Repeated Measures ANOVA

When same set of samples are being observed at different conditions, it is known as repeated measures. In this case, standard ANOVA is not appropriate because it fails to model the correlation between the repeated measures, it means the data violate the ANOVA assumption of independence. This leads to biased variance estimate and thus misleading statistical inference.

Within-Subject Factors: When a dependent variable is measured repeatedly for all sample members across a set of conditions, this set of conditions is called a within-subject factor.

Between-Subject Factors: When a dependent variable is measured for different set of samples for each condition, so that each subject is exposed to single condition.

Analysis can have either only within-subjects factors, between-subjects factors, or both.

Univariate Approach

In Univariate approach of repeated measures ANOVA, we treat each response as a separate (univariate) observation and treat “subject” as a (random) blocking factor. In this form, there are assumptions about the nature of the within-subject correlation that are not met fairly frequently, but standard “adjustments” help. Univariate tests give you more power but it requires Sphericity assumption. Otherwise, the p-value reported by repeated measures ANOVA will be too small.

Multivariate Approach

In Multivariate approach, we treat the several responses on one subject as a single “multivariate” response and model the correlation between the components of that response. The main statistics (SS, MS) are now matrices rather than individual numbers. This approach corresponds to results labeled “multivariate” under “repeated measures ANOVA” for most statistical packages.

Multilevel/Mixed Modelling

In this approach, we treat each measurement as univariate, but appropriately

model the correlations via a “hierarchy” of effects. This is a more modern univariate approach called “mixed models” that subsumes a variety of models in a single unified approach.
Advantages of Mixed model: Handle missing data, allow heterogeneous groups, flexible in modeling correlations.

Sphericity Assumption: For each possible pair of levels of a within-subject factor A (e.g., A1 and A2 or A2 and A3), the variance of difference for each pair of levels are equal.
For example:

	A1	A2	A3
Participant 1	8	9	12
Participant 2	6	11	16
Participant 3	9	8	12

	A1-A2	A1-A3	A2-A3
Participant 1	-1	-4	-3
Participant 2	-5	-10	-10
Participant 3	1	-3	-3

Sphericity Assumption basically implies that the covariance matrix of the repeated measures are compound symmetry (CS).
Compound symmetry means all the covariances (the off-diagonal elements of the covariance matrix) are equal and all the variances are equal in the population.

TEST OF SPHERICITY

To test the assumption, we can use Mauchly's Sphericity Test. The Null hypothesis for the test is that the Sphericity assumption is satisfied.

But this test lacks statistical power for small n, also it is not robust to violations of assumptions such as normality and also it doesn't reveal the degree of violation.

If test of sphericity is significant($p>0.05$), sphericity assumption will be violated. There are two correction methods for this violation, Greenhouse-Geisser and Huynh-Feldt corrections.

These two methods will correct the dfs for the F test. The correction factor is called epsilon (ε). The Greenhouse-Geisser (GG) correction is a conservative correction (i.e. yield a larger p value) while the Huynd-Feldt (HF) correction is more liberal (yield a smaller p value).

If estimated epsilon (ε) <0.75, use the GG correction. Otherwise use HF correction.

MANOVA test:

It is a multivariate approach called Multivariate analysis of variance (MANOVA). This test does not require sphericity assumption and treats repeated measure as correlated multiple responses. However, this MANOVA test is less powerful than repeated measure ANOVA and it requires a large sample size.

Example1: **One-way Repeated Measures ANOVA**

Consider we have scores of 5 students for weekly exams (for 5 weeks).

Subject	Week1	Week2	Week3	Week4	Week5
1	6	10	8	4	5
2	2	4	8	5	6
3	2	4	8	5	5
4	4	5	8	10	7
5	4	7	9	7	11

The given data is in wide format, we have to convert it into long format(stack the repeated measures in one variable.

For example:

person week score

1	1	6
1	2	10
1	3	8
1	4	4
1	5	5
2	1	2…..

Now, run the repeated measures Anova test.

R code:

Firstly, we will create dataframe or import the table externally and convert it into long format.

```
#Create dataframe
Week1<- c(6,2,2,4,4)
Week2<- c(10,4,4,5,7)
Week3<- c(8,8,8,8,9)
Week4<- c(4,5,5,10,7)
Week5<- c(5,6,5,7,11)

data <- data.frame(Week1, Week2, Week3, Week4, Week5)
data


#Convert this wide format into long format
dataLong <- reshape(data= data,
                    varying = c("Week1", "Week2", "Week3",
                    v.names = "Scores",
                    timevar = "week",
                    idvar = "subject",
                    direction = "long"
)


dataLong <- dataLong[order(dataLong$subject),]
View(dataLong)

dataLong$week <- factor(dataLong$week)
dataLong$subject <- factor(dataLong$subject)
```

```
> Week1<- c(6,2,2,4,4)
> Week2<- c(10,4,4,5,7)
> Week3<- c(8,8,8,8,9)
> Week4<- c(4,5,5,10,7)
> Week5<- c(5,6,5,7,11)
> data <- data.frame(Week1, Week2, Week3, Week4, Week5)
> data
  Week1 Week2 Week3 Week4 Week5
1     6    10     8     4     5
2     2     4     8     5     6
3     2     4     8     5     5
4     4     5     8    10     7
5     4     7     9     7    11
> #Convert this wide format into long format
> dataLong <- reshape(data= data,
+                     varying = c("Week1", "Week2", "Week3", "Week4", "Week5
"),
+                     v.names = "Scores",
+                     timevar = "week",
+                     idvar = "subject",
+                     direction = "long"
+ )
> dataLong <- dataLong[order(dataLong$subject),]
> View(dataLong)
> dataLong$week <- factor(dataLong$week)
> dataLong$subject <- factor(dataLong$subject)
>
```

Fig. (1)

After creating dataframe in long format, find mean value for all varying factor.

```r
aggData <- aggregate(dataLong$Scores,
                     by = list(week= dataLong$week),
                     FUN = mean)
xtabs(x ~ ., data= aggData)
```

```
> aggData <- aggregate(dataLong$Scores,
+                      by = list(week= dataLong$week),
+                      FUN = mean)
> xtabs(x ~ ., data= aggData)
week
  1   2   3   4   5 
3.6 6.0 8.2 6.2 6.8 
> 
```

Fig. (2)

Now, use anova measures model and analyze it.

```r
model <- aov_ez(id ="subject",
                dv = "Scores",
                data = dataLong,
                within = "week")

names(model)
summary(model)
```

```
[1] "anova_table" "aov"         "Anova"       "lm"         
[5] "data"       
> summary(model)

Univariate Type III Repeated-Measures ANOVA Assuming Sphericity

            Sum Sq num Df Error SS den Df  F value    Pr(>F)    
(Intercept) 948.64      1    29.36      4 129.2425 0.0003414 ***
week         55.76      4    56.24     16   3.9659 0.0201417 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Mauchly Tests for Sphericity

     Test statistic  p-value
week      0.0004514 0.063106


Greenhouse-Geisser and Huynh-Feldt Corrections
 for Departure from Sphericity

      GG eps Pr(>F[GG])  
week 0.47126    0.06811 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

        HF eps Pr(>F[HF])
week 0.8777199 0.02649646
> 
```

(Top Level) R Script

Fig. (3)

Here, we can see that p value is less than 0.05 and therefore all effects are significant.

Also, for sphericity Mauchly test, p value is 0.0631 which is greater than 0.05, therefore sphericity assumption is satisfied. Results also include Greenhouse-Geisser and Huynh-Feldt corrections.

Let’s see the results for MANOVA test,

```
model$Anova
```

```
> model$Anova

Type III Repeated Measures MANOVA Tests: Pillai test statistic
            Df test stat approx F num Df den Df    Pr(>F)    
(Intercept)  1   0.96998   129.24      1      4 0.0003414 ***
week         1   0.99965   721.16      4      1 0.0279202 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> 
```

Fig. (4)

Now, let’s see mean plots.

```
afex_plot(model, x="week", error="within")

afex_plot(model, x="week",
          error="within",
          factor_levels = list(week = c("week1", "week2", "week3
  ggplot2:: geom_line(ggplot2::aes(group = 1))
```

(Top Level) ÷ R Script ÷

nent History Connections

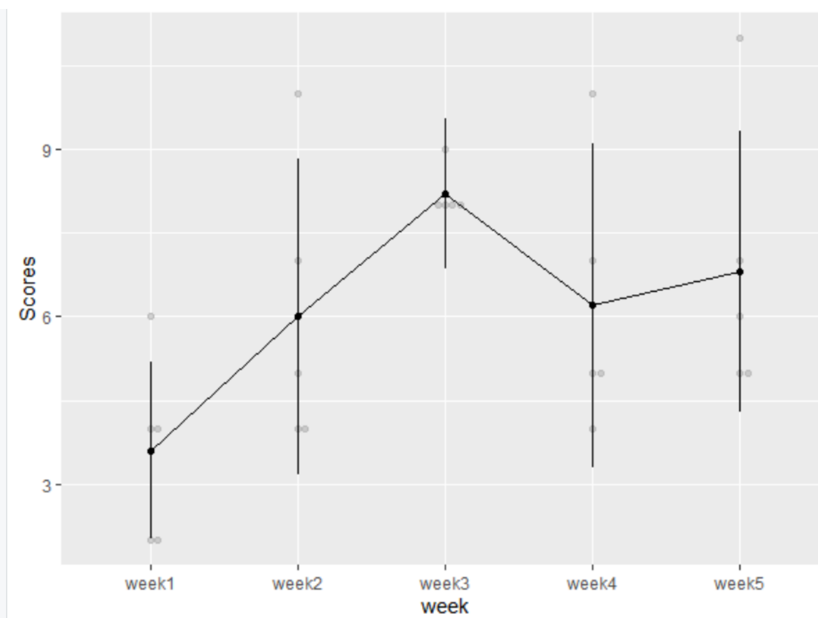

Fig. (5)

Example 2: **Mixed Factorial Design**

Consider we have data of scores of students for 4 weeks. Students are divided into two groups, first group is undergoing training and second group is control group and we want to know if this extra training affect scores of students. Dataset can be downloaded from:

https://github.com/palak-j/Hypothesis-testing-Anova/blob/main/Repeated_measures_ANOVA/Repeated_measures.csv

Firstly, we will create dataframe or import the table externally and convert it into long format.

```r
data <- read.csv("D:/data/Repeated_measures.csv",
                 fileEncoding = 'UTF-8-BOM',header=TRUE)
data

View(data)
names(data)

data$Subject <- factor(data$Subject)

dataLong <- reshape(data= data,
                    varying = c("Week1", "Week2", "Week3", "Wee
                    v.names = "Scores",
                    timevar = "Week",
                    idvar = "subject",
                    direction = "long")
View(dataLong)

dataLong <- dataLong[order(dataLong$subject),]
dataLong

dataLong$Week <- factor(dataLong$Week)
dataLong$Subject <- factor(dataLong$Subject)
dataLong$Group <- factor(dataLong$Group)

aggData <- aggregate(dataLong$Scores,
                     by = list(time= dataLong$Week),
                     FUN = mean)
xtabs(x ~ ., data= aggData)
```

```
14.4    14   2   4    0   14
15.1    15   2   1   35   15
15.2    15   2   2    8   15
15.3    15   2   3    0   15
15.4    15   2   4    0   15
16.1    16   2   1    7   16
16.2    16   2   2    0   16
16.3    16   2   3    9   16
16.4    16   2   4   37   16
17.1    17   2   1   51   17
17.2    17   2   2   53   17
17.3    17   2   3    8   17
17.4    17   2   4   26   17
18.1    18   2   1   25   18
18.2    18   2   2    0   18
18.3    18   2   3    0   18
18.4    18   2   4   15   18
19.1    19   2   1   59   19
19.2    19   2   2   45   19
19.3    19   2   3   11   19
19.4    19   2   4   16   19
20.1    20   2   1   40   20
20.2    20   2   2    2   20
20.3    20   2   3   33   20
20.4    20   2   4   16   20
> xtabs(x ~ ., data= aggData)
time
    1     2     3     4
24.60 19.75 15.75 16.60
>
```

Fig. (6)

Now, use repeated measures ANOVA ignoring the group factor.

```r
model1 <- aov_ez(id ="subject",
                 dv = "Scores",
                 data = dataLong,
                 within = "Week")

names(model1)
summary(model1)

#MANOVA for model1
model1$Anova

#Mean plot for model1
afex_plot(model1, x="Week", error="within")

afex_plot(model1, x="Week",
          error="within",
          factor_levels = list(Week = c("Week1", "Week2", "Week
  ggplot2:: geom_line(ggplot2::aes(group = 1))
```

```
+                within = "Week")
> names(model1)
[1] "anova_table" "aov"         "Anova"       "lm"
[5] "data"
> summary(model1)

Univariate Type III Repeated-Measures ANOVA Assuming Sphericity

             Sum Sq num Df Error SS den Df F value    Pr(>F)
(Intercept) 29414.4      1  11482.6     19 48.6716 1.202e-06 ***
Week          962.4      3   8920.6     57  2.0499    0.1171
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Mauchly Tests for Sphericity

     Test statistic p-value
Week        0.68824   0.251


Greenhouse-Geisser and Huynh-Feldt Corrections
 for Departure from Sphericity

      GG eps Pr(>F[GG])
Week 0.79526     0.1324

        HF eps Pr(>F[HF])
Week 0.9171932   0.123045
>
```

Fig. (7)

```r
model1 <- aov_ez(id ="subject",
                 dv = "Scores",
                 data = dataLong,
                 within = "Week")

names(model1)
summary(model1)

#MANOVA for model1
model1$Anova

#Mean plot for model1
afex_plot(model1, x="Week", error="within")

afex_plot(model1, x="Week",
          error="within",
          factor_levels = list(Week = c("Week1", "Week2", "Week
  ggplot2:: geom_line(ggplot2::aes(group = 1))
```

```
> model1$Anova

Type III Repeated Measures MANOVA Tests: Pillai test statistic
            Df test stat approx F num Df den Df    Pr(>F)
(Intercept)  1   0.71923   48.672      1     19 1.202e-06 ***
Week         1   0.17926    1.238      3     17    0.3269
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

Fig. (8)

we can see value of results from summary of Anova and p value for sphericity is greater than 0.05 and therefore, sphericity assumption is not violated and we can use univariate approach.

Now, add the group factor and analyze all the differences.

```r
#Repeated measures anova with group factor

model2 <- aov_ez(id="subject", dv="Scores",
                 data = dataLong,
                 between = c("Group"),
                 within= c("Week"))



summary(model2)
```

```
Univariate Type III Repeated-Measures ANOVA Assuming Sphericity

             Sum Sq num Df Error SS den Df F value    Pr(>F)
(Intercept) 29414.4      1  11314.4     18 46.7955 2.111e-06 ***
Group         168.2      1  11314.4     18  0.2676  0.611252
Week          962.4      3   7184.2     54  2.4114  0.076812 .
Group:Week   1736.3      3   7184.2     54  4.3503  0.008121 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Mauchly Tests for Sphericity

           Test statistic p-value
Week              0.67745 0.26026
Group:Week        0.67745 0.26026


Greenhouse-Geisser and Huynh-Feldt Corrections
 for Departure from Sphericity

           GG eps Pr(>F[GG])
Week       0.8081    0.09149 .
Group:Week 0.8081    0.01373 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

             HF eps  Pr(>F[HF])
Week       0.942953 0.080905270
Group:Week 0.942953 0.009487549
```

(Top Level) R Script

Fig. (9)

```r
#MANOVA for model2
model2$Anova

#Mean plot for mixed factorial design
afex_plot(model2,
          x="Week",
          trace="Group",
          error="none",
          factor_levels = list(Week = c("Week1",'
                               Group = c("Trainin
```

```
> #MANOVA for model2
> model2$Anova

Type III Repeated Measures MANOVA Tests: Pillai test statistic
            Df test stat approx F num Df den Df    Pr(>F)
(Intercept)  1   0.72220   46.795      1     18 2.111e-06 ***
Group        1   0.01465    0.268      1     18    0.6113
Week         1   0.25093    1.787      3     16    0.1902
Group:Week   1   0.48583    5.039      3     16    0.0120 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

Fig. (10)

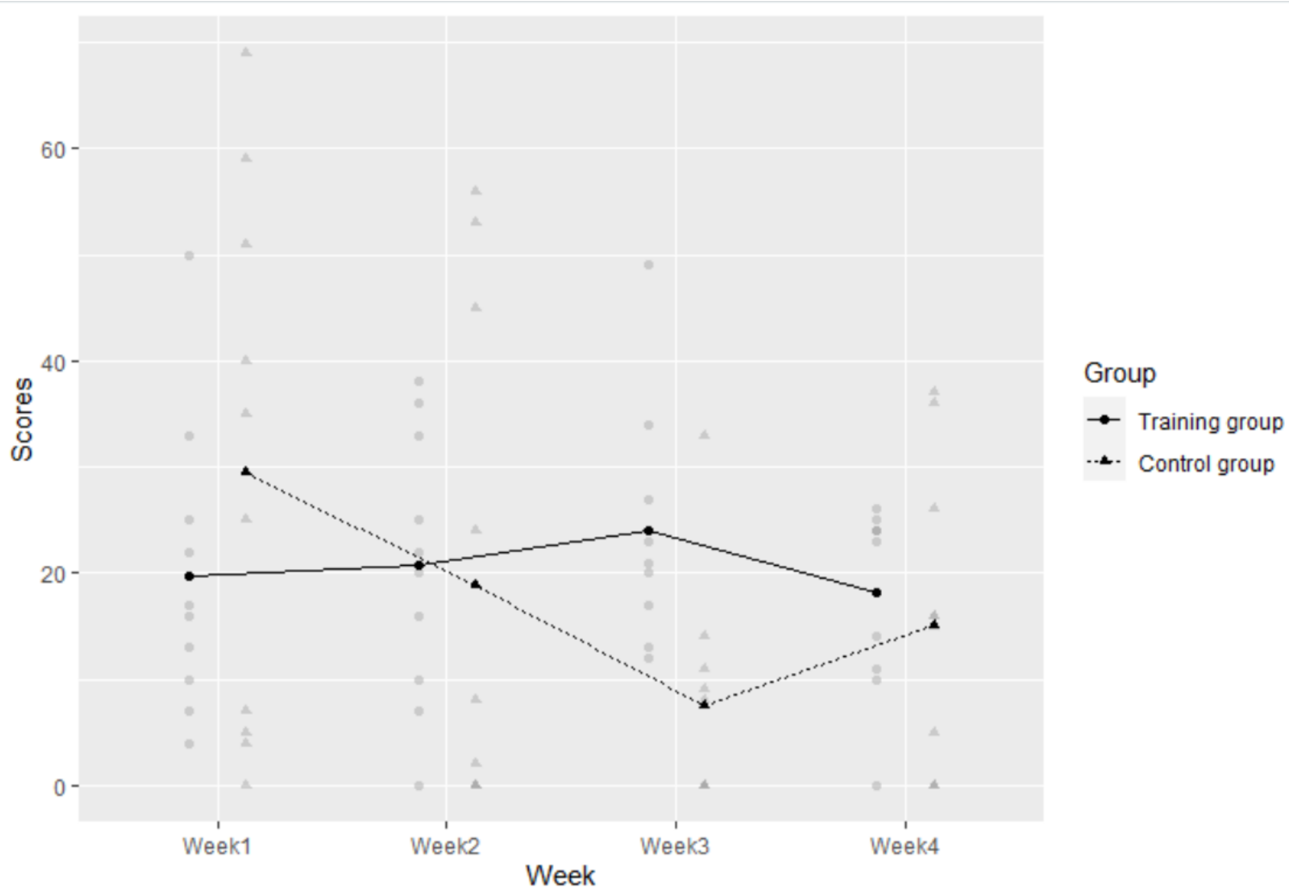

Fig. (11)

After adding both within and between factors, anova model p values are significant and therefore, we can say mean will be different for at least one of the group.

And from the mean plot in, we can see the variation in means for both the groups at different times.

Now, similar to all multiple group design, here also we can perform post-hoc tests. I am using Tukey's HSD here.

```r
#Post-hoc tests
lsmeans(model2,
        pairwise ~Group,
        by = c("Week"),
        adjust ="tukey")
```

```
+          adjust = "tukey")
$lsmeans
Week = X1:
 Group lsmean   SE   df lower.CL upper.CL
 1       19.7 5.07 42.4     9.47     29.9
 2       29.5 5.07 42.4    19.27     39.7

Week = X2:
 Group lsmean   SE   df lower.CL upper.CL
 1       20.7 5.07 42.4    10.47     30.9
 2       18.8 5.07 42.4     8.57     29.0

Week = X3:
 Group lsmean   SE   df lower.CL upper.CL
 1       24.0 5.07 42.4    13.77     34.2
 2        7.5 5.07 42.4    -2.73     17.7

Week = X4:
 Group lsmean   SE   df lower.CL upper.CL
 1       18.1 5.07 42.4     7.87     28.3
 2       15.1 5.07 42.4     4.87     25.3

Warning: EMMs are biased unless design is perfectly balanced
Confidence level used: 0.95

$contrasts
Week = X1:
 contrast estimate   SE   df t.ratio p.value
 1 - 2        -9.8 7.17 42.4  -1.367  0.1788

Week = X2:
 contrast estimate   SE   df t.ratio p.value
 1 - 2         1.9 7.17 42.4   0.265  0.7923

Week = X3:
 contrast estimate   SE   df t.ratio p.value
 1 - 2        16.5 7.17 42.4   2.302  0.0263

Week = X4:
 contrast estimate   SE   df t.ratio p.value
 1 - 2         3.0 7.17 42.4   0.419  0.6777
```

(Top Level) ÷   R Script ÷

Fig. (12)

To determine which group has significant difference, post-hoc test has been used(Tukey test) and from, we can see the p value for group 1 and 2 for week x3 only is 0.0263 which is less than 0.05 and therefore, we can say there is mean difference between both the groups for all weeks except week3.

For R code:

Visit Github