

**MKT 791--Research II  
Spring 1996**

**SESSION 14**

**NONMETRIC DATA ANALYSIS POTPOURRI:  
The linear probability model, Cluster Analysis,  
and Multidimensional Scaling**

**READINGS**

**The Linear Probability Model**

Hair, Joseph F., Jr., Rolph E. Anderson, Ronald L. Tatham, and William C. Black (1995),  
*Multivariate Data Analysis with Readings*, 4th ed., Englewood Cliffs, NJ: Prentice-Hall.

◆ "Regression With a Binary Dependent Variable," pp. 129-133.

Aldrich, John H. and Forrest D. Nelson (1986), "Logit and Probit Models for Multivariate  
Analysis with Qualitative Dependent Variables,' in *New Tools for Social Scientists: Advances  
and Applications in Research Methods*, eds. William D. Berry and Michael S. Lewis-Beck, pp.  
115-155.

**Cluster Analysis**

Hair, Joseph F., Jr., Rolph E. Anderson, Ronald L. Tatham, and William C. Black (1995),  
*Multivariate Data Analysis with Readings*, 4th ed., Englewood Cliffs, NJ: Prentice-Hall.

◆ Chapter 8, "Cluster Analysis"

Aldenderfer, Mark S. and Roger K. Blashfield (1984), *Cluster Analysis*, Newbury Park, CA.:  
Sage, QASS #44.

Punj, G. and David W. Stewart (1983), "Cluster Analysis in Marketing Research: Review and  
Suggestions for Application," *Journal of Marketing Research*, 20 (May), 134-148.

Funkhouser G. Ray (1983), "A Note on the Reliability of Certain Clustering Algorithms," *Journal  
of Marketing Research*, 20 (February), 99-102.

**Multidimensional Scaling**

Hair, Joseph F., Jr., Rolph E. Anderson, Ronald L. Tatham, and William C. Black (1995),  
*Multivariate Data Analysis with Readings*, 4th ed., Englewood Cliffs, NJ: Prentice-Hall.

◆ Chapter 9, "Multidimensional Scaling"

**SAS RESOURCES**

These SAS PROCs can analyze categorical data:

CATMOD--Fits a wide range of linear models to functions of categorical data.

CORRESP--Performs simple and multiple correspondence analyses.

FREQ--Frequencies, contingency tables and measures of association (e.g.,  $\chi^2$ )

LOGISTIC--Fits linear logistic regression models for binary or ordinal response data with maximum-likelihood methods.

PROBIT--Fits a PROBIT model to binary or ordinal data.

## YOUR TURN

First, use PROC FREQ's TABLES feature and analyze an  $m$  by  $n$  contingency table. Interpret the results.

Second, use PROC LOGISTIC to do a logit analysis on a binary dependent variable. Include two or more predictors.

Third, run a cluster analysis (PROC CLUSTER). Experiment with different methods to see how the results differ. Use PROC TREE to obtain a graphical representation of the cluster solutions.

If feasible, perform all three analyses on the HATCO data *and* your data. Bring all analyses to class.

## ADDITIONAL RESOURCES

### Nonparametrics

Books I've found very useful:

Hollander, Myles and Douglas A. Wolfe (1973), *Nonparametric Statistical Methods*, New York: John Wiley and Sons.

Siegel, Sidney (1956), *Nonparametric Statistics for the Behavioral Sciences*, New York: McGraw-Hill.

Siegel, Sidney and N. John Castellan, Jr. (1988), *Nonparametric Statistics for the Behavioral Sciences*, 2nd ed., New York: McGraw-Hill

The Siegel book is a classic--the basic reference in the field. It provides superb coverage of the most commonly used nonparametric tests. I reproduced on the reverse side Siegel's categorization of the tests discussed in his book. The Hollander and Wolfe book affords a nice complement to Siegel. They discuss less common procedures that can be lifesavers in certain situations.

### Categorical Data Analysis

Agresti, Alan (1990), *Categorical Data Analysis*, New York: Wiley.

### Cluster Analysis

- DeSarbo, Wayne S. and Geert De Soete (1984), "On the Use of Hierarchical Clustering for the Analysis of Nonsymmetric Proximities," *Journal of Consumer Research*, 11 (June), 601-610.
- Fader, Peter S. and Leonard M. Lodish (1990), "A Cross-Category Analysis of Category Structure and Promotional Activity for Grocery Products," *Journal of Marketing*, 54 (October), 52-65.
- Klastorin, T. D. (1983), "Assessing Cluster Analysis Results," *Journal of Marketing Research*, 20 (February), 92-98.
- Singh, Jagdip (1990), "A Typology of Consumer Dissatisfaction Response Styles," *Journal of Retailing*, 66 (Spring), 57-99.
- Wedel, Michel Steenkamp, E. M. Jan-Benedict (1991), "A Clusterwise Regression Method for Simultaneous Fuzzy Market Structuring and Benefit Segmentation," *Journal of Marketing Research*, 28 (November), 385-96.

## **MDS**

- Denison, Daniel R., and Claes Fornell (1990), "Modeling Distance Structures in Consumer Research: Scale Versus Order in Validity Assessment," *Journal of Consumer Research*, 16 (March), 479-89
- DeSarbo, Wayne S. and Donna L. Hoffman (1987), "Constructing MDS Joint Spaces from Binary Choice Data: a Multidimensional Unfolding Threshold Model for Marketing Research," *Journal of Marketing Research*, 24 (February), 40-54.
- Malhotra, Naresh K. (1987), "Validity and Structural Reliability of Multidimensional Scaling," *Journal of Marketing Research*, 24 (May), 164-173.
- Malhotra, Naresh K., Arun K. Jain, and Christian Pinson (1988), "The Robustness of MDS Configurations in the Case of Incomplete Data," *Journal of Marketing Research*, 25 (February), 95-102.

## **Correspondence Analysis:**

- Carroll, J. Douglas, Paul E. Green, and Catherine M. Schaffer (1986), "Interpoint Distance Comparisons in Correspondence Analysis," *Journal of Marketing Research*, 23 (August), 271-280.
- Carroll, J. Douglas, Paul E. Green, and Catherine M. Schaffer (1987), "Comparing Interpoint Distances in Correspondence Analysis: A Clarification," *Journal of Marketing Research*, 24 (November), 445-450.
- Carroll, J. Douglas, Paul E. Green, and Catherine M. Schaffer (1987), "Reply to Greenacre's Commentary on the Carroll-Green-Schaffer Scaling of Two-Way Correspondence Analysis Solutions," *Journal of Marketing Research*, 26 (August), 366-368.
- Carroll, J. Douglas and Paul E. Green (1988), "An INDSCAL-Based Approach to Multiple Correspondence Analysis," *Journal of Marketing Research*, 25 (November), 193-203.

Greenacre, Michael J. (1989), "The Carroll-Green-Schaffer Scaling in Correspondence Analysis: A Theoretical and Empirical Appraisal," *Journal of Marketing Research*, 26 (August), 358-265.

Hoffman, Donna L. and Geore R. Franke (1986), "Correspondence Analysis: Graphical Representation of Categorical Data in Marketing Research," *Journal of Marketing Research*, 23 (August), 213-227.

a:\spring94\readings\n\_sess14.s94