Understand the six-step approach to multivariate model building. The six-step model-building process provides a framework for developing, interpreting, and validating any multivariate analysis.

1. Define the research problem, objectives, and multivariate technique to be used.
2. Develop the analysis plan.
3. Evaluate the assumptions.
4. Estimate the multivariate model and evaluate fit.
5. Interpret the variates.
6. Validate the multivariate model.

This chapter introduced the exciting, challenging topic of multivariate data analysis. The following chapters discuss each of the techniques in sufficient detail to enable the novice researcher to understand what a particular technique can achieve, when and how it should be applied, and how the results of its application are to be interpreted.

Questions

1. In your own words, define multivariate analysis.
2. Name the most important factors contributing to the increased application of techniques for multivariate data analysis in the last decade.
3. List and describe the multivariate data analysis techniques described in this chapter. Cite examples for which each technique is appropriate.
4. Explain why and how the various multivariate methods can be viewed as a family of techniques.
5. Why is knowledge of measurement scales important to an understanding of multivariate data analysis?
6. What are the differences between statistical and practical significance? Is one a prerequisite for the other?
7. What are the implications of low statistical power? How can the power be improved if it is deemed too low?
8. Detail the model-building approach to multivariate analysis, focusing on the major issues at each step.

Suggested Readings

A list of suggested readings illustrating issues and applications of multivariate techniques in general is available on the Web at www.prenhall.com/hair.

References

tests, from graphical portrayals to empirical measures, is available to determine whether assumptions are met. Researchers are faced with what may seem to be an impossible task: satisfy all of these statistical assumptions or risk a biased and flawed analysis. These statistical assumptions are important, but judgment must be used in how to interpret the tests for each assumption and when to apply remedies. Even analyses with small sample sizes can withstand small, but significant, departures from normality. What is more important for the researcher is to understand the implications of each assumption with regard to the technique of interest, striking a balance between the need to satisfy the assumptions versus the robustness of the technique and research context.

**Determine the best method of data transformation given a specific problem.** When the statistical assumptions are not met, it is not necessarily a "fatal" problem that prevents further analysis. Instead, the researcher may be able to apply any number of transformations to the data in question that will solve the problem and enable the assumptions to be met. Data transformations provide a means of modifying variables for one of two reasons: (1) to correct violations of the statistical assumptions underlying the multivariate techniques, or (2) to improve the relationship (correlation) between variables. Most of the transformations involve modifying one or more variables (e.g., compute the square root, logarithm, or inverse) and then using the transformed value in the analysis. It should be noted that the underlying data are still intact, just their distributional character is changed so as to meet the necessary statistical assumptions.

**Understand how to incorporate nonmetric variables as metric variables.** An important consideration in choosing and applying the correct multivariate technique is the measurement properties of the dependent and independent variables. Some of the techniques, such as discriminant analysis or multivariate analysis of variance, specifically require nonmetric data as dependent or independent variables. In many instances, the multivariate methods require that metric variables be used. Yet nonmetric variables are often of considerable interest to the researcher in a particular analysis. A method is available to represent a nonmetric variable with a set of dichotomous variables, known as dummy variables, so that it may be included in many of the analyses requiring only metric variables. A dummy variable is a dichotomous variable that has been converted to a metric distribution and represents one category of a nonmetric independent variable.

Considerable time and effort can be expended in these activities, but the prudent researcher wisely invests the necessary resources to thoroughly examine the data to ensure that the multivariate methods are applied in appropriate situations and to assist in a more thorough and insightful interpretation of the results.

**Questions**

1. Explain how graphical methods can complement the empirical measures when examining data.
2. List potential underlying causes of outliers. Be sure to include attributions to both the respondent and the researcher.
3. Discuss why outliers might be classified as beneficial and as problematic.
4. Distinguish between data that are missing at random (MAR) and missing completely at random (MCAR). Explain how each type affects the analysis of missing data.
5. Describe the conditions under which a researcher would delete a case with missing data versus the conditions under which a researcher would use an imputation method.
6. Evaluate the following statement: In order to run most multivariate analyses, it is not necessary to meet all the assumptions of normality, linearity, homoscedasticity, and independence.
7. Discuss the following statement: Multivariate analyses can be run on any data set, as long as the sample size is adequate.