

ANOVA:  
Analysis of Variance

# An example ANOVA problem

---

25 individuals split into three between-subject conditions: A, B and C

- A: 5,6,6,7,7,8,9,10 [8 participants, mean: 7.25]
- B: 7,7,8,9,9,10,10,11 [8 participants, mean: 8.875]
- P: 7,9,9,10,10,10,11,12,13 [9 participants, mean: 10.11]

Are the differences between the conditions significant?

# What does ANOVA do?

---

ANOVA tests the following hypotheses:

- $H_0$  (null hypothesis): The means of all the groups are equal.
- $H_a$ : Not all the means are equal
  - doesn't say how or which ones differ.
  - Can follow up with “multiple comparisons”

# Notation for ANOVA

- $n$  = number of individuals all together
- $i$  = number of groups
- $\bar{x}$  = mean for entire data set is

Group  $i$  has

- $n_i$  = # of individuals in group  $i$
- $x_{ij}$  = value for individual  $j$  in group  $i$
- $\bar{x}_i$  = mean for group  $i$
- $s_i$  = standard deviation for group  $i$

# How ANOVA works

ANOVA measures two sources of variation in the data and compares their relative sizes

- variation BETWEEN groups
  - for each data value look at the difference between its group mean and the overall mean

$$(\bar{x}_i - \bar{x})^2$$

- variation WITHIN groups
  - for each data value we look at the difference between that value and the mean of its group

$$(x_{ij} - \bar{x}_i)^2$$

# F-score

---

- The ANOVA F-statistic is a ratio of the Between Group Variaton divided by the Within Group Variation:

$$F = \frac{\textit{Between}}{\textit{Within}}$$

- A large F is evidence *against*  $H_0$ , since it indicates that there is more difference between groups than within groups.

# ANOVA Output for Our Example

Analysis of Variance summary

Source	DF	SS	MS	F	P
Treatment [between groups]	2	34.74	17.37	6.45	0.006
Error [within groups]	22	59.26	2.69		
Total	24	94.00			

# ANOVA Output for Our Example

Analysis of Variance summary

Source	DF	SS	MS	F	P
Treatment [between groups]	2	34.74	17.37	6.45	0.006
Error [within groups]	22	59.26	2.69		
Total	24	94.00			

1 less than # of groups

# of data values - # of groups

(df for each group added together)

1 less than # of individuals

# ANOVA Output for Our Example

Analysis of Variance summary

Source	DF	SS	MS	F	P
Treatment [between groups]	2	34.74	17.37	6.45	0.006
Error [within groups]	22	59.26	2.69		
Total	24	94.00			

$$\sum (x_{ij} - \bar{x}_i)^2$$

$$\sum (x_{ij} - \bar{x})^2$$

$$\sum (\bar{x}_i - \bar{x})^2$$

SS = "sum of squares"

$$\text{MSG} = \text{SSG} / \text{DFG}$$
$$\text{MSE} = \text{SSE} / \text{DFE}$$

# ANOVA Output for Our Example

Analysis of Variance summary

Source	DF	SS	MS	F	P
Treatment [between groups]	2	34.74	17.37	6.45	0.006
Error [within groups]	22	59.26	2.69		
Total	24	94.00			

$$34.74 / 2 = 17.37$$

$$F = \text{MSG} / \text{MSE}$$

P-value  
comes from  
 $F(\text{DFG}, \text{DFE})$

# Post-hoc analysis

---

- ANOVA indicates that the groups do not all appear to have the same means... what next? How do we know what the differences really are?
- If we only had two groups, then we're done, we know the difference between them is significant.
- If we have three or more groups, then a post hoc test is needed to determine which groups are significantly different from each other

A: 5,6,6,7,7,8,9,10

[8 participants, mean: 7.25]

B: 7,7,8,9,9,10,10,11

[8 participants, mean: 8.875]

P: 7,9,9,10,10,10,11,12,13

[9 participants, mean: 10.11]

# Post-hoc analysis

---

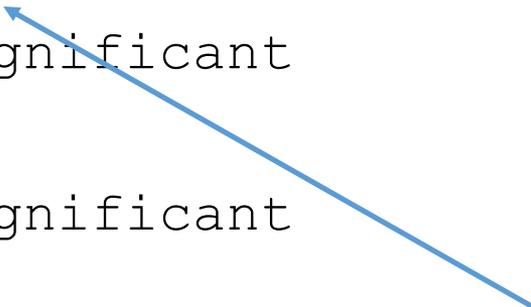
- Multiple post hoc analysis methods exist
- We most commonly see the Tukey test
- Results for our example dataset:

HSD[.05]=2.02; HSD[.01]=2.61

M1 vs M2      nonsignificant

M1 vs M3      P<.01

M2 vs M3      nonsignificant



HSD = the absolute (unsigned) difference between any two sample means required for significance at the designated level.

# Assumptions of ANOVA

---

- The distribution of data in each group is approximately normal
  - check this by looking at histograms and/or normal quantile plots
  - can handle some non-normality, but not severe outliers
- Standard deviations of each group are approximately equal
  - rule of thumb: ratio of largest to smallest sample st. dev. must be less than 2:1

# Our case study...

---

- Our case study has many similarities to the above example, but in that case it's a two-way ANOVA. I leave it to you to decide whether that is the appropriate test and what conclusions can be drawn from it based on the way it was conducted.

ANOVA Summary					
A = row variable (Mobile Robot / No Mobile Robot)					
B = column variable (No Social / Social)					
Subj = subjects					
Source	Sum of Squares	df	Mean Square	F	p
<u>Subjects</u>	134.4	9			
<u>Within Subjects</u>					
A	592.9	1	592.9	70.5833	<.0001
Subj x A	75.6	9	8.4		
B	115.6	1	115.6	13.7076	0.004902
Subj x B	75.9	9	8.4333		
A x B	160	1	160	17.6686	0.002295
Subj x A x B	81.5	9	9.0556		
TOTAL	1235.9	39			