Categorical Regression in Marketing

Kenneth L. Bernhardt, *Georgia State University*
Thomas C. Kinnear, *The University of Michigan*[1]

Market segmentation studies are currently analyzed by many sophisticated analysis techniques, among which are: regression, Multiple Classification Analysis (MCA), Automatic Interaction Detector (AID), cluster analysis, factor analysis, discriminant analysis, canonical correlation and multidimensional scaling. Frank, Massey and Wind [4] present a scheme to indicate when most of these procedures should be utilized in market segmentation analysis.

This paper presents a new multivariate analysis technique that has great potential for use in market segmentation analysis. Multivariate Nominal Scale Analysis (MNA) is a new data analysis technique developed by Frank M. Andrews and Robert C. Messenger at the University of Michigan's Institute for Social Research [1]. Essentially, it is an extension of the Multiple Classification Analysis (MCA) program [3] that has been utilized in a number of marketing studies [8, 9, 10]. MCA accepts nominally scaled independent variables and assumes an intervally scaled dependent variable. MNA accepts both nominal independent and dependent variables, in the context of an additive model.

The MNA Procedure

The primary objective of MNA is to allow the use of a regression type procedure with both nominal independent and dependent variables [1, 4–5]. A nominal criterion variable may be examined in the context of a set of nominal predictors with all the advantages of a multivariate regression procedure. In this context, MNA can provide useful information about:

(a) the effect of all predictors together on the dependent variable,
(b) the effect of a specific predictor on the dependent variable while holding constant all other predictors,
(c) the marginal effect of a specific predictor over and above all other predictors,
(d) the predicted dependent variable classification of any subject and
(e) the relationship between actual and predicted classificaton.

Related objectives include ease of input and ease of interpretation of the output. Specifically, the unique features of MNA include: (a) Dummy variables are automatically created by the MNA program thus saving the user considerable effort if he were to attempt a similar procedure himself. (b) The output of MNA is easy to understand. (c) A coefficient is provided for each category of every predictor, and the coefficients are just additions to or subtractions from the grand mean. In other dummy variable procedures, one arbitrary category of each predictor must be set to zero and other coefficients are expressed as deviations from this arbitrarily omitted category. (d) MNA coefficients are more interpretable and provide more complete information. (e) MNA also handles nonlinear relationships automatically. It is not necessary for the user to search for some best transformation with resulting difficulties in interpretation of the results.

Assumptions MNA makes a number of what are, essentially, regression based assumptions including: the absence of strong multicollinearity among the predictors, the absence of interaction (an additive model), plus the statistical assumptions that the expected value of the error term is zero, the variance of the error term is constant, error terms are uncorrelated, and that the predictors are not correlated with the error term The interaction problem may be overcome by building a pattern variable into the predictor set which would take into account both the additive and interaction effects of variables.

Limitations MNA procedes by forming dummy or 0–1 dependent variables. It is a well known property of regression that a 0–1 dependent variable results in the variance of the error term not being constant. It varies with the dependent variable. Therefore, one of the statistical assumptions of the system is violated and as a result the variance of the coefficients are no longer minimum—the results are

unbiased but inefficient which would be a major problem in statistical inference However, the second limitation of the method is that no statistical inference exists for MNA. The authors of MNA indicate that they are more interested in the strength of relationship (as measured by the size of the coefficient) than statistical significance [1:71] and that it is their experience with large data sets that important relationships are almost always statistically significant [1:37]. Another limitation relates to the loss of metric information for intervally scaled predictors. All predictors are treated as nominal. The positive side of this aspect is that nonlinear relationships are found automatically which would be hidden by a metric procedure.

Mathematical Overview The MNA procedure is a relatively new one, so most readers are likely to be unfamiliar with it. Therefore, the next section of this article presents a mathematical overview of how MNA works (see [1:21–30] for details).

MNA is based on the repeated application of least squares dummy variable regression [11]. Specifically, the set of original predictor variables (X_1, X_2, . . . , X_p) is transformed into a set of dummy predictor variables (x_1, x_2, . . . , x_{c_1}, . . . x_r) by treating every nonempty code of each predictor as a new dummy variable and by assigning a value of 1 when the code appears and 0 when it does not appear.

The resulting data set of dummy predictors has one linear dependency for each set of dummy predictors associated with an original predictor. These yield a singular matrix which prevents proper least squares estimation from being carried out. Therefore, the linear dependencies must be eliminated by omitting one dummy predictor from each set. This procedure yields a set of r = c-p independent predictors, where c = the total number of categories in the dependent variables and p = the number of predictors.

This procedure is completely analogous to multistate dummy in regular regression where an attribute has more than two levels. For example, the four category prediction variable geographic region would have the following codes available:

Northeast	100
Midwest	010
South	001
West	000

We note the removal of the linear dependency with the assignment of all zeros to the West dummy variable.

The dependent variable is also dummyized to form a set of G dummy dependent variables where G is the number of nonempty

dependent variable codes. Then, the set of r predictors is applied successively to the complete set of G dummy dependent variables, using the criterion of minimizing the error sum of squares, which forms the least squares criterion, given by:

(1) $$ESS_l = \Sigma w_k \quad (y_{kl} - \hat{y}_{kl})^2 \quad (l = 1, 2, \ldots G)$$

where

ESS_l = error sums of squares for the lth dummy dependent variable,

w_k = individual k's weight,

y_{kl} = individual k's score on the lth dummy dependent variable,

$\hat{y}_{kl}$ = individual k's predicted score for the lth dummy dependent variable

and where

$$\hat{y}_{kl} = B_{lo} + B_{l1}\, x_{k1} + B_{l2}\, X_{k2} + \ldots + B_{lr}\, x_{kr} \quad (l = 1, 2, \ldots G)$$

here,

x_{km} = the mth dummy predictor score for kth individual.

and

B = the regression coefficients.

Partial derivatives of the ESS's with respect to the B coefficients are then calculated. These partials are then set to zero, yielding the G normal equation sets [1:26].

In mathematical notation:

(2)

$$\begin{array}{ccc} \dfrac{\partial ESS l}{\partial B_{lo}} & & 0 \\ \dfrac{\partial ESS l}{\partial B_{l1}} & = & 0 \qquad l = (1, 2, \ldots G) \\ \bullet & & \bullet \\ \bullet & & \bullet \\ \bullet & & \bullet \\ \dfrac{\partial ESS l}{\partial B_{lr}} & & 0 \end{array}$$

yields the relevant normal equations.

Solutions of these G equations give the B values for the predictive equations and a set of forecasts of individual scores [$\hat{y}_{k1}, \hat{y}_{k2}, \ldots, \hat{y}_{kG}$]. This solution yields values expressed as deviations from the one dummy prediction that was omitted from each set. It is possible to present the predictive equations in a more easily understood form, while at the same time assigning values to the previously omitted codes. MNA does this by transforming the results to a form where coefficients are expressed as deviations from the mean of the *l*th dependent variable [1:27–8]. Here,

$$(3) \quad \hat{y}_l = \bar{y}_l + A_{l1} x_1 + A_{l2} x_2 + \ldots A_{lc} x_c \quad (l = 1, 2, \ldots, G)$$

where

$\bar{y}_l$ = the mean of the *l*th dependent variable,

and

A_{lm} = mth transformed dummy predictor regression coefficient for *l*th dummy dependent variable.

The A_{lm}'s are expressed as deviations from the grand means [$\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_G$]. This system yields forecasts that are identical to the previous system for all individuals and has coefficients attached to all categories of all independent variables. Coefficients are deviations from the grand mean $\bar{y}_l$, and not from the arbitrarily omitted category of each nominal independent variable as they would be if the equation was not transformed.

Statistics Generated by MNA

MNA generates both bivariate and multivariate statistics. Two bivariate statistics are produced to measure the strength of the relationship between the dependent variable and each predictor. The first is the oneway analysis of variance eta-squared statistic which is calculated for each dummy dependent variable and then summarized into a generalized eta-squared. Eta-squared measures the explained variance of each code and the generalized eta-squared statistic measures the explained variance across all codes, i.e., the ratio of explained sums of squares to total sums of squares.

A more useful bivariate statistic, the bivariate theta (Θ_y), is a relatively new statistic formulated by Messenger [6, 7] to measure the strength of association with correct placement in the dependent variable code as the criterion. Theta is defined as the proportion of the sample correctly classed when using a prediction-to-the-modal category strategy in each frequency distribution of each category of

the predictor variable. For example, Table 1 presents a set of data from the cross-tabulation of a 3 code dependent variable Y, with a 3 code independent variable X_i. The number in the cells are the number of people in the sample assigned to the cells. If we knew nothing about the effect of X_i on Y, our best prediction concerning Y would be Y_2, the modal category. That is, $\Theta_y = 400/1000 = .40$ and we will have correctly classified subjects 40 percent of the time. Knowledge of X_i allows for improved classifications. Specifically, if we know the subject is in X_1, the best guess is Y_1, etc. Then,

$$\Theta_Y/X_i = (300 + 300 + 200)/1000 = .80 \tag{4}$$

and we have correctly classified 80 percent of the subjects.[2]

Table 1. An Illustration of Bivariate Theta

		Y			
		1	2	3	Total
X_i	1	300	0	0	300
	2	50	300	50	400
	3	0	100	200	300
Total		350	400	250	1000

The multivariate statistics generated by MNA parallel the bivariate statistics described above. These are the generalized multiple R^2 and the multivariate theta statistic. The latter statistic is defined as the proportion correctly classed using a decision rule of predicting each individual as being in that dependent variable category having the *maximum forecast* value for that individual and written as:

$$\Theta_Y/X_1, X_2, \ldots, X_n, \text{ or } \Theta_M \tag{5}$$

It is the probability of placing a subject in the correct nominal category of the dependent variable, Y, given knowledge of the code values of the independent variables, $X_1, X_2, \ldots, X_n$, when using a prediction to the modal category strategy. It should be noted that multivariate theta (or bivariate theta in the case of only one predictor variable) could be applied to the classification matrix generated by discriminant analysis. This matrix is identical in concept to the predicted versus actual category comparison undertaken by MNA.

The MNA technique is a series of parallel MCA runs using each of the dummy variables in turn as the dependent variable. For each of the dependent variable codes, a predicted probability (Θ_m) of each subject being in that category is calculated. Each subject is predicted to fall in the dependent variable category for which he has the highest calculated probability. A comparison is made between the category each subject is predicted to be in and the category he is actually in, and the proportion correctly classified is then calculated.

Relation to Other Techniques

In their classification schemes for multivariate data analysis methods, Sheth [11] and Kinnear and Taylor [5] noted the absence of any technique to easily accomplish a regression type analysis with all nominal variables. However, procedures other than MNA do exist. First, one could dichotomize the dependent variable, code it 0–1 and do an analysis with MCA or dummy variable multiple regression. As noted previously, such a dependent variable gives inefficient estimators but can be corrected in this instance with a much more complex generalized least squares procedure. This whole option is limited to a dichotomous dependent variable.

A second option is available if the dependent variable has two or more categories. Discriminant analysis may be used with dummy independent variables. This procedure and MNA yield identical predictive results. However, the results of a dummy variable discriminant analysis are very difficult to interpret. The coefficients of a discriminant analysis are a new set of statistically derived predictors that are different from the predictors input by the researcher [2:474]. From a conceptual, and interpretative point of view these coefficients are extremely hard to explain. MNA has the advantage of providing coefficients showing the effects of the original predictors (see [2:12] for details). MNA also offers a much easier input procedure. The user of dummy variable discriminant analysis must create his own dummy variables which MNA does automatically. In dummy variable discriminant analysis some independent variable could be left continuous with the resulting retention of metric information. However, if nonlinear relationships exist between this variable and the dependent variable, the user must search for the proper transformation. MNA finds this nonlinearity automatically. Other less well known techniques are also possibilities. They are not discussed here, but are compared to MNA by Andrews and Messenger [1:36–50]

An Illustration[3]

An example should help illustrate the type of analysis MNA can do for marketers. Suppose one is interested in studying the characteristics of consumers who purchase particular types of brands in the detergent market. On the basis of cluster analysis of time series purchase data and of attitude data, five consumer typologies for the detergent market were developed. These are: (1) nonphosphate brand users (2) major brand users (3) private brand users (4) cents off brand users (5) bonus brand users (towels inside, etc.). These five categories of buyers are the dependent variable in our MNA analysis. The management of this company found this type of scheme as a useful way to classify their market They were prepared to develop strategies against these typologies and wanted to know what consumer characteristics were related to each segment. The company had five independent variables of interest, all of which had been categorized. These variables are: (1) occupation of head of household (2) stage of life cycle (3) level of self confidence of head of household (4) family income, and (5) education. We note that occupation is a truly nominal variable, whereas life cycle, self confidence, education are probably ordinal and income could be treated as interval if left uncategorized. The point of using MNA with this data is that the nominal variable can be used and possible nonmonotonisity or nonlinearity in the other variables can be found automatically, without a troublesome and complex search for the best transformations. The data used in this study were collected by means of a diary panel located in ten cities in the United States.

Table 2 presents the MNA results for the detergent user types as dependent variable and the available predictors (N = 1970). The first finding to note is the overall percentage distribution of respondents over the five categories of the dependent variable. We note that 10.4 percent were classified as nonphosphate brand users, 26.4 percent as major brand users, etc. The modal category was major brand users yielding Θ_y equal to .264. If we knew nothing about the characteristics of the respondents, we could predict the modal category and be right 26.4 percent of the time. The independent variables, X_1, $X_2, \ldots, X_p$ serve to increase our ability to predict above this base level.

The strength of relationship between the set of independent variables and the dependent variable can be correlated in three ways. First, the generalized R^2 equals .16, which indicates that approximately 16 percent of the variance in the dependent variable is explained. Second, we can examine the category specific R^2's. This

Table 2: MNA Results

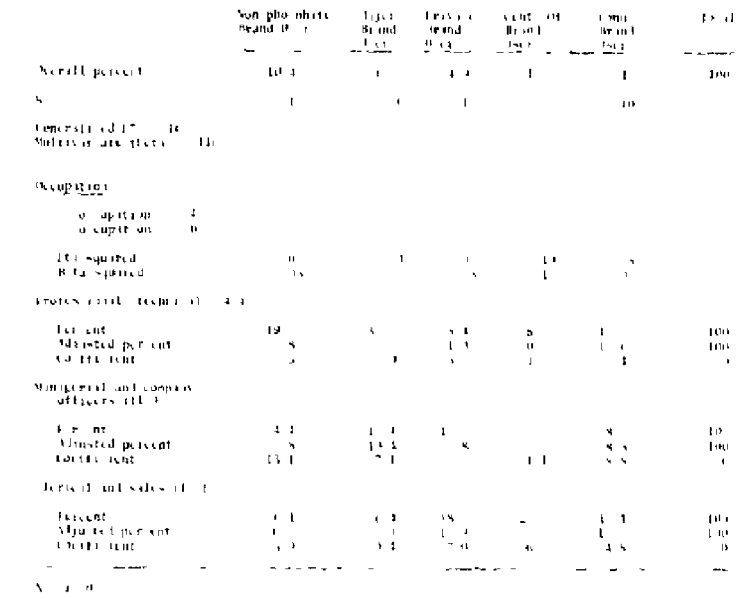

Table 2: MNA Results Continued

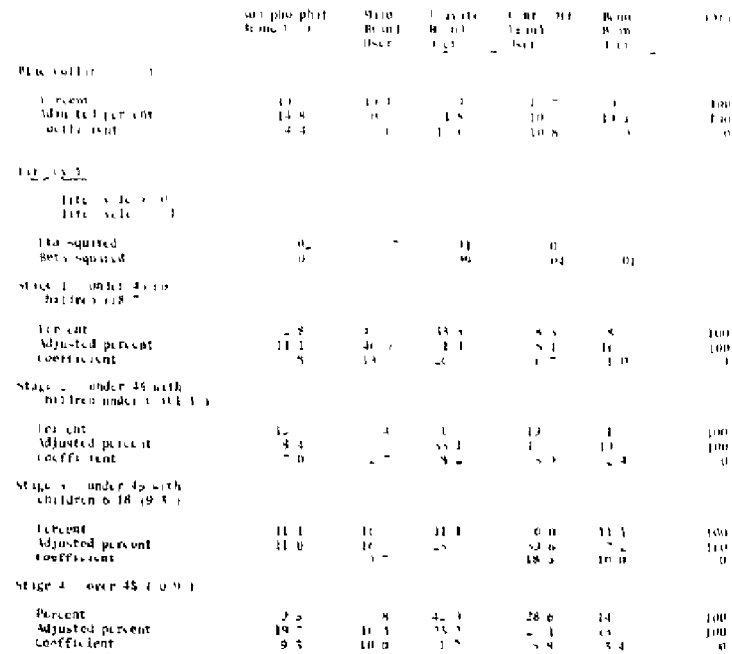

Table 2: MNA Results Continued

	Non phosphate Brand User	Major Brand User	Private Brand User	Cents Off Brand User	Bonus Brand User	Total
Self confidence level						
[illegible] self confidence [illegible]						
[illegible] self confidence [illegible]						
Eta squared	.00	[illegible]	[illegible]	[illegible]	[illegible]	
Beta squared	.00	[illegible]	[illegible]	.00	[illegible]	
Low self confidence [illegible]						
Percent	[illegible]	[illegible]	[illegible]	[illegible]	18	[illegible]
Adjusted percent	[illegible]	[illegible]	[illegible]	[illegible]	18.2	[illegible]
Coefficient	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Medium self confidence [illegible]						
Percent	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Adjusted percent	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	100
Coefficient	[illegible]	[illegible]	[illegible]	[illegible]		[illegible]
High self confidence [illegible]						
Percent	[illegible]	[illegible]	[illegible]		[illegible]	[illegible]
Adjusted percent	[illegible]		[illegible]		[illegible]	[illegible]
Coefficient	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Family Income						
[illegible] family income [illegible]						
[illegible] family income [illegible]						
Eta squared	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
Beta squared	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
Less than $10,000 [illegible]						
Percent	[illegible]	[illegible]	[illegible]	[illegible]		100
Adjusted percent		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Coefficient		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Table 2: MNA Results Continued

	Non phosphate Brand User	Major Brand User	Private Brand User	Cents Off Brand User	Bonus Brand User	Total
$10,000 - $11,999 [illegible]						
Percent	11.1	[illegible]	[illegible]	13.[illegible]	[illegible]	[illegible]
Adjusted percent	11.9	34.[illegible]	[illegible]	15.5	[illegible]	100
Coefficient	1.[illegible]	[illegible]	1		4	0
$1[illegible],000 - $14,999 (39.4%)						
Percent	[illegible]	[illegible]	15.[illegible]	[illegible]	[illegible]	[illegible]
Adjusted percent	13.1	3	15	[illegible]	13.[illegible]	[illegible]
Coefficient	[illegible]	10.8	6.4	[illegible]	[illegible].8	0
$15,000 and over [illegible]						
Percent	[illegible]	[illegible]	[illegible]	[illegible]	1	100
Adjusted percent	[illegible].6	44.6	31.5	1	9	100
Coefficient	[illegible].8	18	[illegible]	9.[illegible]	8.[illegible]	0
Education						
[illegible] education [illegible]						
[illegible] education [illegible]						
Eta squared	.04	[illegible]	[illegible]	.04	.01	
Beta squared	.07	.01	[illegible]	.01	.00	
High School or less (48.[illegible])						
Percent	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	100
Adjusted percent	[illegible].0	[illegible]	[illegible]	[illegible]	18	100
Coefficient	[illegible].4	[illegible].8	[illegible]	5.[illegible]	1.[illegible]	0
Vocational School or some college (23.[illegible]%)						
Percent	6.7	4.4	[illegible]	28.9	17.8	100
Adjusted percent	8.0	6.[illegible]	19.1	[illegible]	18.3	100
Coefficient	4	[illegible]	3.8	0.[illegible]	1.2	0
College graduate ([illegible].5%)						
Percent	0.0	37.4	23.6	9.1	10.[illegible]	100
Adjusted percent	23.1	5	25.0	6.[illegible]	13.[illegible]	100
Coefficient	1[illegible]	6.1		15	3	0

examination indicates that the cents off brand user category was best predicted by the independent variable ($R^2 = .22$) and that the bonus brand user category was the least well predicted ($R^2 = .10$).

Another way to examine the overall relationship between the dependent and independent variables is to note the multivariate theta value. This value is the percentage of respondents that could be correctly classified with knowledge of the independent variables. Multivariate theta, $\Theta_y/X_1, X_2, \ldots, X_p$, equals .440. By comparing this value to Θ_y, .264, we note that these independent variables allow us to increase our correct prediction level by 17.6 percentage points (44.0 − 26.4).[4]

MNA also produces a number of predictor specific calculations and statistics. The generalized eta-squared and the bivariate theta are utilized to indicate the strength of the bivariate association between an independent variable and the dependent variable. For example, for occupation, η^2 is .04 and Θ is .36, indicating that occupation explains .04 percent of the variance and correctly classifies 36 percent of the sample. MNA also gives category-specific eta-squareds and beta-squareds for each predictor. The latter statistic is an approximation of the ability of a predictor to explain variance of each category of the dependent variable while holding constant all other predictor variables.

The details of how each category of an independent variable is associated with each category of the dependent variable are also available. MNA produces three sets of figures for each category of each independent variable to show these relationships. The "percent" figures give the bivariate percentage distribution of respondents across the categories of the dependent variable By comparing rows of percents we can see, for example, that professional and technical respondents are more likely to use nonphosphate detergents than the other occupational categories (19.2 percent versus 4.4 for the managerial category, 6.1 for the clerical and sales and 10.3 for the blue collar category). In a similar fashion we note that blue collar respondents are more likely to purchase both bonus brands and private brands than other occupational classifications. Other independent variable categories can be examined in the same way.

The "coefficient" figures give the effect of being a specific category of a predictor variable on the likelihood of a respondent being in each category of the dependent variable. These coefficients are the heart of the multivariate analysis. An individual's predicted probability of being in a specific category of the dependent variable

is equal to "overall percent" for that category plus the coefficients across all predictor categories relevant to that respondent and that dependent variable category. The coefficients can be interpreted as the amount of increase or decrease in likelihood of dependent variable category membership after holding constant all other predictor variables. The "adjusted percent" figures are formed by adding the coefficient for that category of the dependent variable to the relevant "overall percent." The result is the percentage distribution of respondents across categories of the dependent variable after allowance has been made for the effects of other predictors.

Examination of the results presented in Table 2 allows the marketer to form a portrait of those using particular types of detergent brands. For example, major brand users can be described as tending to have the following characteristics:

	Percent	*Adjusted Percent*
1. Employed in professional and technical occupations	36.2	32.3
2. Under 45 with no children	47.2	46.0
3. Medium level of self confidence	34.6	31.5
4. Earnings of from $12,000 to	36.1	37.2
5. College graduate	36.4	32.5

It is possible to select a specific consumer profile and determine the likelihood that people with those characteristics are major brand users. Specifically, those having the above characteristics have a predicted probability of using a major detergent brand of .714, up from the base probability of .264. This predicted probability is calculated by adding the coefficients in these categories to the overall percent using major brands.[5] Similar descriptions and calculation can be undertaken for other detergent user types.

Table 3 presents a classification matrix that compares actual classification on the dependent variable with the categories predicted by MNA. The diagonal elements indicate the proportion correctly classified for each dependent variable category. Table 3 also shows the nature of the misclassifications that did occur. We note that for major brand users 60.8 percent were correctly predicted as being major brand users, none were incorrectly predicted as nonphosphate brand users, 15.7 were incorrectly predicted to be private brand users, etc. The users of misclassifications for the other

Table 3: Classification Matrix

Actual	Predicted				
	Non-phosphate Brand User	Major Brand User	Private Brand User	Cents Off Brand User	Bonus Brand User
Non-phosphate brand user (N = 200)	20.0%	35 0%	10 0%	20 0%	15.0%
Major brand user (N = 510)	0.0	60 8	15.7	19 6	3 9
Private brand user (N = 480)	2 1	22 9	45 8	14 6	14 6
Cents off brand user (N = 410)	0 0	22 0	19 5	48 8	9 8
Bonus brand user (N = 330)	0.0	24 2	30 3	21 2	24.2

detergent segments are also available from Table 3. The nature of the misclassifications that do occur gives the marketer a view of the extent to which segments overlap in terms of characteristics. Table 3 also allows the user to determine the degree to which he is able to predict membership in particular segments. In this illustration, major brand users were the most successfully predicted, and non-phosphate users were the least successfully predicted.

Potential Uses

Differences found in the socioeconomic profile of each consumer typology could be used to help select appropriate media vehicles, design packages, select relevant actors for commercials, etc. In addition to MNA's usefulness in segmentation analysis and planning, it has great potential for some of the purposes described in the following examples.

1. The sales force for a company could be evaluated, with each salesman being put in a performance category such as well above average, slightly above average, average, etc. MNA could then be used to examine a number of determinants in an effort to describe which variables are associated with each of the categories of salesman performance.

2. MNA could be used to identify and describe geographical territories which offer the greatest profit opportunities The company's market areas could be defined into certain categories, similar to those described in the sales force analysis example above, with MNA then used to identify the variables most important in describing the better territories.
3. MNA could be utilized to assist in media planning. A number of media choices could be analyzed using a variety of demographic, psychographic, and product usage variables. The result would be useful in comparing various media audiences and a company's target market.
4. The new product development process is another area where MNA has potential usefulness. After several different prototypes had been developed, the importance of a variety of independent variables could be examined utilizing MNA to determine which variables were associated with preference for each of the prototypes. This analysis would indicate which segments of the market would offer the highest potential for each of the prototypes.
5. Demographic and personal characteristics associated with users of different brands could be analyzed using MNA. This would enable the marketer to examine user profiles for each of a number of competitive brands. Similarly, users of a product could be divided into several usage categories, varying from very heavy user to very light user, with MNA being used to determine the profiles of those in each category.
6. The analysis of credit risks is another potential area for the use of MNA Multiple discriminant analysis is often used to evaluate credit risks now, but MNA offers the additional advantage of being able to include nominal level independent variables in the analysis, thus offering the potential for better description of the various risks categories.

Summary and Conclusions

Since MNA is capable of handling nominal variables as both predictors and criterion variables, it is a useful technique for marketing researchers. Now categorical dependent variables such as brand choice, consumer typologies, switching patterns, media choices, etc., can be handled in a regression type analysis with categorical independent variables, such as region of country, occupation, sex, race, exposure to marketing programs, stage of life cycle, etc. The potential uses of the MNA technique in marketing seem extensive.

Footnotes

[1]The authors wish to thank the *JBR* reviewers and Editor for their most helpful comments

[2]Θ_{Y/X_I} is really just a more intuitively appealing form of the Goodman and Kruskal Lambda statistic, λ_I, which is defined as the proportion of reduction in error given predictor X_I's codes

$$\begin{aligned}\lambda_I &= (\Theta_{Y/X_I} - \Theta_Y) / (1 - \Theta_Y)\\ &= (.80 - .40) / (1 - .40)\\ &= .67\end{aligned}$$

In this example error has been reduced from .6 to .2 by the knowledge of the association between X_I and Y. It can be seen that it represents a .67 reduction in error as calculated by λ above.

[3] The data used in this illustration are the disguised results of a real study. The essence of the type of design used and nature of results remains the same.

[4]We note here that $\lambda = \dfrac{.440 - .264}{1 - .264} = .24$

[5]

Base percent		26.4
Category Effects		
Professional, technical	5.9	
Under 45, no children	19.6	
Medium self confidence	5.1	
$12,000 to $14,999	10.8	
College graduate	6.1	
		47.5
		73.9

Base probability = .264
Predicted probability for respondents having these characteristics = .739
Increase in probability due to knowledge of predictors = .739 − .264 = .475

Reference

1. Andrews, F.M. and Messenger, R.C. *Multivariate Nominal Scale Analysis.* Ann Arbor, Michigan: Institute for Social Research, University of Michigan, 1973.
2. Andrews, F.M., Morgan, J.N., and Messenger, R.C. "Comments on Reviews by Jagdish N. Sheth of MNA and THAID," *Journal of Marketing Research* 11 (November 1974): 473–475.
3. Andrews, F.M. and Sonquist, J.A. *The Multiple Classification Analysis Program.* Ann Arbor, Michigan: Institute for Social Research, University of Michigan, 1967.
4. Frank, R.E., Massy, W.F., and Wind, Y. *Market Segmentation.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1972, 139–69.
5. Kinnear, T.C. and Taylor, J.R. "Multivariate Methods in Marketing Research: A Further Attempt at Classification." *Journal of Marketing* 34 (October 1971): 56–9.
6. Messenger, R.C. *Theta User's Guide.* Unpublished manuscript, Ann Arbor, Michigan: Institute for Social Research, University of Michigan, 1971.
7. Messenger, R.C. and Mandell, L.M. "Modal Search Technique for Predictive Nominal Scale Multivariate Analysis." *Journal of the American Statistical Association* 67 (December 1972).

8 Newman, J.W and Staelin, R "Multivariate Analysis of Differences in Buyer Decision Time," *Journal of Marketing Research* 8 (May 1971), 192–8
9. Newman, J W "Prepurchase Information Seeking for New Cars and Major Household Appliances" *Journal of Marketing Research* 9 (August 1972) 249–57
10 Peters, W H. "Using MCA to Segment New-Car Markets" *Journal of Marketing Research* 7 (August 1970) 360–3
11 Sheth, J N "The Multivariate Revolution in Marketing Research" *Journal of Marketing* 34 (January 1971) 13–9
12 Sheth, J.N "Book Review of *Multivariate Nominal Scale Analysis* and *THAID A Sequential Analysis Program for The Analysis of Nominal Scale Dependent Variable*" *Journal of Marketing Research* 11 (May 1974) 228–232
13. Suits, D B. "Use of Dummy Variables in Regression Equation" *Journal of American Statistical Association* 52 (1957), 548–51