Multivariate Analysis of Variance II

Adding Predictors

Andrew Zieffler

Read in the minneapolis.csv data

```
## Read in the data
> mpls = read.csv("http://www.tc.umn.edu/~zief0002/Data/minneapolis.csv")
```

Packages Needed

- **ez**
- **ggplot2**
- **reshape2**

```
> mpls2 = mpls[complete.cases(mpls), ]
```

Remove rows with missing data

Reshape Wide to Long Data

```
## Use the reshape2 package
> library(reshape2)

## Melt the data to the long format
> mplsLong = melt(
    mpls2,
    id = c("studentID", "female"),
    measure = c("grade.5", "grade.6", "grade.7", "grade.8")
    )
```

The `id=` argument

keep these variables as columns

The `measure=` argument

Change these variables into two new ones...`variable` and `value`

```
> head(mplsLong)

  studentID female variable value
1         1      1   read.5   172
2         3      0   read.5   191
3         5      1   read.5   207
4         6      0   read.5   191
5         7      0   read.5   199
6         9      1   read.5   149
```

- **Change the column names from "variable" and "value" (to "grade" and "read")**
- **Change the level names of the new "grade" column**
- **Coerce "female" into a factor**
- **Coerce "studentID" into a factor**

```
> head(mplsLong)

  studentID female   grade read student
1         1 Female grade.5  172       1
2         3   Male grade.5  191       3
3         5 Female grade.5  207       5
4         6   Male grade.5  191       6
5         7   Male grade.5  199       7
6         9 Female grade.5  149       9
```

FIT THE MANOVA MODEL

- Bind the repeated measures together
- Fit the multivariate model using the bound repeated measures as the outcome. In this model we will also include the predictor of "female".

```
# Bind together the outcome columns from the wide data
dvm = cbind(mpls$read.5, mpls$read.6, mpls$read.7, mpls$read.8)

# Fit the multivariate model
mlm.1 = lm(dvm ~ 1 + mpls$female)
mlm.1

Call:
lm(formula = dvm ~ 1 + mpls$female)

Coefficients:
             [,1]      [,2]      [,3]      [,4]
(Intercept)  203.8571  208.8571  215.1429  218.5714
mpls$female   -3.0000    0.5714   -1.8571   -1.1429
```

In fitting the MV model, the predictor needs to be pulled from the data frame (i.e., use the $ notation)

- Set up the intra-subject design which is the names of the repeated measures variables from the wide data
- Fit the multivariate model using the bound repeated measures as the outcome. In this model we will also include the predictor of "female".

```
# Set up the intra-subject design
grade = factor(c("read.5", "read.6", "read.7", "read.8"))

# Get the output from the MANOVA
library(car)
Anova(mlm.1, idata = data.frame(grade), idesign = ~ grade, test = "Pillai")

Type II Repeated Measures MANOVA Tests: Pillai test statistic
                  Df test stat approx F num Df den Df    Pr(>F)
(Intercept)        1   0.99138  1379.66      1     12 9.307e-14 ***
mpls$female        1   0.00119     0.01      1     12 0.9068611
grade              1   0.80929    14.15      3     10 0.0006262 ***
mpls$female:grade  1   0.25100     1.12      3     10 0.3877099
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- The interaction is non-significant (p = 0.388)
- There is an effect of grade (p < .001), but no effect of sex (p = .907).
- These are similar to the results we found with the RM-ANOVA

- To compute an effect size, examine the Wilk's Lambda value

```
# Get the output from the MANOVA
Anova(mlm.1, idata = data.frame(grade), idesign = ~ grade, test = "Wilks")

Type II Repeated Measures MANOVA Tests: Wilks test statistic
                  Df test stat approx F num Df den Df    Pr(>F)
(Intercept)        1   0.00862  1379.66      1     12 9.307e-14 ***
mpls$female        1   0.99881     0.01      1     12 0.9068611
grade              1   0.19071    14.15      3     10 0.0006262 ***
mpls$female:grade  1   0.74900     1.12      3     10 0.3877099
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Female: 1–0.99881 = 0.001
- Grade: 1 – 0.19071 = .809 (roughly 80% of the variation in reading scores is explained through the mean patterns across grade