Chapter 2

Analysis of Groups: Multivariate Analysis of Variance

Section 2.1

Introduction to Multivariate Analysis of Variance

Objectives

- Understand when to use multivariate analysis of variance (MANOVA).
- Review concepts and definitions related to matrices and vectors in multivariate statistics.
- Understand the assumptions of MANOVA.
- Recognize multivariate test statistics and how they are calculated.

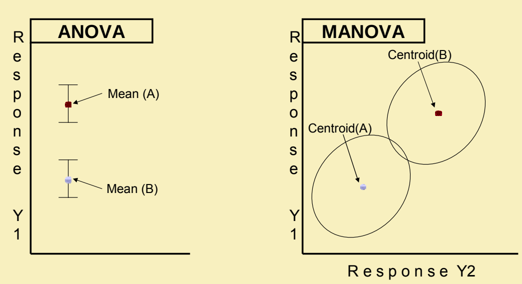
3

What Is MANOVA?

- A linear model.
- A statistical method for identifying group differences on a set of dependent variables.
- A method that incorporates the interrelationships among dependent variables in examining group differences.

4

ANOVA versus MANOVA



5

Statistical Advantages of MANOVA

Compared to ANOVA with multiple dependent variables, MANOVA

- reduces overall type-I error rate
- accounts for important information such as correlation among the dependent variables
- accounts for joint effects in the responses that would be missed otherwise in univariate tests (MANOVA increases power)
- allows you to examine multiple scores to screen for overall differences without combining scores into a single composite.

(Stevens 1996)

6

The MANOVA Model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{E}$$

where

Y the $n \times p$ matrix of p dependent variables for n observations

X the model matrix

β the parameter matrix

E the error (residual) matrix.

7

Statistics Used in Linear Models

Recall the *F*-statistic used to test hypotheses in linear models and the coefficient of determination, R^2, used to identify the proportion of variance accounted for by terms in the linear model:

$$F = \frac{SS_B / df_B}{SS_W / df_W} \text{ or } \frac{\text{Variance}_{\text{Between-groups}}}{\text{Variance}_{\text{Within-groups}}}$$

$$R^2 = \frac{SS_B}{SS_T} \text{ or } \frac{\text{Between-groups variability}}{\text{Total variability}}$$

In multivariate models, you will see statistics that are multivariate generalizations of familiar univariate statistics.

8

Partitioning Variances for MANOVA

H the matrix of hypothesized effects. This is analogous to the SS_B in univariate ANOVA.

E the error, or residual, matrix. This is analogous to the SS_W in univariate ANOVA.

T the total variability matrix. This is analogous to the corrected total sum of squares in ANOVA.

9

Test Statistics for MANOVA

$$\text{Wilks' Lambda} = \frac{|\mathbf{E}|}{|\mathbf{T}|}$$

$$\text{Pillai's Trace} = trace(\mathbf{HT}^{-1})$$

$$\text{Hotelling-Lawley Trace} = trace(\mathbf{E}^{-1}\mathbf{H})$$

$$\text{Roy's Maximum Root} = \lambda, \text{ or}$$

$$\text{largest eigenvalue of } \mathbf{E}^{-1}\mathbf{H}$$

10

Matrices for MANOVA Test Statistics

H the matrix of hypothesized effects. This is analogous to the SS_B in univariate ANOVA.

E the error, or residual, matrix. This is analogous to the SS_W in univariate ANOVA.

T the total variability matrix. This is analogous to the corrected total sum of squares in ANOVA.

11

Test Statistics for MANOVA

$$\text{Wilks' Lambda} = \frac{|\mathbf{E}|}{|\mathbf{T}|}$$

or

$$\prod_{i=1}^{j} \frac{1}{1+\lambda_i}$$

$$\text{where } j = \min(p,\ k - 1)$$

(where λ_i is the i^{th} eigenvalue of $\mathbf{E}^{-1}\mathbf{H}$)

12

Test Statistics for MANOVA

$$\text{Pillai's Trace} = tr(\mathbf{HT}^{-1})$$

or

$$\sum_{i=1}^{j} \frac{\lambda_i}{1+\lambda_i}$$

$$\text{where } j = \min(p,\ k-1)$$

(where λ_i is the i^{th} eigenvalue of $\mathbf{E}^{-1}\mathbf{H}$)

13

Test Statistics for MANOVA

$$\text{Hotelling-Lawley Trace} = tr(\mathbf{E}^{-1}\mathbf{H})$$

or

$$\sum_{i=1}^{j} \lambda_i$$

$$\text{where } j = \min(p,\ k-1)$$

(where λ_i is the i^{th} eigenvalue of $\mathbf{E}^{-1}\mathbf{H}$)

14

Test Statistics for MANOVA

Roy's Maximum Root = λ, or

largest eigenvalue of $\mathbf{E}^{-1}\mathbf{H}$

15

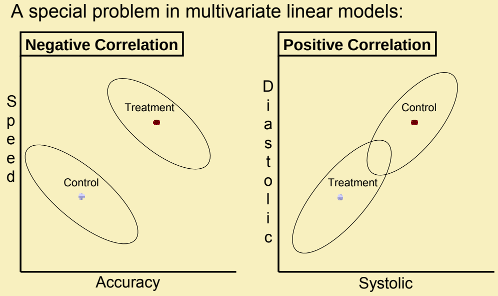
Assumptions of MANOVA

- Random sample
- Independent observations
- Multivariate normality
- Homogeneity of covariance matrices

16

Collinear Dependent Variables

A special problem in multivariate linear models:



Typing Skills **Blood Pressure**

Bray and Maxwell 1985

17

Sample Size

Most multivariate analyses are large-sample procedures.

Rules of thumb for **minimum** sample size:

- greater of 100 observations or 5 times the number of parameters

or

- 20+ observations per group.

For small-effects sizes and large variances, larger samples are necessary for adequate statistical power.

18