

Class 6: Introduction of Analysis of Variance (ANOVA)

Contents

1	Design Example	2
1.1	Comparing 3 groups	2
2	Analysis of Variance	2
2.1	ANOVAs vs. t-tests	2
2.1.1	Types of One-way ANOVAs	3
2.2	Terms	3
2.3	Types of studies	3
2.4	Logic of ANOVA	3
2.4.1	Logic of F-testing	3
2.4.2	Between-Subjects Variance	4
2.4.3	Within-Subjects Variance	4
2.4.4	F-Ratio Logic	4
3	Analysis Procedures	5
3.1	By Hand Calculation	5
3.1.1	Between-Subject Variance	5
3.1.2	Within-Subject Variance	5
3.1.3	F-value	6
3.1.4	Significance	6
3.2	Helper Table	6
3.3	Effect size	6
3.3.1	Understanding η^2 (or ω^2)	7
3.3.2	Effect Size for Power	8
3.4	By R Functions	8
3.4.1	Build our data frame	9
3.4.2	Check the means and SS calculations with dplyr	9
3.4.3	Run ANOVA	9
4	Assumptions	10
5	APA style Report	12
5.1	Plot ANOVA	12
6	References	12

Note: There is some very complex R code used to generate today's lecture. I have hidden it in the PDF file. If you see `echo=FALSE` inside the RMD file, it means that is the code you are not expected to understand or learn. I will explain the functions you will need to learn.

1 Design Example

Interested in how 'trusting' people feel for their partner after engaging in various levels of joint action. 2 People enter the lab (one is a confederate) and are asked to either a) move chairs independently, b) Move a couch together, and confederate helps c) Move a couch together and confederate hinders. Moving a couch requires the two people to coordinate their action. After the study ends, people rate how trustworthy their partner seems (1-7, with 6 = high trust).

No Joint Action	Helpful Joint Action	Disrupted Joint Action
2	5	2
3	4	3
5	7	3
3	4	1
2	5	1
M = 3	M = 5	M = 2
SS = 6	SS = 6	SS = 4

1.1 Comparing 3 groups

So, to do t-tests, I would have to compare G1 vs. G2, G2 vs. G3, G1 vs. G3. This becomes a problem when I have more 4+ groups. There are statistical reasons we cannot do this, such as each time we do this we risk committing a type I error (A dead Salmon will teach us this lesson in a week). So, we need some way to test all three groups at once to know if at least one group is different from another. So how do I compare three groups (or more) at once? I know let's just fix the t-test formula!

$$t = \frac{M_1 - M_2 - M_3}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2} + \frac{S_p^2}{n_3}}}$$

What is wrong with my fix?

2 Analysis of Variance

- Analysis of variance (ANOVA) is the procedure to test hypotheses with two or treatments
- ANOVA can handle multiple samples while t-tests can only be calculated for two samples
- ANOVA used to decide whether
 - The treatments have no effect
 - Treatments do have differences

2.1 ANOVAs vs. t-tests

ANOVAs are closely related to t-tests, $F = t^2$ (This is why the F-distribution thus has no tails). Hypothesis testing with ANOVAs is just like with t-tests. t-tests are essentially a ratio of [mean difference over noise (error/chance)], and ANOVA is an F ratio (variance caused by you over variance due to noise)

2.1.1 Types of One-way ANOVAs

- Fixed-factor ANOVAs which will be the focus of the rest of the semester
- Random-factor ANOVAs have been replaced by Mixed-Models which I teach a whole course on

2.2 Terms

- FACTOR: Independent or quasi-independent variable (the manipulated variable)
- Factors can have *LEVELS* that make up the factor

For example, you are testing how temperature (factor) affects concentration. You test four groups at four different temperatures (levels) to see what temperature has the strongest effect.

2.3 Types of studies

- Like t-tests, you can do a repeated measures design or independent measures design
 - Repeated measure design is where the same subject is tested multiple times (over the course of time)
 - Independent measures design is where a separate group is used for each factor (we will focus on these for a few weeks)
- Mixed design: a combination of repeated measures and independent measures (we will cover review this design after we learn about the repeated measures and independent measures ANOVA)

2.4 Logic of ANOVA

ANOVA asks is “One of these things not like the other”



H_0 All groups have equal means

H_1 At least 1 group is different from the other

It does not tell which is different and it does not have tails (for reasons you will see below). We will report to how to figure out which is different next class.

2.4.1 Logic of F-testing

- First, we need to analyze the variance (it's in them of the ANOVA right?)

- There two concepts we need to solve

$$F = \frac{S_{Treatment}^2}{S_{Error}^2}$$

- * Between-treatments variance: we will calculate the variance between groups! *This is the variance we **caused**.*
- * Within-treatment variance: the variance within each cell is the error term. *That is the variance due to **chance**.* We need to assume (just like our T-test) that variance in group 1 = group 2 = group 3 (Assumption: Homogeneity of variance: HOV)
- We can rewrite our F-test to be, $F = \frac{S_B^2}{S_W^2}$

Note: ANOVA uses a more technical name for variance mean squared deviation scores (or *MS* for short). Thus, when you see *MS* its **equivalent conceptually** to S^2 . [Root-MS **equivalent conceptually** S]. In both cases, I cannot say $MS = S^2$ for mathematical reasons...

- So, in most people will recognize:

$$F = \frac{MS_B^2}{MS_W^2}$$

2.4.2 Between-Subjects Variance

- Two explanations for between-treatment variance:
 - Treatment effect
 - Chance
 - * Individual differences: what if one person likes it cold and others function better at very warm temps
 - * Experimental error: just by taking the measurement you can cause an error

2.4.3 Within-Subjects Variance

- One explanation for Within-treatment variance [if any only if HOV is meet and you have no confounding or nuisance variables]
 - Noise/chance differences (individual differences) > However, imagine you had different RAs running each condition and they each screwed up the instructions differently (condition is confounded with RA which could affect the mean differences or simply add variance you should have controlled for)

2.4.4 F-Ratio Logic

$$F = \frac{MS_B^2}{MS_W^2}$$

$$F = \frac{Treatment + Noise}{Noise}$$

if Treatment = 0

$$F = \frac{0 + Noise}{Noise} = 1$$

Thus, $F = 1$ you have no treatment effect.

But if Treatment = 3, Noise = 1

$$F = \frac{Treatment + Noise}{Noise} = \frac{3 + 1}{1} = 3$$

Thus, your signal is 3 times your noise!

3 Analysis Procedures

We return to our study data. We will run our analysis first by hand and examine each stage of the calculation. We will next do it in R.

No Joint Action	Helpful Joint Action	Disrupted Joint Action
2	5	2
3	4	3
5	7	3
3	4	1
2	5	1
M = 3	M = 5	M = 2
SS = 6	SS = 6	SS = 4

3.1 By Hand Calculation

Simplified formulas for equal sample sizes.

3.1.1 Between-Subject Variance

Remember that $S^2 = \frac{SS}{n-1} = \frac{\Sigma(X-M)^2}{n-1}$. We can transform this, $MS_B = \frac{SS_B}{k-1} = \frac{n\Sigma(M-G)^2}{df_{BT}}$, where $G = \frac{\Sigma M}{K}$ [Grand mean] and $K = \#$ Levels of the Factor, $n = \#$ of people per cell [Note: This formula works in equal sample sizes only]

3.1.1.1 Calculation

1. $K = 3$
2. $G = \frac{3+5+2}{3} = 3.33$
3. $df_{BT} = 3 - 1 = 2$
4. $n = 5$
5. $SS_{BT} = 5 * [(3 - 3.33)^2 + (5 - 3.33)^2 + (2 - 3.33)^2] = 23.33$
6. $MS_{BT} = \frac{SS_B}{df_B} = \frac{23.33}{2} = 11.67$

3.1.2 Within-Subject Variance

Here we simply calculate the $MS_{WT} = \frac{SS_W}{N-K} = \frac{\Sigma(SS_{cell})}{df_W}$. $SS_W = \Sigma(X - M)^2$ per group and sum it.

3.1.2.1 Calculation

1. $N = n * k = 5 * 3 = 15$
2. $df_W = 15 - 3 = 12$
3. $SS_W = 6 + 6 + 4 = 16$
4. $MS_W = \frac{16}{12} = 1.33$

3.1.3 F-value

$$1. F = \frac{MS_B}{MS_W} = \frac{11.67}{1.33} = 8.75$$

3.1.4 Significance

You will need to look up two things on F-table: df_{Num} & df_{dem} . The degrees of freedom that when into the numerator and the degrees of freedom that go into the denominator. See hand out for tables and this SHINY APP: <https://shiny.rit.albany.edu/stat/fdist/>

The F_{crit} changes with df as in t-test. When $K = 2$, $F_{crit} = t_{crit}^2$

3.2 Helper Table

That is hard to track, so we build a source table to help us track what we are doing

Source	SS	DF	MS	F
Between	SS_B	df_B	MS_B	F
Within	SS_W	df_W	MS_W	
Total	SS_T	df_T		

With the formulas (formal equations)

Source	SS	DF	MS	F
Between	$n \sum_{i=1}^k (M_i - G)^2$	$K - 1$	$\frac{SS_B}{df_B}$	$\frac{MS_B}{MS_W}$
Within	$\sum_{i=1}^K \sum_{j=1}^n (X_{ij} - M_i)^2$	$N - K$	$\frac{SS_W}{df_W}$	
Total	$\sum_{j=1}^N (X_j - G)^2$	$N - 1$		

With the formulas (conceptually simplified)

Source	SS	DF	MS	F
Between	$nSS_{treatment}$	$K - 1$	$\frac{SS_B}{df_B}$	$\frac{MS_B}{MS_W}$
Within	$\sum SS_{within}$	$N - K$	$\frac{SS_W}{df_W}$	
Total	SS_{scores}	$N - 1$		

Note: $SS_B + SS_W = SS_T$ & $df_B + df_W = df_T$

3.3 Effect size

we cannot use Cohen's d , instead we will use eta-squared = percentage of variance the treatment accounts for. Remember SS_B is the variation caused by treatment and SS_T is the variation of Treatment + Noise.

$$\eta^2 = \frac{SS_B}{SS_B + SS_W} = \frac{SS_B}{SS_T}$$

This eta-squared formula for one-way ANOVA is equal to the more modern generalized-eta squared η_g^2 . So R will report η_g^2 (Bakeman, 2005). We will report to these formulas in the future.

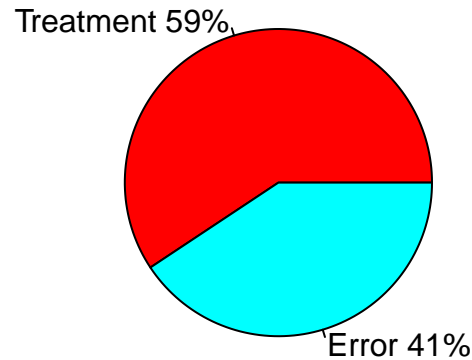
Note: η_g^2 will always yield values larger than population, and there are corrections (but they are not often reported). They are ω^2 and ω_g^2 .

$$\omega^2 = \frac{SS_B - (K - 1)MS_W}{SS_T + MS_W}$$

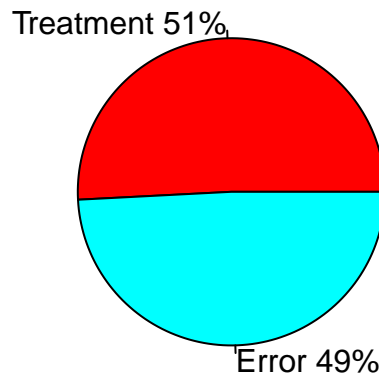
3.3.1 Understanding η^2 (or ω^2)

η^2 is a useful method for understanding how your ANOVA “parsed” your variance. In fact, $\eta^2 = R^2$ from regression (which is how much your line explains the dots). The reason we can get the percentage of variance accounted for is because ANOVA is actually a special case of regression.

Pie Chart of Parsed Variance: Eta-Sq



Pie Chart of Parsed Variance: Omega-Sq



3.3.2 Effect Size for Power

For power analysis, we will need f , which is what Cohen developed (like d for t-tests)

$$f = \sqrt{\frac{K-1}{K} \frac{F}{n}}$$

ES	Value	Size
f	.1	Small effect
f	.25	Medium effect
f	.4	Large effect

$f_{observed} = 1.08$, Large effect!

3.4 By R Functions

The formulas I provided will work when there are equal sample sizes and no missing data. When you have unbalanced cells, choices have to be made about how to parse the variance because ANOVA is a special case of regression. We will use Type III Sums of squares. I will explain in future lectures the different methods to parse the variance.

No Joint Action	Helpful Joint Action	Disrupted Joint Action
2	5	2
3	4	3
5	7	3
3	4	1
2	5	1

No Joint Action	Helpful Joint Action	Disrupted Joint Action
M = 3	M = 5	M = 2
SS = 6	SS = 6	SS = 4

3.4.1 Build our data frame

We can read in the data into R via csv file, or we can hand type the data frame. This is what I have done below.

```
n=5; k=3;
JA.Study<-
  data.frame(SubNum=seq(1:(k*n)),
             Condition=ordered(
               c(rep("No",n),
                 rep("Helpful",n),
                 rep("Disrupted",n))
             ),
             DV=c(2,3,5,3,2,5,4,7,4,5,2,3,3,1,1))
```

You can View(JA.Study) it see as well

3.4.2 Check the means and SS calculations with dplyr

```
library(dplyr)
Means.Table<-JA.Study %>%
  group_by(Condition) %>%
  summarise(N=n(),
            Means=mean(DV),
            SS=sum((DV-Means)^2),
            SD=sd(DV),
            SEM=SD/N^.5)
knitr::kable(Means.Table, digits = 2)
# Note remember knitr::kable makes it pretty, but you can just call `Means.Table`
```

Condition	N	Means	SS	SD	SEM
Disrupted	5	2	4	1.00	0.45
Helpful	5	5	6	1.22	0.55
No	5	3	6	1.22	0.55

Good are means and SS match, so I data entered it correctly.

3.4.3 Run ANOVA

afex is a package developed to allow people to do SPSS-like ANOVA in R (type III: R defaults to Type I, which will talk about later). It had many output formats, but for us the most useful are (Anova, nice). These output formats allows to see different things and use the objects it creates in different ways.

3.4.3.1 Anova style output

- This style is close to our ANOVA source table, but instead its formatted like a regression F-source table.

- You can ignore the intercept line (we will cover that next semester).
- Condition = SS_B and the name will always match your IV name.
- Residual = Error. Residual in regression means leftover (and this is the error term: chance)
- There is no total (you can do that with addition)

```
library(afex)

ANOVA.JA.Table<-aov_car(DV~Condition + Error(SubNum),
                        data=JA.Study, return='Anova')
ANOVA.JA.Table

## Anova Table (Type III tests)
##
## Response: dv
##           Sum Sq Df F value    Pr(>F)
## (Intercept) 166.667  1  125.00 1.059e-07 ***
## Condition    23.333  2    8.75 0.004531 **
## Residuals    16.000 12
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.4.3.2 Nice style output

- This style reports everything you need to report in APA format! That's why it's called **nice**, cause it's nice to you!

```
library(afex)

Nice.JA.Table<-aov_car(DV~Condition + Error(SubNum),
                      data=JA.Study, return='nice')
Nice.JA.Table

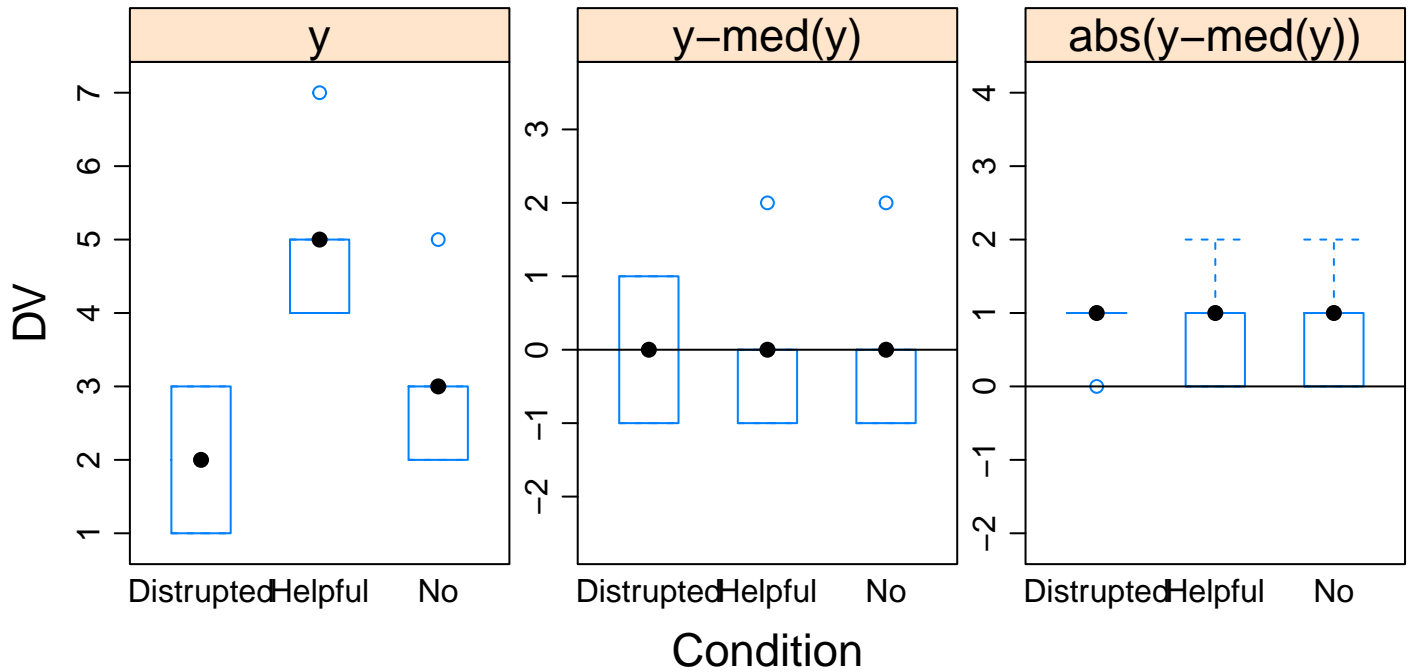
## Anova Table (Type 3 tests)
##
## Response: DV
##      Effect    df  MSE      F ges p.value
## 1 Condition 2, 12 1.33 8.75 ** .59    .005
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1
```

4 Assumptions

We assume HOV and normality. We can test for HOV using Brown-Forsythe formula. Brown-Forsythe formula subtracts the median within each cell and recalculated the ANOVA. If there is not variance imbalance we should see a result very similar to the middle plot below (the pvalue > .05: non-significant). If violate this assumption, there is special ANOVA called the Brown-Forsythe ANOVA (but no one reports it).

Using the HH package, we can visualize our violations. As you can see the middle plot show no deviation.

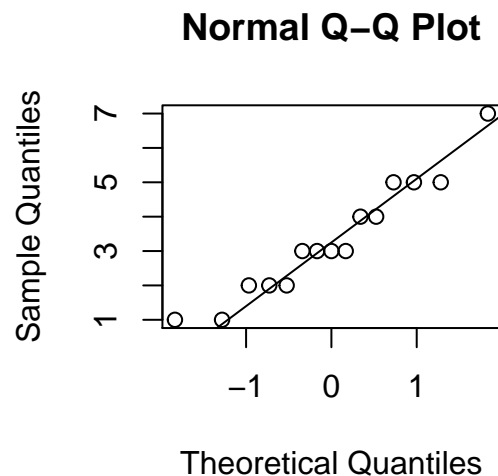
```
library(HH)
hov(DV~Condition, data=JA.Study)
hovPlot(DV~Condition, data=JA.Study)
```



```
##
##  hov: Brown-Forsyth
##
## data:  DV
## F = 5.7778e-32, df:Condition = 2, df:Residuals = 12, p-value = 1
## alternative hypothesis: variances are not identical
```

Testing for normality with such a small sample does not make sense. In general normality tests tend to be overly sensitive (Shapiro-Wilk test). Often we examine QQ plots of the DV. If the data points are very different from the line, it suggests violations of normality.

```
qqnorm(JA.Study$DV)
qqline(JA.Study$DV)
```



5 APA style Report

$$F(df_n, df_d) = x.xx, p = 0.xx, \eta_g^2 = 0.xx$$

A one-way between subjects ANOVA was conducted to compare the effect of joint action on feelings of trust in no, helpful, and disrupted joint action. There was a significant effect of joint action on feelings of trust, $F(2, 12) = 8.75, p < 0.01, \eta_g^2 = 0.59$.

5.1 Plot ANOVA

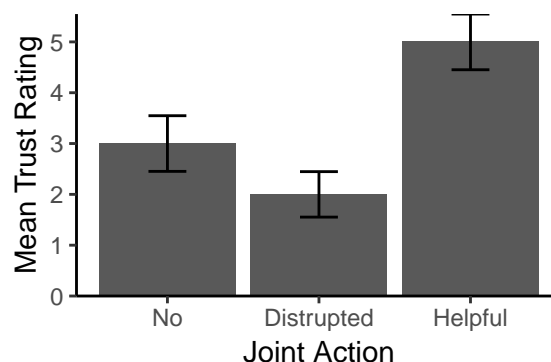
Normally we plot bar graphs with 1 SEM unit. I used the object I created `Means.Table` that has the means. Also, I want to re-set the order using the `factor` command and set the order I want. We will apply some fancy code to make our plot more APA!

```
library(ggplot2)

Means.Table$Condition<-factor(Means.Table$Condition,
                              levels = c("No","Distruped","Helpful"),
                              labels = c("No","Distruped","Helpful"))

Plot.1<-ggplot(Means.Table, aes(x = Condition, y = Means))+
  geom_col()+
  scale_y_continuous(expand = c(0, 0))+ # Forces plot to start at zero
  geom_errorbar(aes(ymax = Means + SEM, ymin= Means - SEM),
               position=position_dodge(width=0.9), width=0.25)+
  xlab('Joint Action')+
  ylab('Mean Trust Rating')+
  theme_bw()+
  theme(panel.grid.major=element_blank(),
        panel.grid.minor=element_blank(),
        panel.border=element_blank(),
        axis.line=element_line(),
        legend.title=element_blank())

Plot.1
```



6 References

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. Behavior research methods, 37(3), 379-384.