A Brief Review of Univariate Statistics

Today . . .

- Extending ANOVA to repeated measures (the M matrix)
- Combining group comparisons (L matrix) with repeated measures (the M matrix)

Analysis of variance designs can be extended to include measures collected repeatedly for each person . . .

	Time 1	Time 2	Time 3	. . .	Time t
Treatment					
Control					

and multiple treatment levels that are given to each person . . .

	Treatment A_1		Treatment A_2	
	Time 1	Time 2	Time 1	Time 2
Treatment B_1				
Treatment B_2				

In these designs, the total variability (SS_{Total}) is still partitioned into parts representing systematic and random variability, but now there are more parts. Some new assumptions must be satisfied as well.

There are a variety of ways to get the ANOVA model with repeated measures in R. Depending on the method used, the data will need to be in either wide (traditional) format or long format.

Data in wide format have all information for each case in a single row; multiple columns contain the repeated measures.

Data in long format have each repeated measure in a separate row. A single column contains all of the repeated measures values.

```
##    Subject Group DV1 DV2 DV3 G1 G2 G3
## 1        1     1   5   6   5  1  0  0
## 2        2     1   4   5   7  1  0  0
## 3        3     1   5   3   4  1  0  0
## 4        4     1   4   5   6  1  0  0
## 5        5     1   3   6   5  1  0  0
## 6        6     1   5   4   4  1  0  0
## 7        7     1   5   4   7  1  0  0
## 8        8     1   6   5   5  1  0  0
## 9        9     1   7   7   8  1  0  0
## 10      10     1   6   4   7  1  0  0
## 11      11     1   6   5   5  1  0  0
## 12      12     1   6   3   5  1  0  0
```

Wide Format

```
##    Subject Group G1 G2 G3 Time DV T1 T2 T3
## 1        1     1  1  0  0    1  5  1  0  0
## 2        1     1  1  0  0    2  6  0  1  0
## 3        1     1  1  0  0    3  5  0  0  1
## 4        2     1  1  0  0    1  4  1  0  0
## 5        2     1  1  0  0    2  5  0  1  0
## 6        2     1  1  0  0    3  7  0  0  1
## 7        3     1  1  0  0    1  5  1  0  0
## 8        3     1  1  0  0    2  3  0  1  0
## 9        3     1  1  0  0    3  4  0  0  1
## 10       4     1  1  0  0    1  4  1  0  0
## 11       4     1  1  0  0    2  5  0  1  0
## 12       4     1  1  0  0    3  6  0  0  1
```

Long Format

We will treat these data as coming from a 3 x 3 (Group x Time) design.

The traditional approach using the aov() function requires the data in long format. Group, Time, and Subject must be specified as factors. In a repeated measures design, the repeated measure, Time, is considered to be nested within the Subject factor (i.e., each case has its own profile of scores). The Error( ) designation is needed to define the error terms correctly.

```
ANOVA_1 <- aov(Data_Long$DV ~ Time_F + Group_F + Time_F:Group_F +
    Error(Subject_F/Time_F), data = Data_Long)
```

The standard ANOVA model assumes the factors are fixed. What does that mean?

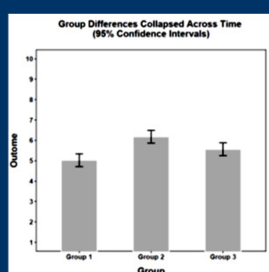
The aov( ) function produces Type I sums of squares rather than Type II or Type III sums of squares. What does that mean and when will it matter?

```
Error: Subject_F
          Df Sum Sq Mean Sq F value   Pr(>F)
Group_F    2   49.4   24.69    13.2 0.000013
Residuals 72  134.5    1.87

Error: Subject_F:Time_F
                Df Sum Sq Mean Sq F value  Pr(>F)
Time_F           2   66.7    33.4    42.7 2.7e-15
Time_F:Group_F   4   71.5    17.9    22.9 1.2e-14
Residuals      144  112.5     0.8
```

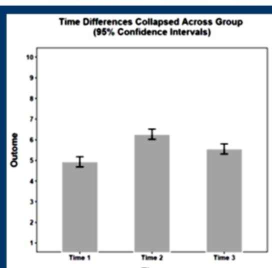
We have three significant effects. What can we tell about them from this information? Why do we have two different error terms?

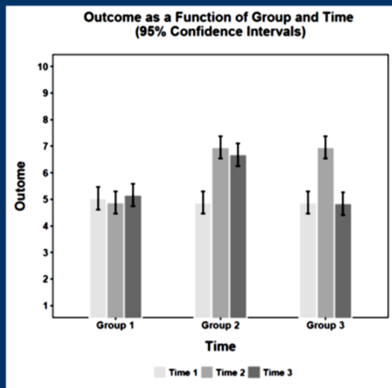
```
CLD(Group_emm, alpha = 0.05, adjust = "holm", details = TRUE)

## $emmeans
##  Group_F emmean     SE df lower.CL upper.CL .group
##  Group 1  5.027 0.1578 72    4.640    5.414  1
##  Group 3  5.560 0.1578 72    5.173    5.947   2
##  Group 2  6.173 0.1578 72    5.786    6.560    3
```

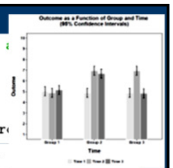
Pairwise comparisons can help isolate the nature of the main effects and interaction. The emmeans package provides a convenient way to do this.



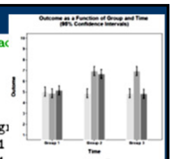

```
CLD(Time_emm, alpha = 0.05, adjust = "holm", details = TRUE)

## $emmeans
##  Time_F emmean     SE    df lower.CL upper.CL .group
##  Time 1  4.933 0.1235 179.9    4.635    5.232  1
##  Time 3  5.560 0.1235 179.9    5.262    5.858   2
##  Time 2  6.267 0.1235 179.9    5.968    6.565    3
```







```
CLD(Group_x_Time_emm, by = "Group_F", alpha = 0.05, a
    details = TRUE)

## $emmeans
## Group_F = Group 1:
##  Time_F emmean     SE    df lower.CL upper.CL .gr
##  Time 2   4.88 0.2139 179.9     4.28     5.48  1
##  Time 1   5.04 0.2139 179.9     4.44     5.64  1
##  Time 3   5.16 0.2139 179.9     4.56     5.76  1
##
## Group_F = Group 2:
##  Time_F emmean     SE    df lower.CL upper.CL .group
##  Time 1   4.88 0.2139 179.9     4.28     5.48  1
##  Time 3   6.68 0.2139 179.9     6.08     7.28   2
##  Time 2   6.96 0.2139 179.9     6.36     7.56   2
##
## Group_F = Group 3:
##  Time_F emmean     SE    df lower.CL upper.CL .group
##  Time 3   4.84 0.2139 179.9     4.24     5.44  1
##  Time 1   4.88 0.2139 179.9     4.28     5.48  1
##  Time 2   6.96 0.2139 179.9     6.36     7.56   2
```



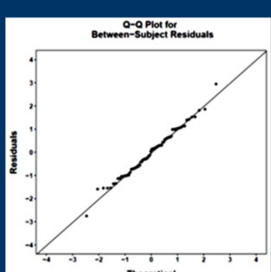
```
CLD(Group_x_Time_emm, by = "Time_F", alpha = 0.05, ad
    details = TRUE)

## $emmeans
## Time_F = Time 1:
##  Group_F emmean     SE    df lower.CL upper.CL .gr
##  Group 2   4.88 0.2139 179.9     4.28     5.48  1
##  Group 3   4.88 0.2139 179.9     4.28     5.48  1
##  Group 1   5.04 0.2139 179.9     4.44     5.64  1
##
## Time_F = Time 2:
##  Group_F emmean     SE    df lower.CL upper.CL .group
##  Group 1   4.88 0.2139 179.9     4.28     5.48  1
##  Group 2   6.96 0.2139 179.9     6.36     7.56   2
##  Group 3   6.96 0.2139 179.9     6.36     7.56   2
##
## Time_F = Time 3:
##  Group_F emmean     SE    df lower.CL upper.CL .group
##  Group 3   4.84 0.2139 179.9     4.24     5.44  1
##  Group 1   5.16 0.2139 179.9     4.56     5.76  1
##  Group 2   6.68 0.2139 179.9     6.08     7.28   2
```

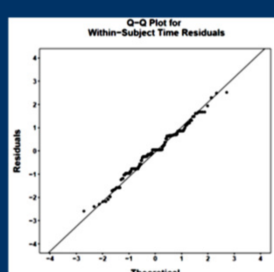
The analysis of variance makes several assumptions. The residuals are assumed to be normally distributed. The variance-covariance matrices for the repeated measures are assumed to be homogeneous across between-subjects conditions (Group in this case). The repeated measures are assumed to satisfy the sphericity assumption.

```r
resid_data_b <- residuals(ANOVA_1$Subject_F)
resid_data_b <- as.data.frame(resid_data_b)
names(resid_data_b) <- c("R")

resid_data_w_Time <- residuals(ANOVA_1$`Subject_F:Time_F`)
resid_data_w_Time <- as.data.frame(resid_data_w_Time)
names(resid_data_w_Time) <- c("R")

# Shapiro-Wilk test.
shapiro.test(resid_data_b$R)

##
##  Shapiro-Wilk normality test
##
## data:  resid_data_b$R
## W = 0.99, p-value = 0.8

shapiro.test(resid_data_w_Time$R)

##
##  Shapiro-Wilk normality test
##
## data:  resid_data_w_Time$R
## W = 0.99, p-value = 0.1
```

Both sets of residuals are normally distributed.

The close approximation to the normal distribution is especially clear in Q-Q plots.





Each group in a repeated measures design has a variance-covariance matrix for the multiple measures. These matrices are assumed to be homogeneous across the groups.

	Time 1	Time 2	Time 3
Group 1			
Group 2			
Group 3			

→

	Time 1	Time 2	Time 3
Time 1	σ^2_{11}	σ_{12}	σ_{13}
Time 2	σ_{21}	σ^2_{22}	σ_{23}
Time 3	σ_{31}	σ_{32}	σ^2_{33}

```
Box <- boxM(Data_Wide[3:5], Data_Wide$Group_F)

Box

##
##  Box's M-test for Homogeneity of Covariance Matrices
##
## data:  Data_Wide[3:5]
## Chi-Sq (approx.) = 4.2, df = 12, p-value = 1

Box$cov

## $`Group 1`
##         DV1   DV2   DV3
## DV1 0.9567 0.255 0.285
## DV2 0.2550 1.193 0.145
## DV3 0.2850 0.145 1.390
##
## $`Group 2`
##         DV1    DV2    DV3
## DV1 1.0267 0.4950 0.4183
## DV2 0.4950 1.1233 0.3617
## DV3 0.4183 0.3617 0.8933
##
## $`Group 3`
##         DV1    DV2    DV3
## DV1 1.2767 0.3700 0.5633
## DV2 0.3700 1.1233 0.3683
## DV3 0.5633 0.3683 1.3067
```

```
Box$pooled

##         DV1    DV2    DV3
## DV1 1.0867 0.3733 0.4222
## DV2 0.3733 1.1467 0.2917
## DV3 0.4222 0.2917 1.1967
```

In repeated measures designs, the homogeneity assumption is more general, now including covariances in addition to variances. Box's test can be used (but assumes normality).

Repeated measures designs also must satisfy the *sphericity assumption*. The sphericity assumption is met when the pooled variance-covariance matrix exhibits compound symmetry, but that is unlikely to happen in a repeated measures design.

	Time 1	Time 2	Time 3
Time 1	σ^2_{11}	σ_{12}	σ_{13}
Time 2	σ_{21}	σ^2_{22}	σ_{23}
Time 3	σ_{31}	σ_{32}	σ^2_{33}

Instead, measures that are spaced farther apart are usually correlated less strongly than measures that are collected more closely in time.

If (Time 2-Time 1) = (Time 3-Time 2)

$$(\sigma_{12}^2 = \sigma_{23}^2) > \sigma_{13}^2)$$

	Time 1	Time 2	Time 3
Time 1	σ^2_{11}	σ_{12}	σ_{13}
Time 2	σ_{21}	σ^2_{22}	σ_{23}
Time 3	σ_{31}	Σ_{32}	σ^2_{33}

More generally, sphericity is met when there is homogeneity of the variances for all possible difference scores:

$$\sigma^2_{Time_2 - Time_1} = \sigma^2_{Time_3 - Time_1} = \sigma^2_{Time_3 - Time_2}$$

A reminder that all repeated measures analyses are difference score analyses.

A test for sphericity is available with the ezANOVA( ) function in the ez package.

```
ANOVA_2 <- ezANOVA(Data_Long, dv = DV, wid = Subject_F, within = .(Time_F),
    between = Group_F, detailed = TRUE, return_aov = TRUE, type = 2)
ANOVA_2$ANOVA

##            Effect DFn DFd     SSn   SSd       F          p p<.05
## 1     (Intercept)   1  72 7022.44 134.5 3759.04 6.924e-64     *
## 2         Group_F   2  72   49.39 134.5   13.22 1.289e-05     *
## 3          Time_F   2 144   66.75 112.5   42.74 2.690e-15     *
## 4 Group_F:Time_F   4 144   71.47 112.5   22.88 1.199e-14     *
##       ges
## 1 0.9660
## 2 0.1667
## 3 0.2128
## 4 0.2244

ANOVA_2$`Mauchly's Test for Sphericity`

##            Effect      W      p p<.05
## 3          Time_F 0.9836 0.5558
## 4 Group_F:Time_F 0.9836 0.5558
```

Sphericity is met for these data.

Violations of sphericity are especially problematic. They can inflate the Type I error rate. Remedies include:

- Single df contrasts
- Adjusted df tests (e.g., Geisser-Greenhouse)
- Multivariate tests (no sphericity assumption)
- Resampling procedures (no assumptions)

```
ANOVA_2$`Sphericity Corrections`

##            Effect    GGe     p[GG] p[GG]<.05   HFe     p[HF]
## 3          Time_F 0.9839 4.341e-15         * 1.011 2.690e-15
## 4 Group_F:Time_F 0.9839 1.901e-14         * 1.011 1.199e-14
##   p[HF]<.05
## 3         *
## 4         *
```

An alternative is to use the lmer( ) function from the lme4 package. This package is used for multilevel models, of which a repeated measures design is an example. In this case, the repeated measures are treated as nested within subjects. The dummy codes for time and group are used in a no-intercept model so that the coefficients are the cell means for all Group x Time combinations.

```
ANOVA_3 <- lmer(Data_Long$DV ~ -1 + G1:T1 + G1:T2 + G1:T3 + +G2:T1 +
    G2:T2 + G2:T3 + +G3:T1 + G3:T2 + G3:T3 + (1 | Subject), data = Data_Long)
```

Specifying a random intercept for subjects assumes Time is fixed to match the aov( ) assumption.

```
Fixed effects:
      Estimate Std. Error t value
G1:T1    5.040      0.214    23.6
G1:T2    4.880      0.214    22.8
G1:T3    5.160      0.214    24.1
T1:G2    4.880      0.214    22.8
T2:G2    6.960      0.214    32.5
T3:G2    6.680      0.214    31.2
T1:G3    4.880      0.214    22.8
T2:G3    6.960      0.214    32.5
T3:G3    4.840      0.214    22.6
```

The regression coefficients are the Group x Time means.

```
Correlation of Fixed Effects:
      G1:T1 G1:T2 G1:T3 T1:G2 T2:G2 T3:G2 T1:G3 T2:G3
G1:T2 0.317
G1:T3 0.317 0.317
T1:G2 0.000 0.000 0.000
T2:G2 0.000 0.000 0.000 0.317
T3:G2 0.000 0.000 0.000 0.317 0.317
T1:G3 0.000 0.000 0.000 0.000 0.000 0.000
T2:G3 0.000 0.000 0.000 0.000 0.000 0.000 0.317
T3:G3 0.000 0.000 0.000 0.000 0.000 0.000 0.317 0.317
```

Note the compound symmetry assumption imposed on the data.

Because we have means that vary across groups and repeated measures, we have the potential to create linear combinations of groups (L) and measures (M) at the same time.

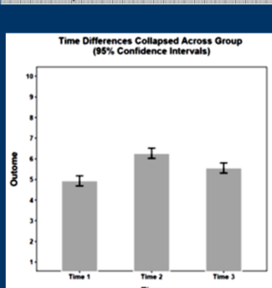
Our means are in a single vector instead of a matrix (as in **LXM**). That means we need to combine the **L** and **M** weights in a vector as well.

Any combination of means can be examined. For example, previously we found a significant 2 df omnibus test for Time. Now we can see if the separate linear and quadratic components are significant.

```
LM_Matrix

##                        [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9]
## Overall Linear            1    0   -1    1    0   -1    1    0   -1
## Overall Quadratic         1   -2    1    1   -2    1    1   -2    1
## G3:Linear                 0    0    0    0    0    0    1    0   -1
## G1:Quadratic              1   -2    1    0    0    0    0    0    0
## G1 vs G2                  1    1    1   -1   -1   -1    0    0    0
## Linear: G1 vs G2          1    0   -1   -1    0    1    0    0    0
## Quadratic: G1 vs G3      -1    2   -1    0    0    0    1   -2    1
## T1: G1 and G2 vs G3       1    0    0    1    0    0   -2    0    0
```

Each is a single-df test that does not require sphericity.

```
Linear Hypotheses:
                         Estimate Std. Error z value Pr(>|z|)
Overall Linear == 0        -1.880      0.433   -4.34  5.6e-05
Overall Quadratic == 0     -6.120      0.750   -8.16  1.8e-15
```

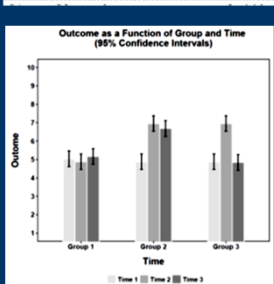


Both the overall linear effect and the overall quadratic effect are significant.

Perhaps we have a more focused question about these temporal trends. We can easily test them in single groups.

```
LM_Matrix

##                      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9]
## Overall Linear          1    0   -1    1    0   -1    1    0   -1
## Overall Quadratic       1   -2    1    1   -2    1    1   -2    1
## G3:Linear               0    0    0    0    0    0    1    0   -1
## G1:Quadratic            1   -2    1    0    0    0    0    0    0
## G1 vs G2                1    1    1   -1   -1   -1    0    0    0
## Linear: G1 vs G2        1    0   -1   -1    0    1    0    0    0
## Quadratic: G1 vs G3    -1    2   -1    0    0    0    1   -2    1
## T1: G1 and G2 vs G3     1    0    0    1    0    0   -2    0    0
```

```
Linear Hypotheses:
                          Estimate Std. Error z value Pr(>|z|)
Overall Linear == 0         -1.880      0.433   -4.34  5.6e-05
Overall Quadratic == 0      -6.120      0.750   -8.16  1.8e-15
G3:Linear == 0               0.040      0.250    0.16     1.00
G1:Quadratic == 0            0.440      0.433    1.02     0.93
```



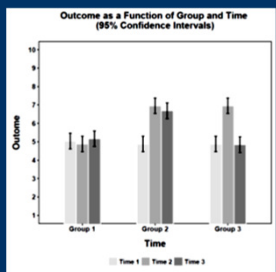
There is no meaningful linear trend in Group 3 and no meaningful quadratic trend in Group 1.

We can easily compare trends between groups.

```
LM_Matrix

##                      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9]
## Overall Linear          1    0   -1    1    0   -1    1    0   -1
## Overall Quadratic       1   -2    1    1   -2    1    1   -2    1
## G3:Linear               0    0    0    0    0    0    1    0   -1
## G1:Quadratic            1   -2    1    0    0    0    0    0    0
## G1 vs G2                1    1    1   -1   -1   -1    0    0    0
## Linear: G1 vs G2        1    0   -1   -1    0    1    0    0    0
## Quadratic: G1 vs G3    -1    2   -1    0    0    0    1   -2    1
## T1: G1 and G2 vs G3     1    0    0    1    0    0   -2    0    0
```

Linear Hypotheses:

	Estimate	Std. Error	z value	Pr(>|z|)
Linear: G1 vs G2 == 0	1.680	0.353	4.75	1.0e-05
Quadratic: G1 vs G3 == 0	-4.640	0.612	-7.58	2.4e-13



Groups 1 and 2 differ in their linear trends; Groups 1 and 3 differ in their quadratic trends.

The move to repeated measures expands the complexity of the designs that can be tested, but all can be thought of inherently as tests of linear combinations of means. That fact allows us to isolate questions of particular importance despite large numbers and complex arrangements of means. The weights in ANOVA are chosen (theory-driven).

Linear combinations in our first multivariate method will be empirically derived.

Next time . . .

Principal components analysis