Multivariate Analysis of Variance - MANOVA

Dr. J. Kyle Roberts

Southern Methodist University
Simmons School of Education and Human Development
Department of Education Policy and Leadership

Null Hypothesis for ANOVA and MANOVA

- In the case of the Univariate ANOVA, the null hypothesis was stated as:

$$H_0 : \overline{X}_1 = \overline{X}_2 = \overline{X}_3 = \cdots = \overline{X}_K$$

- where K represents the total number of "levels" in the "way" for one independent variable.
- For the One-Way MANOVA, the null hypothesis would be:

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \boldsymbol{\mu}_3 = \cdots = \boldsymbol{\mu}_k$$

- where $\boldsymbol{\mu}$ is a vector of means for a given number of dependent variables.

Null Hypothesis for MANOVA

- We could test to see if the vector of means of the dependent variables is equal for multiple independent groups and our new null would be:

$$H_0 : \begin{bmatrix} \overline{X}_{11} \\ \overline{X}_{21} \\ \vdots \\ \overline{X}_{p1} \end{bmatrix} = \begin{bmatrix} \overline{X}_{12} \\ \overline{X}_{22} \\ \vdots \\ \overline{X}_{p2} \end{bmatrix} = \begin{bmatrix} \overline{X}_{13} \\ \overline{X}_{23} \\ \vdots \\ \overline{X}_{p3} \end{bmatrix} = \cdots = \begin{bmatrix} \overline{X}_{1k} \\ \overline{X}_{2k} \\ \vdots \\ \overline{X}_{pk} \end{bmatrix}$$

- where p represents the total number of dependent variables for k levels.

The Test Statistic for MANOVA

- Recall that for the univariate ANOVA, we compute F as (MS_b/MS_w) and then compute the probability of rejecting H_0 given F and our df.
- For MANOVA, our test statistic is Λ and is computed as:

$$\Lambda = \frac{|\boldsymbol{W}|}{|\boldsymbol{T}|} = \frac{|\boldsymbol{W}|}{|\boldsymbol{B} + \boldsymbol{W}|}$$

- where $\boldsymbol{W}$ and $\boldsymbol{T}$ are determinants of the within and total sum of squares and cross-product matrices.
- This means that if the between effect, $\boldsymbol{B}$, is very large, then Λ approaches 0.
- However, if $\boldsymbol{B}$, is very small or even 0, then Λ approaches 1.

Example from Stevens (2002) p. 213

- Create the following dataset in `R` where there are 2 dependent variables and three groups.

K_1		K_2		K_3	
y_1	y_2	y_1	y_2	y_1	y_2
2	3	4	8	7	6
3	4	5	6	8	7
5	4	5	7	10	8
2	5			9	5
				7	6
$\overline{y}_{11} = 3$	$\overline{y}_{21} = 4$	$\overline{y}_{12} = 5$	$\overline{y}_{22} = 7$	$\overline{y}_{13} = 8.2$	$\overline{y}_{13} = 6.4$

Calculation of $\boldsymbol{W}$

- Put simply, $\boldsymbol{W} = \boldsymbol{W}_1 + \boldsymbol{W}_2 + \boldsymbol{W}_3$ where

$$\boldsymbol{W}_1 = \begin{bmatrix} ss_1 & ss_{12} \\ ss_{21} & ss_2 \end{bmatrix}$$

- and

$$ss_1 = \sum_{j=1}^{4} (y_{1(j)} - \overline{y}_{11})^2$$

$$ss_2 = \sum_{j=1}^{4} (y_{2(j)} - \overline{y}_{21})^2$$

$$ss_{12} = ss_{21} = \sum_{j=1}^{4} (y_{1(j)} - \overline{y}_{11})(y_{2(j)} - \overline{y}_{21})$$

Further Calcualtion of $\boldsymbol{W}_1$

- Using the previous dataset, we can calculate

$$
\begin{aligned}
ss_1 &= (2-3)^2 + (3-3)^2 + (5-3)^2 + (2-3)^2 = 6 \\
ss_2 &= (3-4)^2 + (4-4)^2 + (4-4)^2 + (5-4)^2 = 2 \\
ss_{12} = ss_{21} &= (2-3)(3-4) + (3-3)(4-4) + (5-3)(4-4) \\
&\quad + (2-3)(5-4) = 0
\end{aligned}
$$

- This means that the three within matrices are:

$$\boldsymbol{W}_1 = \begin{bmatrix} 6 & 0 \\ 0 & 2 \end{bmatrix}, \boldsymbol{W}_2 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \boldsymbol{W}_3 = \begin{bmatrix} 6.8 & 2.6 \\ 2.6 & 5.2 \end{bmatrix}$$

Calculation of $\boldsymbol{W}$ based on $\boldsymbol{W}_1$, $\boldsymbol{W}_2$, and $\boldsymbol{W}_3$

$$\boldsymbol{W} = \boldsymbol{W}_1 + \boldsymbol{W}_2 + \boldsymbol{W}_3 = \begin{bmatrix} 14.8 & 1.6 \\ 1.6 & 9.2 \end{bmatrix}$$

- Given the computation of the within matrix, we can further compute the between matrix as

$$\boldsymbol{B} = \begin{bmatrix} 61.87 & 24.40 \\ 24.40 & 19.05 \end{bmatrix}$$

- This means that we can compute the total matrix as:

$$\boldsymbol{T} = \begin{bmatrix} 61.87 & 24.40 \\ 24.40 & 19.05 \end{bmatrix} + \begin{bmatrix} 14.8 & 1.6 \\ 1.6 & 9.2 \end{bmatrix} = \begin{bmatrix} 76.72 & 26.00 \\ 26.00 & 28.25 \end{bmatrix}$$

Calculation of Wilk's Λ

$$\Lambda = \frac{|\boldsymbol{W}|}{|\boldsymbol{T}|} = \frac{\begin{vmatrix} 14.8 & 1.6 \\ 1.6 & 9.2 \end{vmatrix}}{\begin{vmatrix} 76.72 & 26.00 \\ 26.00 & 28.25 \end{vmatrix}} = \frac{14.8(9.2) - 1.6^2}{76.72(28.25) - 26^2} = 0.0897$$

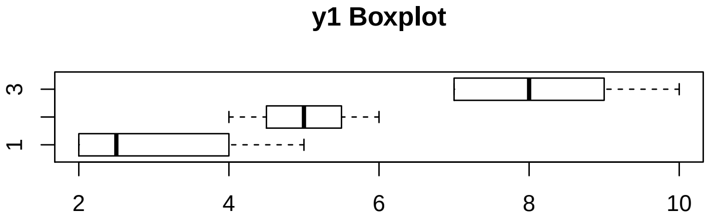
and the χ^2 test statistic is

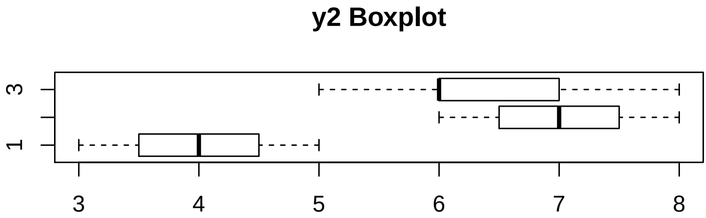
$$\begin{aligned} \chi^2 &= -[(N-1) - 0.5(p+k)] \ln \Lambda \\ &= -[(12-1) - 0.5(2+3)] \ln(0.0897) \\ &= 20.4987 \text{ with } p(k-1) \text{ df} \end{aligned}$$

Setting Up the Data in R

```
> manova.data <- data.frame(group = as.factor(rep(1:3,
+     c(4, 3, 5))), y1 = c(2, 3, 5, 2, 4, 5, 6,
+     7, 8, 10, 9, 7), y2 = c(3, 4, 4, 5, 8, 6,
+     7, 6, 7, 8, 5, 6))
> with(manova.data, tapply(y1, group, mean))

  1   2   3
3.0 5.0 8.2

> with(manova.data, tapply(y2, group, mean))

  1   2   3
4.0 7.0 6.4
```

Graphical Analysis of manova.data





Running the MANOVA in R

```
> (m1 <- manova(cbind(y1, y2) ~ group, manova.data))

Call:
   manova(cbind(y1, y2) ~ group, manova.data)

Terms:
                   group Residuals
resp 1          61.86667  14.80000
resp 2             19.05      9.20
Deg. of Freedom        2         9

Residual standard error: 1.282359 1.01105
Estimated effects may be unbalanced

> summary(m1, test = "Wilks")

          Df    Wilks approx F num Df den Df    Pr(>F)
group      2 0.089674   9.3575      4     16 0.0004271
Residuals  9
```

One-Way MANOVA Homework

Create data for a one-way MANOVA with 4 dependent variables and 4 levels in the way.

1. Create the data in such a way that there is no difference between the first three mean vectors, but there is a difference between the fourth mean vector and the other three. Prove this to me with a Lattice-style boxplot, the Wilk's Lambda, and with post-hoc ANOVAs (last part is not recommended, but just for illustrative purposes).
2. Create the data in such a way that the mean vectors are similar for groups 1, 2, and 4, but in group 3 ONLY one of the means from any of the 4 dependent variables is different from the means of the other 3 groups. Again, prove this to me with a Lattice-style boxplot, the Wilk's Lambda, and with post-hoc ANOVAs (last part is not recommended, but just for illustrative purposes). The results from this analysis should have a statistically significant Wilk's Lambda.