MANOVA: Multivariate ANOVA

- Suppose, a client was interested in testing if there was a significant difference between the sexes for blood pressure (1-way ANOVA or t-test).

And they were interested in testing if there was a significant difference between the sexes for cholesterol (1-way ANOVA or t-test).

- In biological studies, it wouldn’t be unreasonable to think we have more than one response variable measured on each individual.

- It’s possible to gain some power for detecting differences between the sexes by considering both responses simultaneously in a MANOVA.
• A MANOVA is similar to an ANOVA because predictors are still factors, but we have more than one continuous-variable response or outcome on each experimental unit. For example, $\mathbf{y}_i = (y_{i1}, y_{i2})$.

This semester examples:

**Danielle Nauman** (4 outcomes):
1) BCP, 2) IOAP,
   3) BCP duration, 4) IOAP duration

**Ben Miller**: (4 outcomes)
1) DCI, 2) OSBD, 3) Oucher, 4) PRCD

• One option is to perform a separate 1-way ANOVA analysis for each outcome, but another option is to unify the model using a MANOVA.
• 1-way ANOVA null hypothesis (1 factor with 2 levels):

\[ H_o : \mu_1 = \mu_2 \]

• 1-way MANOVA null hypothesis (1 factor with 2 levels, and 2 responses):

\[ H_o : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \quad \text{or} \quad H_o : \begin{pmatrix} \mu_{11} \\ \mu_{12} \end{pmatrix} = \begin{pmatrix} \mu_{21} \\ \mu_{22} \end{pmatrix} \]

• 1-way MANOVA null hypothesis (1 factor with 2 levels, and 3 responses):

\[ H_o : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \quad \text{or} \quad H_o : \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \end{pmatrix} = \begin{pmatrix} \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{pmatrix} \]

• In a MANOVA, we test for the equality of mean vectors, not just means. Rejecting the null says there is a difference in at least one of the responses between the factor groups.
Some comments from others...

- What can we gain by approaching this as a MANOVA?
  - It can protect against Type I errors that might occur if multiple ANOVAs were conducted independently. In MANOVA, we get one overall test to ask if any of the treatment groups were significant for any of the outcomes.
  - By measuring several dependent variables in a single experiment, there is a better chance of discovering which factor is truly important.
  - By looking at things in a higher dimension, it can reveal differences not discovered by ANOVA tests.

- MANOVA works well in situations where there are moderate correlations between DVs (very high correlation can cause estimation problems, and very low correlation says you used degrees of freedom for no real gain).
EXAMPLE*:

\[ N = 33 \] subjects are randomly assigned to one of three groups.

Group 1 receives technical dietary information interactively from an on-line website.

Group 2 receives the same information in from a nurse practitioner.

Group 3 receives the information from a video tape made by the same nurse practitioner.

The researcher looks at three different ratings of the presentation recorded by each individual: difficulty \((y_1)\), useful \((y_2)\) and importance \((y_3)\), to determine if there is a difference in the transfer of information.

*UCLA Academic Technology Services on-line example.
In particular, the researcher is interested in whether the interactive website is superior because that is the most cost-effective way of delivering the information.

The balanced data:

<table>
<thead>
<tr>
<th>Obs</th>
<th>GROUP</th>
<th>USEFUL</th>
<th>DIFFICULTY</th>
<th>IMPORTANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>19.6000</td>
<td>5.1500</td>
<td>9.5000</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>15.4000</td>
<td>5.7500</td>
<td>9.1000</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>22.3000</td>
<td>4.3500</td>
<td>3.3000</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>24.3000</td>
<td>7.5500</td>
<td>5.0000</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>22.5000</td>
<td>8.5000</td>
<td>6.0000</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>20.5000</td>
<td>10.2500</td>
<td>5.0000</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>14.1000</td>
<td>5.9500</td>
<td>18.8000</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>13.0000</td>
<td>6.3000</td>
<td>16.5000</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>14.1000</td>
<td>5.4500</td>
<td>8.9000</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>16.7000</td>
<td>3.7500</td>
<td>6.0000</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>16.8000</td>
<td>5.1000</td>
<td>7.4000</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>17.1000</td>
<td>9.0000</td>
<td>7.5000</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>15.7000</td>
<td>5.3000</td>
<td>8.5000</td>
</tr>
<tr>
<td>26</td>
<td>3</td>
<td>19.2000</td>
<td>4.8500</td>
<td>8.3000</td>
</tr>
<tr>
<td>27</td>
<td>3</td>
<td>12.0000</td>
<td>8.7500</td>
<td>9.0000</td>
</tr>
<tr>
<td>28</td>
<td>3</td>
<td>13.0000</td>
<td>5.2000</td>
<td>10.3000</td>
</tr>
<tr>
<td>29</td>
<td>3</td>
<td>11.9000</td>
<td>4.7500</td>
<td>8.5000</td>
</tr>
<tr>
<td>30</td>
<td>3</td>
<td>12.0000</td>
<td>5.8500</td>
<td>9.5000</td>
</tr>
<tr>
<td>31</td>
<td>3</td>
<td>19.8000</td>
<td>2.8500</td>
<td>2.3000</td>
</tr>
<tr>
<td>32</td>
<td>3</td>
<td>16.5000</td>
<td>6.5500</td>
<td>3.3000</td>
</tr>
<tr>
<td>33</td>
<td>3</td>
<td>17.4000</td>
<td>6.6000</td>
<td>1.9000</td>
</tr>
</tbody>
</table>
Individual one-way ANOVA plots:

Taken individually, we can see some suggestion of differences between the groups for each response, but there is a lot of overlap in the range of values.
Bivariate plots:

 Sometimes, when we look at a response vector rather than an individual response, like \( \mathbf{y}_i = (y_{i1}, y_{i2}, y_{i3}) \), differences between groups can become more apparent (see top right picture above for IMPORTANCE/USEFUL).
Drawing ellipsoids:

1. Using *dataEllipse()* from ‘car’ package:

```r
> library(car)
> attach(manovaData)  ## structure shown earlier in SAS.
> col.vec=c(1,3,4)
> pch.vec=c(1,7,17)

# IMPORTANCE VS. DIFFICULTY
> plot(IMPORTANCE,DIFFICULTY,xlim=c(10,26),ylim=c(2,11),
    col=col.vec[GROUP],pch=pch.vec[GROUP],
    main="DIFFICULTY/IMPORTANCE by GROUP")
> dataEllipse(IMPORTANCE[GROUP==1],DIFFICULTY[GROUP==1],
    levels=c(.2,.4,.6),add=TRUE,plot.points=FALSE,
    col=1,center.pch=4,fill=TRUE, fill.alpha=0.2)
> dataEllipse(IMPORTANCE[GROUP==2],DIFFICULTY[GROUP==2],
    levels=c(.2,.4,.6),add=TRUE,plot.points=FALSE,
    col=3,center.pch=4,fill=TRUE, fill.alpha=0.2)
> dataEllipse(IMPORTANCE[GROUP==3],DIFFICULTY[GROUP==3],
    levels=c(.2,.4,.6),add=TRUE,plot.points=FALSE,
    col=4,center.pch=4,fill=TRUE, fill.alpha=0.2)
> points(mean(IMPORTANCE),mean(DIFFICULTY),cex=2,pch=16)
> legend(10,12,c("Group 1","Group 2","Group 3"),
       pch=pch.vec,col=col.vec)
```
2. Using `stat_ellipse()` from ‘ggplot2’ package:

```r
> library(ggplot2)
> attach(manovaData) ## structure shown earlier in SAS.
> manovaData$GROUP <- as.factor(manovaData$GROUP)

> ggplot(manovaData, aes(USEFUL, DIFFICULTY, fill=GROUP)) +
  geom_point() +
  xlim(5, 30) +
  ylim(-2, 15) +
  stat_ellipse(level=0.95) +
  stat_ellipse(geom = "polygon", level=0.95, alpha=0.3)
```
In the following SAS code, we request a MANOVA analysis using all three response variables.

\[ H_0 : \mu_1 = \mu_2 = \mu_3 \quad \text{or} \quad H_0 : \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \end{pmatrix} = \begin{pmatrix} \mu_{21} \\ \mu_{22} \\ \mu_{23} \end{pmatrix} = \begin{pmatrix} \mu_{31} \\ \mu_{32} \\ \mu_{33} \end{pmatrix} \]

We also perform a hypothesis test for the internet group vs. the traditional groups, and a comparison of the two traditional groups.

/***** MANOVA in SAS *************/

proc glm data=m.manova;
   class group;
   model useful difficulty importance=group;
   contrast '1 vs 2 & 3' group 2 -1 -1;
   contrast '2 vs 3' group 0 1 -1;
   /* MANOVA using all three responses:*/
   manova h=_all_;
run;
The first part of the output provides a 1-way ANOVA for each response variable:

### Dependent Variable: USEFUL

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>52.9242378</td>
<td>26.4621189</td>
<td>2.70</td>
<td>0.0835</td>
</tr>
<tr>
<td>Error</td>
<td>30</td>
<td>293.9654425</td>
<td>9.7988481</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Tot</td>
<td>32</td>
<td>346.8896803</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contrast</th>
<th>DF</th>
<th>Contrast SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 vs 2&amp;3</td>
<td>1</td>
<td>52.74241913</td>
<td>52.74241913</td>
<td>5.38</td>
<td>0.0273</td>
</tr>
<tr>
<td>2 vs 3</td>
<td>1</td>
<td>0.18181870</td>
<td>0.18181870</td>
<td>0.02</td>
<td>0.8926</td>
</tr>
</tbody>
</table>

### Dependent Variable: DIFFICULTY

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>3.9751512</td>
<td>1.9875756</td>
<td>0.47</td>
<td>0.6282</td>
</tr>
<tr>
<td>Error</td>
<td>30</td>
<td>126.2872767</td>
<td>4.2095759</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Tot</td>
<td>32</td>
<td>130.2624279</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contrast</th>
<th>DF</th>
<th>Contrast SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 vs 2&amp;3</td>
<td>1</td>
<td>3.73469643</td>
<td>3.73469643</td>
<td>0.89</td>
<td>0.3538</td>
</tr>
<tr>
<td>2 vs 3</td>
<td>1</td>
<td>0.24045478</td>
<td>0.24045478</td>
<td>0.06</td>
<td>0.8127</td>
</tr>
</tbody>
</table>
The GLM Procedure

Dependent Variable: IMPORTANCE

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>81.8296936</td>
<td>40.9148468</td>
<td>2.88</td>
<td>0.0718</td>
</tr>
<tr>
<td>Error</td>
<td>30</td>
<td>426.3708962</td>
<td>14.2123632</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Tot</td>
<td>32</td>
<td>508.2005898</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contrast</th>
<th>DF</th>
<th>Contrast SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 vs 2&amp;3</td>
<td>1</td>
<td>80.30060224</td>
<td>80.30060224</td>
<td>5.65</td>
<td>0.0240</td>
</tr>
<tr>
<td>2 vs 3</td>
<td>1</td>
<td>1.52909132</td>
<td>1.52909132</td>
<td>0.11</td>
<td>0.7452</td>
</tr>
</tbody>
</table>

The second part of the output provides the information on the MANOVA.

Recall that the 1-way ANOVA breaks the $SS_{total}$ into $SS_{group}$ and $SSE$, and it is the $MS_{group}$ compared to the $MSE$ that provides a test statistics for the null.

The $SS_{group}$ reflects how far the groups means are from the overall mean $\bar{Y}$. 
In the MANOVA, we again calculate a ‘between SS’ and ‘within SS’, but now we are dealing with vectors instead of scalars. The ‘between SS’ represents how far the group mean vector is from the overall mean vector.
There are a variety of test statistics that compare the ‘between SS’ and ‘within SS’, but I will mention Wilks’ lambda $\Lambda^*$ here.

$$\Lambda^* = \frac{|SS_{within}|}{|SS_{between} + SS_{within}|}$$

and we reject $H_0$ when $\Lambda^*$ is small.

The overall test for the MANOVA in SAS:

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>F Value</th>
<th>Num DF</th>
<th>Den DF</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks’ Lambda</td>
<td>0.52578838</td>
<td>3.54</td>
<td>6</td>
<td>56</td>
<td>0.0049</td>
</tr>
<tr>
<td>Pillai’s Trace</td>
<td>0.47667013</td>
<td>3.02</td>
<td>6</td>
<td>58</td>
<td>0.0122</td>
</tr>
<tr>
<td>Hotelling-Lawley Trace</td>
<td>0.89722998</td>
<td>4.12</td>
<td>6</td>
<td>35.61</td>
<td>0.0031</td>
</tr>
<tr>
<td>Roy’s Greatest Root</td>
<td>0.89198790</td>
<td>8.62</td>
<td>3</td>
<td>29</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

NOTE: F Statistic for Roy’s Greatest Root is an upper bound.

NOTE: F Statistic for Wilks’ Lambda is exact.

This p-value is stronger than any of the 1-way ANOVAs taken alone. At least one of the outcomes is significantly different among the groups.
A profile plot of the means can give insight into the relationships among the outcomes and the groups.

> attach(manovaData)
> ppData1 <- data.frame(group=levels(GROUP),
>                        apply(manovaData[,-1],2,tapply,GROUP,mean))
> ppData1

<table>
<thead>
<tr>
<th></th>
<th>USEFUL</th>
<th>DIFFICULTY</th>
<th>IMPORTANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18.11818</td>
<td>6.190909</td>
<td>8.681818</td>
</tr>
<tr>
<td>2</td>
<td>15.52727</td>
<td>5.581818</td>
<td>5.109091</td>
</tr>
<tr>
<td>3</td>
<td>15.34545</td>
<td>5.372727</td>
<td>5.636364</td>
</tr>
</tbody>
</table>

> ppData <- melt(ppData1,id.vars=1)
> with(ppData,interaction.plot(variable,group,value))

It looks like Groups 2 & 3 behave similarly across the outcomes, and Group 1 is higher for each outcome.
The mean-vector contrasts of interest in the MANOVA:

Group 1 vs. others:

The mean vector of Group 1 is significantly different than the average mean vector of the other two groups.

Group 2 vs. Group 3

The mean vector of Group 2 is not significantly different than the mean vector of Group 3.
• You can also follow-up with hypothesis tests comparing the means for each outcome, but including a multiple-testing adjustment would be best practice (such as Bonferroni).

• The MANOVA analysis does assume we have a multivariate normal distribution and that each group has a similar variance-covariance structure, or $\Sigma_1 = \Sigma_2 = \Sigma_2 = \Sigma$.

In SAS, if you request Bartlett’s test for homogeneity in SAS using all three response variables, the null hypothesis of equal variance-covariance matrices is not rejected.

```sas
proc discrim data=m.manova pool=test;
  class group;
  var useful difficulty importance;
run;
```
Test of Homogeneity of Within Covariance Matrices

Chi-Square DF Pr > ChiSq
10.285847 12 0.5909

Since the Chi-Square value is not significant at the 0.1 level, a pooled covariance matrix will be used in the discriminant function.

- But based on the plots, the assumption of constant variance seems most reasonable if we limit ourselves to the IMPORTANCE and USEFUL response variables.
You can choose to use a subset of your response variables for the MANOVA by simply changing your SAS coding in the `manova` line:

```sas
proc glm data=m.manova;
  class group;
  model useful difficulty importance=group;
  contrast '1 vs 2&3' group 2 -1 -1;
  contrast '2 vs 3' group 0 1 -1;
  /* MANOVA using only useful & importance:*/
  manova h=group m=(1 0 1);
run;
```

Output from MANOVA using just USEFUL and IMPORTANCE outcomes:

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>F Value</th>
<th>Num DF</th>
<th>Den DF</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks’ Lambda</td>
<td>0.53598494</td>
<td>12.99</td>
<td>2</td>
<td>30</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Pillai’s Trace</td>
<td>0.46401506</td>
<td>12.99</td>
<td>2</td>
<td>30</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Hotelling-Lawley Trace</td>
<td>0.86572405</td>
<td>12.99</td>
<td>2</td>
<td>30</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Roy’s Greatest Root</td>
<td>0.86572405</td>
<td>12.99</td>
<td>2</td>
<td>30</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
In R, overall F-test using Wilk’s test:

```r
> head(dt)
GROUP USEFUL DIFFICULTY IMPORTANCE
1 1 19.6 5.15 9.5
2 1 15.4 5.75 9.1
3 1 22.3 4.35 3.3
4 1 24.3 7.55 5.0
5 1 22.5 8.50 6.0
6 1 20.5 10.25 5.0
```

```r
> fit <- manova(cbind(USEFUL,DIFFICULTY,IMPORTANCE)~as.factor(GROUP),dt)
> summary(fit, test="Wilk")
Df Wilks approx F num Df den Df Pr(>F)
as.factor(GROUP) 2 0.52579 3.5382 6 56 0.004859 **
Residuals 30
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Perform the Shapiro-Wilk test for multivariate normality for each group separately:

```r
> library(mvnormtest)
> mshapiro.test(t(dt[dt$GROUP==1,c("USEFUL","DIFFICULTY","IMPORTANCE")]))

Shapiro-Wilk normality test
data:  Z
W = 0.86053, p-value = 0.0585
```

```r
> mshapiro.test(t(dt[dt$GROUP==2,c("USEFUL","DIFFICULTY","IMPORTANCE")]))

Shapiro-Wilk normality test
data:  Z
W = 0.95491, p-value = 0.707
```

```r
> mshapiro.test(t(dt[dt$GROUP==3,c("USEFUL","DIFFICULTY","IMPORTANCE")]))

Shapiro-Wilk normality test
data:  Z
W = 0.89874, p-value = 0.1784
```
Check for equality of variance-covariance matrices using Box’s M-test:

```r
> library(biotoools)
> boxM(dt[,,-1], dt[,1])
```

Box's M-test for Homogeneity of Covariance Matrices

data:  dt[, -1]
Chi-Sq (approx.) = 10.286, df = 12, p-value = 0.5909

- References: