, available on microprocessors, that aid in research design for predictive purposes. Both techniques, the Log-linear model and Path Analysis, provide a framework for specifying all possible models of a process as well as finding the best fitting relationships, and Path Analysis can be interpreted as finding the best fitting causal realtionships. Also differences in design and analysis assumptions that limit the applicability of the two techniques were discussed.

INTRODUCTION

Analysis in Marketing and Economics might be most broadly viewed as concerned with prediction and/or with examining the interdependence between questions, variables, or objects. The analysis of dependence involves prediction or specification of the dependent variable(s) given a set of specific values for the independent variables. A technique for analysis of dependence may be constrained by the minimum properties of both the dependent and the independent variable(s) (i.e. nominal, ordinal or interval).

It is our goal to demonstrate the application of two alternative multivariate analysis of dependence techniques: Path Analysis, which requires interval data, and the Log-Linear Model, which can also be used with nominal data.

There are three critical assumptions limiting the applicability of the Log-linear Model. First, separate observations must be statistically independent. Also, all observations must be identically distributed and, finally, the number of observations must be large.

The second analysis of dependence technique is path, or causal, analysis which requires both the dependent and the independent variables to have at minimum interval properties. The form of Path Analysis under consideration here is a regression technique and thus is constrained by the assumptions for regression analysis.

THE DATA BASE

The study involved an examination of speech compression and message appeal on the relative effectiveness of a radio advertisement. The 2 by 3 design was based on two factors: 1. appeal (humor or serious) and 2. compression level. Each version of the advertisement was compressed using a Lexicon speech compressor. A normal, 20 percent compressed, and 40 percent compressed tape was developed for each version of the advertisement. Subjects were randomly assigned to one of the six experimental conditions. For each cell in the design there were from 22 to 27 subjects for a total of 146 subjects in the experiment. The instructions explained that subjects would be asked to listen to a brief radio advertisement and then asked to respond to what they had heard. After listening to the radio advertisement, subjects were asked to jot down on the following page any thoughts they had while listening to the message. Two minutes was alloted for reporting the cognitive responses(Wright,1973). Subsequently, subjects were asked to turn to the next page of their response booklet where they were asked to evaluate their own cognitive responses on a set of nine-point "favorability" scales. Subjects were then asked to evaluate the advertisement on a set of seventeen semantic differential scales. These scales were designed to assess the evaluative, belief and intention components of attitude (Fishbein, 1975). These scales included such items as persuasive/unpersuasive and good/bad.

Evaluations of their own cognitive responses were used to tabulate the frequency of positive, negative and neutral Cognitive Responses for each subject. Q-mode factor analysis of the subjects on the basis of their semantic differential scale responses was used to group the subjects relative to the three components of attitude: evaluative, belief and intention or behavioral intention. The intent of the model-building and testing in this study was to develop a predictor of the subjects' behavioral intention (or buying intention).

The consumer behavior literature treats both cognitive responses and semantic differential scales such as these, as adequate tools for forecasting buyer behavior. While the semantic differential scales are currently a very popular instrument, cognitive responses may be less likely to force a biased response from the subject.

RESULTS

Log Linear Model and the G^2 Statistic

The Log Linear Model accomodates multinomial data, which regression analysis cannot do without biased measurements of fit. The categories for any variable in the analysis (dependent or independent) must be independent of one another. The analysis of cognitive responses was conducted by assuming that positive, neutral and negative cognitive responses were generated from independent response categories. Thus a 3-way contingency table was generated with one dimension identifying Compression level (C), one dimension identifying Response category (J), and one dimension identifying level of Appeal (A).

A problem arises in using a traditional chi-square analysis in analyzing this matrix. Predictions based on marginal frequencies will confound the effect of the diagonal cells with the effects of the offdiagonal elements. Since this should relate to different aspects of the process to be explained repeated chi-squares with multinomial data are inappropriate. When the number of observations is reasonably large

$$\chi^2 = 2 \sum_{i=1}^{c} f_i \log(f_i/n\pi_i) \qquad (1)$$

$$= 2\left(\sum_{i=1}^{c} f_i \log f_i - \sum_{i=1}^{c} \log \pi_1 - n \log n\right) \qquad (2)$$

is also distributed as a chi-square variable. This is the likelihood-ratio chi-square. Asymptotically, the Pearson chi-square and the likelihood-ratio chi-square are identical (for further discussion see Bishop, Feinberg and Holland, 1975). In the log-linear model, the logarithmic transformation is used so that the predicted value for each cell reduces to a linear sum of component effects. A test against the model of perfect fit is given by the likelihood-ratio chi square:

$$G^2 = -2 \sum f \log(e/f) \qquad (3)$$

where f refers to the observed frequencies and e refers to the expected frequencies.

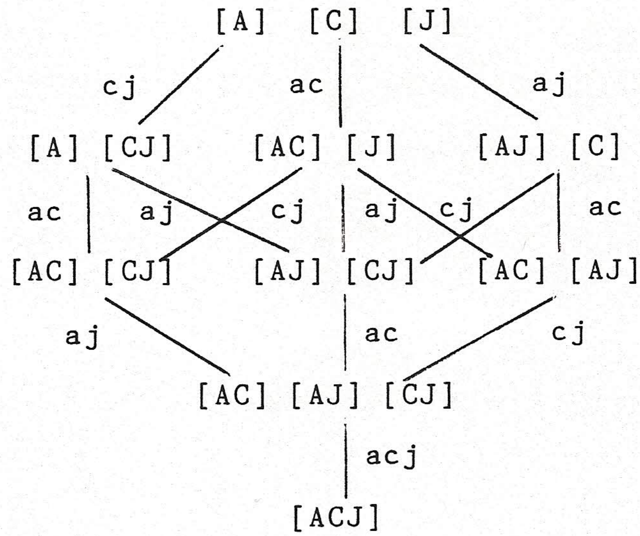
The log-linear model allows hypotheses involving linear contrasts between sets of cells to be tested directly. Further the log-linear model is used here to find a minimum descriptive set of variables which adequately describes the relationships in the data. Analogous to ANOVA methods, significance of a particular effect as well as its magnitude can be tested by the log-linear model. Finally, specific sets of cells can be omitted from the analysis without difficulty because of its additive structure.

Earlier work in speech compression had indicated greater interest, comprehension, and persuasive effects of a message that has been electronically compressed (LaBarbera and MacLachlan, 1979; MacLachlan and Siegel, 1980). Similar differences were reported for different message appeals. Humorous messages have been reported to increase interest and persuasiveness of the message and to decrease comprehension. On the other hand serious messages resulted in less interest and more comprehension (Sternthal and Craig, 1973). One of the goals of this research is to test the applicability of a behavioral intentions model for these effects in addition to the cognitive response model used in the earlier work.

In the following discussion, the symbols for Compression and Cognitive Response effects are C and J, respectively. Message Appeal is represented by A. These are constant throughout all analyses presented here.

A lattice of models can be formed which express all possible hierarchical arrangements of contributing effects. A set of models relevant to these data collected under various Appeals is presented below:

Fig. 1



The two extremes represent the "independence" model, [A] [C] [J], and the saturated model [ACJ]. When terms appear together within the same set of square brackets, an association or interaction exists between those terms. Terms which do not share a common bracket are considered to be independent. The small case letters on the connecting lines between models represent the association or term that defines the difference between any two adjacent models. One procedure for modeling is simply to find a model which describes the data matrix best. The approach used for testing specific effects is to subtract the G^2 values of two adjacent models in the hierarchy to yield the single parameter (represented by small case letters in Figure 1) by which the two models differ.

The best fitting model for this analysis was found by subtracting the degrees of freedom and G^2 for the [A] [CJ] model from the degrees of freedom and G^2 for the [A] [C] [J] model.

Model	Degrees of Freedom	Log Likelihood Ratios (G^2)
[A] [C] [J]	12	20.73
[A] [CJ]	8	10.64
cj	4	10.09

Thus compression is associated with Cognitive Response (S) but appeal is not ($X^2(4)=10.09$; $p<.025$). Examining the [CJ] association by examining the changes in the fit of the model as a single level of each variable is eliminated revealed that the association was determined by the zero compression condition ($G^2(5)=2.67$; $p=.25$) and the negative cognitive responses ($G^2(5)=4.62$; $p=.50$). Therefore, in forecasting buyer behavior this analysis demonstrates that such can be most efficiently done by examining only negative cognitive responses and compression level without regard to the particular appeal used in the message. This analysis used the BMDP program 4F to calculate the log likelihood ratios from which these results were derived. The BMDP series with program 4F is now available for many microprocessors.

Path Analysis

Causal modeling, or path analysis, is a method of developing a complete model and estimating the predicted relationships useful for the purpose of developing a forecasting model. Path analysis is used to estimate the magnitude of the linkages between variables in order to provide information about the underlying causal processes. The simplest way to obtain the path coefficients is to employ ordinary regression techniques. To obtain estimates of the main path coefficients, each endogenous variable is regressed on the variables that directly impinge upon it. For example, the path coefficient between Cognitive Response (J) and Semantic Differential Scales (S) is the standardized regression coefficient of S obtained by regressing J on S. The corresponding residual coefficient can also be calculated from the regression of J on S as

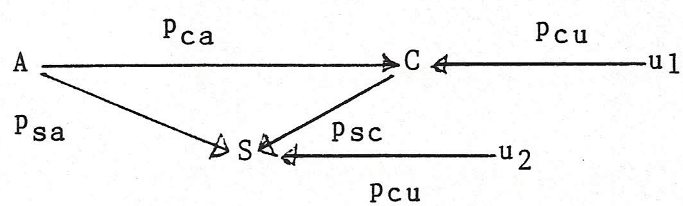
$$P_u = \sqrt{1-R^2}. \qquad (4)$$

The first predictive model to be developed is of the relationships among the 3 variables: Appeals (A), Compression (C) and the Sum of Semantic Differential Scales Response (S). The possible alternative models are the following:

Model 1A. [S] [A] [C]
Model 1B. [S] [AC]
Model 1C. [SA] [C]
Model 1D. [SC] [A]
Model 1E. [CA] [SC]
Model 1F. [CS] [AS]
Model 1G. [AC] [AS]
Model 1H. [SA] [AC] [CS]

The paths among the three variables can be displayed in an arrow diagram. In such a diagram the exogenous variables are those which are not affected by other variables. The remaining measured variables are endogenous. The u's are residual or disturbance terms that impinge upon the endogenous variables. The p_{ij} represent path coefficients.

Fig. 2 Model 1



The direction of the relationships must be developed from theoretical or temporal relationships, not from the path analysis. In particular, if a linkage between two variables has been omitted from the model, then implicitly we have said that we expect the magnitude of the path coefficient to be zero. Thus the test of the model becomes whether the omitted coefficient is indeed zero. The magnitude of the omitted linkage is determined by estimating the model with the linkage under question included.

Any path diagram leads directly to a set of equations which represent the paths as standardized regression coefficients. The Model 1 equations, assuming the given causal directions shown by the direction of the arrows, are as follows:

$$C = p_{ca}A + p_{cu}U_1$$

$$S = p_{sa} + p_{sc}C + p_{cu}U_2$$

The results of computing the path coefficients as the Beta or standarized regression parameters of the above equations for our data are shown below:

$$C = 0.007\ A + 1.0\ U_1$$

$$S = 0.23\ A + -.25\ C + 0.94\ U_2$$

Thus C is unrelated to A and S is weakly related to both A and C. These results refute the model of independence (1A) and models 1B-1E, 1G and 1H. The model which best fits the data is Model 1F, which is a different model from that found to best fit the Cognitive Response (J) data with the log-linear technique (Model 1D).

Comparison of the residual paths under different variable specifications can be used to reduce measurement error. For example, Q-mode factor analysis of the subjects on the seventeen semantic differential scales results in three factors that describe 67% of the variance in the semantic differential scales, S. The three factors are recognizable and interpreted as describing three important aspects of the buyers' thought processes: evaluation, belief and buyer intention.

A weighted average of the three factors (weighted by the fraction of total variance explained: 45%, 14% and 8%, respectively) was constructed and this measure (F) was substituted for S in the path analysis. The path coefficient for the residual term in the second equation increased, indicating that the factor score variable, F, has less predictive power than the measure it replaced.

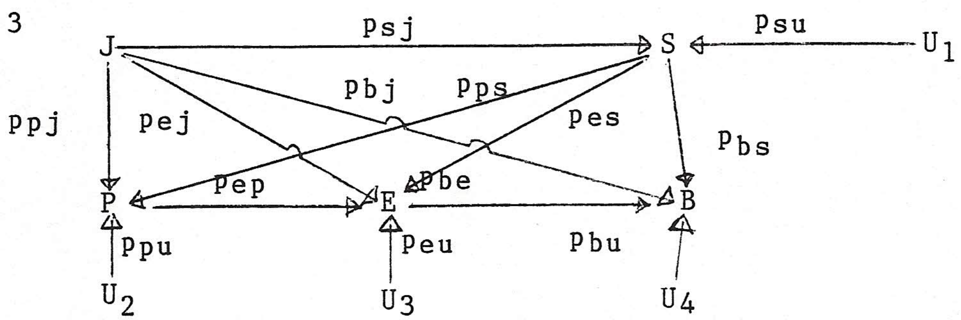
$$F = 0.006\ A - 0.310\ C + 0.95\ U_2$$

Use of the factor score variable, with this model, did not fit the data as well. This analysis, however, suggests the same model as the log-linear analysis found best fit the data.

Redundancy is a built in feature of the semantic differential scales while such redundancy is not built into Cognitive Response measures. Thus it is theoretically more sound to eliminate the redundancy in semantic differential responses by using factor scores in the analysis. In another case a factor score variable might increase explanatory power by reducing the effect of redundant measures on the questionnaire, and thus reducing variance. Regardless of the source of measurement error, comparison of the residual paths under alternate variable specification is a guide toward measurement error reduction. However, this must be tempered by judgement from theoretical models.

A second path analysis was made of the relationships among four variables: Cognitive Response (J), Semantic Differential Scales (S) and the three factors: Belief (P), Evaluation (E) and Buying Intention (B). The J is average favorability rating, instead of the frequency of various categories of response as was used in the log-linear analysis. This allows interpretation of our Cognitive Responses as interval data. Path analysis can then be used on each direct and indirect relationship to examine all possible logical models and eliminate models that have no predictive power for forecasting Buying Intention. The estimation equations are listed following the arrow diagram, Figure 3.

Fig. 3



$$S = p_{sj}J + p_{su}U_1$$

$$P = p_{pj}J + p_{ps}S + p_{pu}U_2$$

$$E = p_{ej}J + p_{es}S + p_{ep}P + p_{eu}U_3$$

$$B = p_{bj}J + p_{bs}S + p_{be}E + p_{bu}U_4$$

The values of the path coefficients, derived from OLS regression analysis of each equation, are as follows:

$$S = 0.39\ J + 0.92\ U_1$$

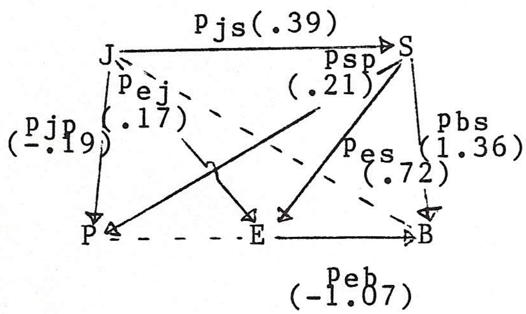
$$P = -0.19\ J + 0.21\ S + 0.97\ U_2$$

$$E = 0.17\ J + 0.72\ S - 0.06\ P + 0.59\ U_3$$

$$B = 0.01\ J + 1.36\ S - 1.07\ E + 0.52\ U_4$$

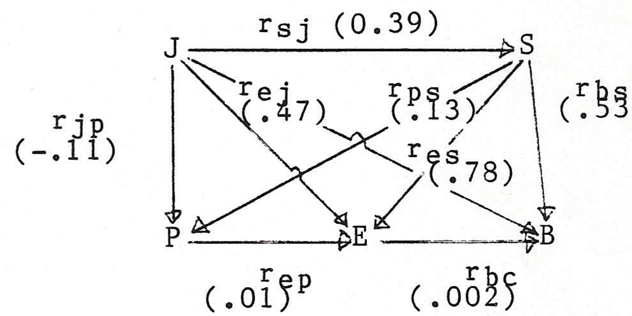
The path coefficients show that information on the sum of the semantic differential scales (S) is a much better predictor of Buyer Intention (B) than is knowledge of Belief or Evaluative responses. It must be recognized that buyer intention scales may still be less related to buying behavior than cognitive responses. On the basis of the available data, the improved model of buyer intention, Figure 4A, eliminates certain linkages that have path coefficients close to zero.

Fig. 4A.



Direct effects, holding other variables constant (Multiple regression analysis).

Fig. 4B.



Direct effects (bivariate regression analysis).

Figure 4B shows the magnitudes of the simple correlations between each pair of two variables in the model, obtained as the Beta coefficients of simple regression. These coefficients can be decomposed into direct and indirect effects that one variable has upon another. Then comparing the magnitude of the direct and indirect effects can be used to identify the causal mechanisms of the model, which will help in prediction.

The simplist way to decompose the coefficients of Fig. 4B into direct and indirect effects is to use Sewall Wright's rules on the arrow diagram.

(a) No path may pass through the same diagram more than once.
(b) No path may go backward on (against) an arrow after the path has gone forward on a different arrow.
(c) No path may pass through a double-headed curved arrow (representing an unanalyzed correlation between exogenous variables) more than once in any single path.

The decomposition of the correlations yields information about the causal processes. It also provides a way to test the adequacy of the model: if the model is specified correctly, then (except for measurement and sampling error) the empirical correlation between any two variables should be numerically equal to the sum of the simple and compound paths linking the two variables. If the equality does not hold, this suggests that the model may be improperly specified and in need of revision.

For example, in Figure 4B, consider the total correlation between S and B, i.e., r_{bs}. This correlation can be divided into direct and indirect relationships using Wright's rules. To do this, imagine trying to move from S to B, in Figure 3, and try to find all the paths from S to B that do not violate any of Wright's rules. This path decomposition is thus found to be:

$$r_{bs} = p_{bs} + p_{su}*p_{bu}$$

In fact, the above equality does hold: 0.53 = 1.36 + 0.72(-1.07) The calculation of the direct and indirect effects from regression coefficients implies that the model is correctly specified.

Another example is the relationship between E and B: $r_{be} = 0.002$. The low magnitude of the total correlation implies that E has no effect on B. This correlation can be decomposed into the direct effect of E on B (-1.07), the effect of E through J and S on B

$$(p_{ej}*p_{sj}*p_{bs})$$

and the effect of E through S on B: $(p_{es}*p_{bs})$.

Both of the latter two effects are logically spurious to the theoretical model because they go against the directions of some of the arrows. However, they do not violate Wright's rules so must be included in the decomposition of r_{be}. In fact, however, the values of the "spurious" relationships are required to balance the equation for the decomposition of the total correlation:

$$r_{be} = p_{be} + p_{ej}*p_{sj}*p_{bs} + p_{es}*p_{bs}$$

$$0.002 = -1.07 + (0.17)(0.39)(1.36) + (0.72)(1.36)$$
$$0.002 = -1.07 + 0.09 + 0.98$$

Thus the path analysis indicates that, for the given data, the model's assumed causal ordering between E and S and between J and E is incorrect. The true causal model would include reciprocal linkages in both cases.

Path analysis, although based on regression analysis, is superior in moving beyond estimation of direct effects. Path analysis allows one to examine the causal processes underlying the observed relationships and to estimate the relative importance of alternative paths of influence. It also provides a means for separating errors in equations from errors in variables and thus permits better diagnosis and theory building in research when compared to regression. Finally, both path and log-linear analysis more clearly specify alternative models of a process, and the requirements for distinguishing the best fitting model. Thus these processes are important tools wherever forecasting depends on accurate specification of a model.

REFERENCES

[1] Asher, H. B., Causal Modeling, (Beverly Hills, CA: Sage Publications, 1976).

[2] Bentler, P. M., "Theory and Implimentation of EQS, A Structural Equations Program" Manual for Program Version 1.0, Los Angeles, CA: BMDP Statistical Software, Inc., Mimeo, 1984.

[3] Bishop, Y. M., Feinberg, S. E., and Holland, P. W., Discreet Multivariate Analysis--Theory and Practice (Cambridge, MA: MIT Press, 1975).

[4] Fishbein, M. and Ajzen, I., Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. (Reading, Mass.: Addison-Wesley, 1975), p. 12.

[5] Hofacker, C., "Statistical Computing," Perspective, Vol. 7, no. 3-4 (April, 1983), pp. 18-27.

[6] LaBarbera, P., and MacLachlan, J., "Time-Compressed Speech in Radio Advertising," Journal of Marketing, Vol. 43 (1979), pp. 30-36.

[7] MacLachlan, J. and Siegel, M.H., "Reducing the Cost of TV Commercials by Use of Time Compressions," Journal of Marketing Research, Vol. 17 (1980), pp. 52-57.

[8] Sternthal, B., and Craig, C.S., "Humor in Advertising," Journal of Marketing, Vol. 37 (October, 1973), pp. 12-18.

[9] Wright, P., "The Cognitive Processes Mediating Acceptance of Advertising," Journal of Marketing Research, Vol. 10 (February, 1973), pp. 53-62.