

MANOVA and the Multivariate GLM

Here we generalize the notation we learned before to the case of several variables within each group.

MANOVA – An Intuitive Introduction

FIGURE 11.12
Marginal distributions of Y_1 and Y_2
and joint distribution of Y_1, Y_2

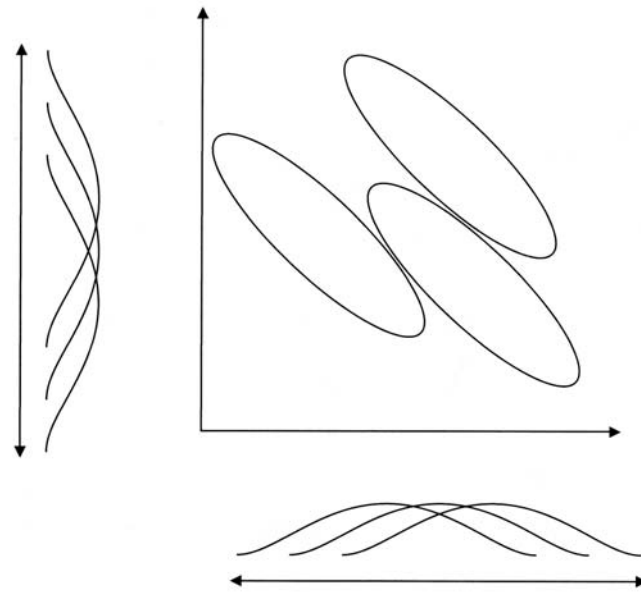
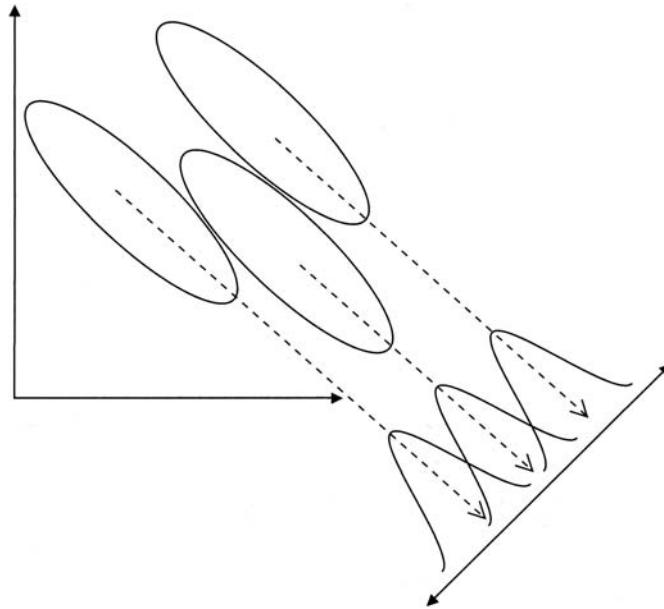


FIGURE 11.13
Linear combination of Y_1 and Y_2 reduces within-group variance relative to differences in group means



The Multivariate GLM

Recall that our univariate GLM is

$$\mathbf{x} = \mathbf{D}_1 \boldsymbol{\mu}^* + \boldsymbol{\varepsilon} \quad (1.1)$$

The multivariate version allows several columns of \mathbf{x} and several columns in $\boldsymbol{\mu}$ and $\boldsymbol{\varepsilon}$. We will assume at the outset that our design matrix is of full rank, so our new notation is

$${}_N \mathbf{X}_v = \mathbf{D}_1 \mathbf{U} + \mathbf{E} \quad (1.2)$$

The multivariate general linear hypothesis is of the form

$$H_0 : {}_p\mathbf{H}'_q \mathbf{U}_v \mathbf{M}_u = \mathbf{0} \quad (1.3)$$

Hence, we can express hypotheses about the various factors in the design, but also, via the \mathbf{M} matrix, express hypotheses about combinations of the variables in the design. In factorial MANOVA, we will confine ourselves to a simpler special case where \mathbf{M} is an identity matrix, so we are hypothesizing the equivalent of univariate

ANOVA effects, except that we are testing them on several variables at once.

This is a multivariate hypothesis. But we could reduce it to a univariate hypothesis by simply taking a linear combination of the population means.

$$H_0 : \mathbf{H}'\mathbf{U}\mathbf{M}\mathbf{a} = \mathbf{0} \quad (1.4)$$

The multivariate hypothesis of Equation (1.3) is true if and only if the univariate hypothesis of Equation (1.4) is true for all choices of \mathbf{a} . Since

\mathbf{Ma} is in effect also taking a linear combination of the columns of \mathbf{X} , we can see that the test statistic for this *univariate* hypothesis is just the same as before, i.e., defining $\mathbf{x}_a = \mathbf{XMa}$, we have

$$F(\mathbf{a}) = \frac{\mathbf{x}'_a \mathbf{D}_1 (\mathbf{D}'_1 \mathbf{D}_1)^{-1} \mathbf{H} \left[\mathbf{H}' (\mathbf{D}'_1 \mathbf{D}_1)^{-1} \mathbf{H} \right]^{-1} \mathbf{H}' (\mathbf{D}'_1 \mathbf{D}_1)^{-1} \mathbf{D}'_1 \mathbf{x}_a}{\frac{p}{\mathbf{x}'_a \left[\mathbf{I} - \mathbf{D}_1 (\mathbf{D}'_1 \mathbf{D}_1)^{-1} \mathbf{D}'_1 \right] \mathbf{x}_a}} \quad (1.5)$$

$$N - q$$

We would accept this single univariate test if $F(\mathbf{a}) \leq F^*$, for some critical value F^* . Since if $\max_{\mathbf{a}} F(\mathbf{a}) \leq F^*$, all other $F(\mathbf{a})$ must also fall below F^* , it follows that $\max_{\mathbf{a}} F(\mathbf{a})$ can be used to test the

multivariate hypothesis, using *some* critical value that needs to be determined.

Using matrix calculus, you can show that the largest possible value of $F(\mathbf{a})$ is given by the largest root of the determinantal equation

$$|\mathbf{A} - \lambda\mathbf{B}| = 0 \quad (1.6)$$

or, more conveniently, the greatest characteristic root of \mathbf{AB}^{-1} , where

$$\mathbf{A} = \mathbf{M}'\mathbf{X}'\mathbf{D}_1(\mathbf{D}_1'\mathbf{D}_1)^{-1}\mathbf{H}\left[\mathbf{H}'(\mathbf{D}_1'\mathbf{D}_1)^{-1}\mathbf{H}\right]^{-1}\mathbf{H}'(\mathbf{D}_1'\mathbf{D}_1)^{-1}\mathbf{D}_1'\mathbf{X}\mathbf{M} \quad (1.7)$$

and

$$\mathbf{B} = \mathbf{X}'\left[\mathbf{I} - \mathbf{D}_1(\mathbf{D}_1'\mathbf{D}_1)^{-1}\mathbf{D}_1'\right]\mathbf{X} \quad (1.8)$$

The distribution of the *greatest root criterion* (g.r.c.) has been determined and tabular lookups are available. However, modern software reports p -values based on this measure, and the table lookup procedure requires several calculations and intervening steps, so we will not discuss it here.

The g.r.c. is only one of several statistics that could be computed. Some other criteria are *Wilks's* Λ , the *Lawley-Hotelling trace statistic*, and the *Pillai trace statistic*.

***Wilks* Λ**

This criterion is derived via the generalized likelihood-ratio principle. This can be written

$$\begin{aligned}\Lambda &= \frac{|\mathbf{B}|}{|\mathbf{A}| + |\mathbf{B}|} \\ &= \frac{1}{|\mathbf{AB}^{-1} + \mathbf{I}|}\end{aligned}\tag{1.9}$$

There are various transformations of this criterion that lead to both approximate and exact p -values. For example, an asymptotic chi-square statistic is

$$\begin{aligned}\chi_{qu}^2 &= \left[N - q - \frac{1}{2}(u - p + 1) \right] \log[1/\Lambda] \\ &= \left[N - q - \frac{1}{2}(u - p + 1) \right] (\log|\mathbf{A} + \mathbf{B}| - \log|\mathbf{B}|) \quad (1.10)\end{aligned}$$

Lawley-Hotelling Trace Statistic

$$T_0^2 = \text{Tr}(\mathbf{A}\mathbf{B}^{-1}) \quad (1.11)$$

Pillai Trace Statistic

$$V = \text{Tr} \left[\mathbf{A} (\mathbf{A} + \mathbf{B})^{-1} \right] \quad (1.12)$$