



---

# **Module 4 Lecture - One-way ANOVA and Multiple Comparison Procedures**

Analysis of Variance

---

Quinton Quagliano, M.S., C.S.P

Department of Educational Psychology

## Table of Contents

<b>1</b>	<b>Overview and Introduction</b>	<b>2</b>
1.1	Objectives . . . . .	2
1.2	Introduction . . . . .	2
<b>2</b>	<b>The F Distribution</b>	<b>3</b>
2.1	Introduction . . . . .	3
2.2	Additional Facts About the F Distribution . . . . .	4
<b>3</b>	<b>The F Ratio</b>	<b>5</b>
3.1	Introduction . . . . .	5
3.2	Between & Within Variation . . . . .	6
3.3	Calculation of the F-ratio/statistic . . . . .	6
3.3.1	Overall F Calculation . . . . .	6
3.3.2	Between-group Calculations . . . . .	6
3.3.3	Within-group Calculations . . . . .	7
3.4	Conclusions with the F-ratio/statistic . . . . .	8
<b>4</b>	<b>One-way ANOVA</b>	<b>9</b>
4.1	Introduction . . . . .	9
4.2	Assumptions . . . . .	9
4.3	Null and Alternative Hypotheses . . . . .	10
4.4	Extensions . . . . .	11
4.5	Effect Size for One-way ANOVA . . . . .	12
4.5.1	Eta-squared . . . . .	12
4.5.2	Omega-squared . . . . .	13
<b>5</b>	<b>Mean Comparisons in One-way ANOVA</b>	<b>13</b>
5.1	Introduction . . . . .	13
5.1.1	Types of Contrasts . . . . .	14
5.1.2	Concept of Weights . . . . .	14
<b>6</b>	<b>Conclusion</b>	<b>15</b>
6.1	Recap . . . . .	15
	<b>Key Terms</b>	<b>16</b>

# 1 Overview and Introduction

## 1.1 Objectives

- Interpret the F probability distribution as the number of groups and the sample size change.
- Discuss one use for the F distribution: one-way ANOVA
- Conduct and interpret one-way ANOVA.
- Appreciate the ANOVA as another test in your toolbox alongside the other tests for dealing with categorical and continuous outcome variables
- Understand the F-distribution as another practical distribution for deriving inferences and conclusions
- Understand the types of effect size usually employed for ANOVA
- Introduce the different mean comparison strategies in a One-way ANOVA, to be followed up on later in next lecture

## 1.2 Introduction

 Discuss: When we needed to compare the averages of two groups on a continuous variable, what statistical test would we use?

- While we've already covered how to test \_\_\_\_\_ between two groups (even with an \_\_\_\_\_ violation!), now we need to consider the scenario with \_\_\_\_\_ or more groups on some outcome
  - Examples:
    - \* Comparing class outcomes for those in hybrid, in-person, or online sections
    - \* Comparing students on three different tracks in high school: regular, advanced, or remedial
- When we are comparing three \_\_\_\_\_ groups on some numeric, outcome, we will employ the one-way ANOVA, standing \_\_\_\_\_ for **Analysis of Variance (ANOVA)**
  - ANOVA is a \_\_\_\_\_ family of techniques with many applications and extensions - many of which we will cover in this class!

- We are looking at the one-way ANOVA as a \_\_\_\_\_ use of this technique

**!** Important

'ANOVA is a t-test, but for 3 categories' is actually a pretty fair way to describe this, as they are actually equivalent if/when comparing only two groups.

- Why not use a bunch of \_\_\_\_\_ across the 3 groups instead of this new complicated method?
  - Problem: It raises our Type I error rate
  - This is something done as part of ANOVA in \_\_\_\_\_ or \_\_\_\_\_ testing

**!** Discuss: Describe what Type I errors are, in your own words

- Previous tests all made use of specific, practical \_\_\_\_\_ for determining significance:
  - z-tests → normal distribution
  - t-tests → t-distribution
  - $\chi^2$  tests →  $\chi^2$  distribution
  - The ANOVA family will introduce a new distribution: the \_\_\_\_\_
- We'll start by describing the F-distribution itself, then the F-ratio/statistic that we compute, and then finally the \_\_\_\_\_ within the One-way ANOVA

## 2 The F Distribution

### 2.1 Introduction

- The **F-distribution** was developed by Sir Ronald Fisher and is a unique distribution applied often for the \_\_\_\_\_ family of statistical tests
  - In \_\_\_\_\_ it is given as  $F \sim F_{df(num), df(denom)}$  where:
  - $df(num) \rightarrow df_{between}$  and  $df(denom) \rightarrow df_{within}$

- Example:  $F \sim F_{2,24}$  has:

- \*  $df_{between} = 2$
- \*  $df_{within} = 24$

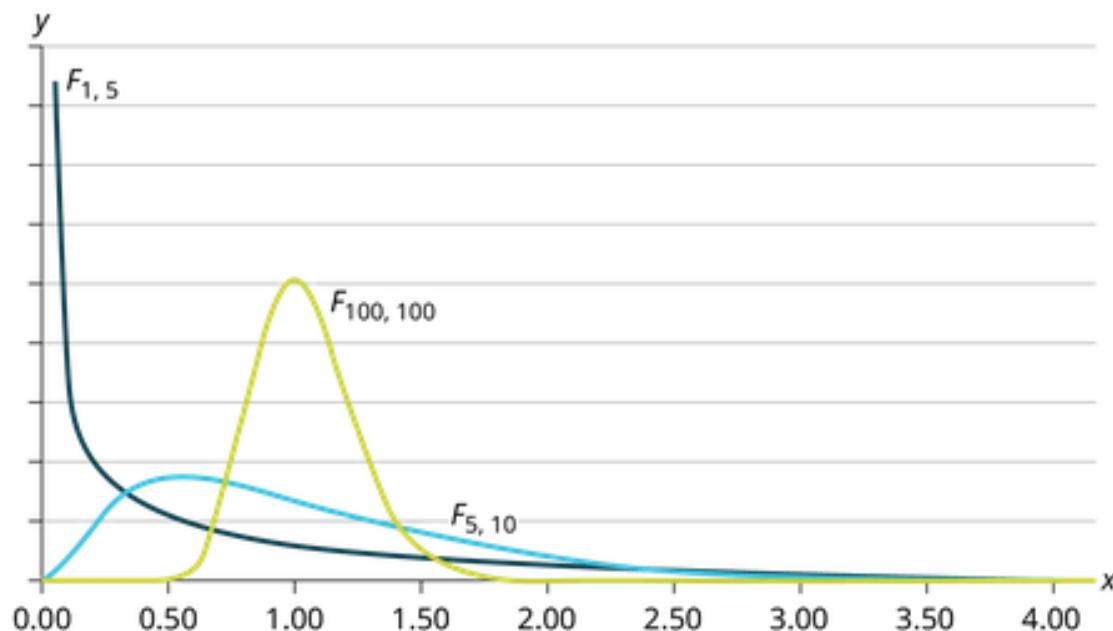
### ! Important

Hold on just a second, \*two\* different degrees of freedom? Yes! We'll discuss more about why that is in our discussion on the F-ratio itself

- The F-distribution actually consists of values that are \_\_\_\_\_ of the t-distribution, and is derived from the t-distribution
  - How exactly Fisher proved that is way beyond the scope of this class

► Discuss: Recall, what exactly is the distinction that makes the t-distribution different and useful over the plain normal distribution? The F-statistic benefits from this same!

## 2.2 Additional Facts About the F Distribution



- The curve is not \_\_\_\_\_ but skewed to the right.
  - Thus, it has a more similar look to the \_\_\_\_\_ distribution, compared to the t-distribution
- There is a different curve for each set of \_\_\_\_\_.
  - Once again, similar to chi-squared
- The F statistic is always greater than or equal to zero.

**Discuss:** After you work through the F-ratio section and calculation below, return to this section and try to explain, using the formula, why this F is always greater than or equal to zero

- As the degrees of freedom for the numerator and for the denominator get larger, the curve approximates the \_\_\_\_\_ distribution.
  - Remember, the  $df_{between} \rightarrow$  numerator
  - The  $df_{within} \rightarrow$  denominator

## 3 The F Ratio

### 3.1 Introduction

- The **F-ratio** is just another name for the F statistic that \_\_\_\_\_ from the test we run
  - The F-ratio/statistic we get is a ratio of:
    - \* The variances \_\_\_\_\_ samples
    - \* The variances \_\_\_\_\_ samples
    - \* Its formula is given as:

$$F = \frac{MS_{between}}{MS_{within}}$$

- Well describe what the  $MS$  means here in a second

### 3.2 Between & Within Variation

- For **Variances between samples**, we are considering how much \_\_\_\_\_ there is between the 3 or more groups we are comparing to one another
  - This may also be known as “variation due to \_\_\_\_\_” or “\_\_\_\_\_ variation”
  - This will be represented \_\_\_\_\_ as the  $MS_{between}$  or the **mean square**
  - This is the \_\_\_\_\_ of the F-ratio formula
- In the case of **Variances within samples**, we example how much variation there is within \_\_\_\_\_ of the groups we are comparing to one another
  - This may also be known as “variation due to \_\_\_\_\_” or “\_\_\_\_\_ variation”
  - This will be represented as  $MS_{within}$  in the following formula
  - This is the \_\_\_\_\_ of the F-ratio formula

### 3.3 Calculation of the F-ratio/statistic

- We are *not* going to calculate F by hand, as it is very time-consuming, albeit plenty possible
  - However, you should still follow along and make sure you can see the flow of information as we substitute numbers in these calculations

#### 3.3.1 Overall F Calculation

$$F = \frac{MS_{between}}{MS_{within}}$$

#### 3.3.2 Between-group Calculations

$$MS_{between} = \frac{SS_{between}}{df_{between}}$$

$$df_{between} = k - 1$$

$$SS_{between} = \sum \left[ \frac{(s_j)^2}{n_j} \right] - \frac{(\sum s_j)^2}{n}$$

$$SS_{total} = \sum x^2 \cdot \frac{(\sum x)^2}{n}$$

- Where

- $SS_{between}$  : Sum of squares between groups
- $SS_{total}$  : Sum of squares total
- $df_{between}$  : degrees of freedom between groups
- $k$  : the number of groups
- $n$  : total sample size
- $n_j$  : the size of  $j^{th}$  group
- $s_j$  : sum of values of  $j^{th}$  group

### 3.3.3 Within-group Calculations

$$MS_{within} = \frac{SS_{within}}{df_{within}}$$

$$df_{within} = n - k$$

$$SS_{within} = SS_{total} - SS_{between}$$

- Where

- $SS_{within}$  : Sum of squares between groups
- $SS_{between}$  : Sum of squares between groups
- $SS_{total}$  : Sum of squares total
- $df_{within}$  : degrees of freedom between groups
- $n$  : total sample size

### 3.4 Conclusions with the F-ratio/statistic

**Discuss:** Review: describe the general processes of hypothesis testing, starting with setting the null and alternative hypotheses and ending with a rejection or retaining of the null hypothesis. Try and use the words 'rare event' at some point in your explanation.

- The F statistic, which we derive from the formula for whatever test we are using, is actually best described as a \_\_\_\_\_ or fraction (we'll return to this ratio/statistic in a bit!)
  - Much like previous test statistics, we use this to determine if our results are \_\_\_\_\_ enough to describe our results as unlikely if the null hypothesis is \_\_\_\_\_
  - As usual, the F statistic has a corresponding \_\_\_\_\_ that tell us the probability of results due to chance under the null hypothesis

**?** What value do we compare p against to determine if we can reject the null hypothesis?

- A) Alpha
- B) Omega
- C) Confidence Level
- D) Beta

Explanation:

- For \_\_\_\_\_ testing of the F-ratio, our one-way ANOVA test will always be \_\_\_\_\_ -tailed due to the ratio nature of statistic, with larger numbers suggest greater variation between groups relative to the variation within groups!
  - Put another way, we are trying to see if it is \_\_\_\_\_ enough out on the right tail to say that it is significant!

🔊 Discuss: Which other distribution also always had a right-tailed application? Explain why that is using the relevant formula for that statistic.

## 4 One-way ANOVA

### 4.1 Introduction

#### ❗ Important

In the prior calculations, we covered the mathematical computation in a one-way ANOVA, but this next part will cover the more conceptual side of things.

- The goal of the one-way ANOVA is to determine if there are \_\_\_\_\_ differences between \_\_\_\_\_ group means
  - This is done by examining the \_\_\_\_\_ of the groups (as we previously discussed as part of [The F.Ratio](#))
- This is how we will practically use the F-distribution

### 4.2 Assumptions

- Much like the other tests, we need to be mindful of several \_\_\_\_\_ that underline this statistical test
- These assumptions are:
  - Each population from which a sample for each group is taken is assumed to be \_\_\_\_\_.
  - All samples are \_\_\_\_\_ selected and independent.
  - The populations are assumed to have \_\_\_\_\_ standard deviations (or variances).

- The factor is a \_\_\_\_\_ variable.
- The response is a \_\_\_\_\_ variable.

### 4.3 Null and Alternative Hypotheses

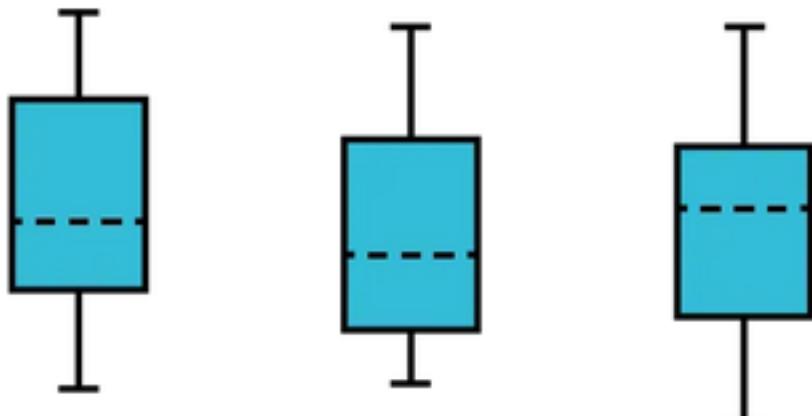
- In a one-way ANOVA, the \_\_\_\_\_ hypothesis is as follows:
  - $H_0 : \mu_1 = \mu_2 = \mu_3 \dots = \mu_k$
  - $H_A$  : At least two of the group means  $\mu_1, \mu_2, \mu_3, \dots, \mu_k$  are not equal

#### ! Important

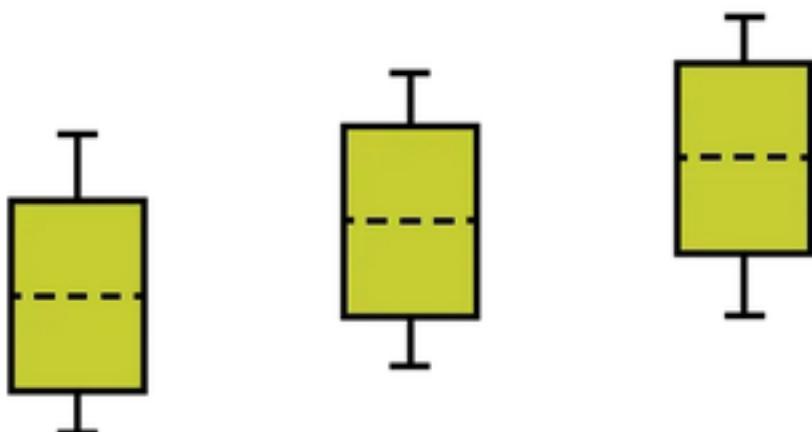
Be \*very\* careful in how you phrase the alternative hypothesis for this test, it looks a little bit different than the other tests we've covered before!

- Effectively, a statistically significant F-statistic only tells us that a difference exists \_\_\_\_\_, but not where that difference lies
  - More explanation on this in the [Extensions](#) section!

#### ► Discuss: What is the name or type of the following plots?



(a)



(b)

#### 4.4 Extensions

- Of course, this is only the tip of the iceberg
- One question that the ANOVA alone \_\_\_\_\_ answer: which of the groups are different?
  - If we have a \_\_\_\_\_ one-way ANOVA test, then we can follow up with **post-hoc testing**, which can tell us which \_\_\_\_\_ exactly are different than one another

---

*"I don't mind not knowing. It doesn't scare me." — Richard P. Feynman*

- There is also the **two-way ANOVA**, which allows us to compare two or more variables and their combined and interacting impact on a single dependent variable in two-way Analysis of Variance.
  - While beyond the scope of this class, \_\_\_\_\_ are an incredibly important part of a lot of more complicated studies
- Finally, there is the **repeated-measures ANOVA**

► Discuss: Which of the t-tests also dealt with comparing two sets of measurements, taken at two separate time points, on the same sample?

## 4.5 Effect Size for One-way ANOVA

- There are two measures of **effect size**: Eta-squared and Omega-squared

### 4.5.1 Eta-squared

- The first effect size estimate is **eta-squared**, or  $\eta^2$ 
  - A measure of effect size for one-way ANOVA to describe \_\_\_\_\_ of variance in the dependent variable explained by the independent variable in this sample

$$\eta^2 = \frac{SS_{Between}}{SS_{Total}}$$

Discuss: What other effect size coefficient is also described as 'variance explained in one variable by another'. Hint: think of comparisons between two continuous numeric variables

#### 4.5.2 Omega-squared

- The second effect size estimate is **omega-squared** or  $\omega^2$ 
  - A measure of effect size for one-way ANOVA to describe proportion of variance in the dependent variable explained by independent variable in the target \_\_\_\_\_; \_\_\_\_\_ effect size estimate

$$\omega^2 = \frac{SS_{Between} - (k - 1)(MS_{within})}{SS_{total} + MS_{within}}$$

##### Important

Because omega-squared tries to speak to the \*population\*, I generally prefer it as an effect size indicator over eta-squared

## 5 Mean Comparisons in One-way ANOVA

### 5.1 Introduction

- One of the core \_\_\_\_\_ of an ANOVA test, is that it doesn't provide any information by itself on \_\_\_\_\_ groups are different than one another
  - To do these comparisons, we use **contrast/comparison tests**

- Using these, we can test \_\_\_\_\_ hypotheses about differences between our 3 or more groups
- However, you should only pursue doing mean comparisons *if* your overall one-way ANOVA is statistically significant.

! Important

Think about what our null hypothesis is under ANOVA. Very limited, right?  
Contrasts and comparisons are how we can move past that

### 5.1.1 Types of Contrasts

- We have two types of contrasts:
  - A Priori / Planned / Orthogonal Contrasts - Planned \_\_\_\_\_ of time
  - A Posteriori / Post-hoc - Performed \_\_\_\_\_ ANOVA
- Both of these types of contrasts tests make the same assumptions as the ANOVA itself

! Important

Planned comparisons are more powerful (less stringent) than post hoc comparisons which are typically more conservative. For this reasons, I'd usually say it's best to try to make predictions prior to analyzing the data.

### 5.1.2 Concept of Weights

- For contrast tests, **weights** are \_\_\_\_\_ to means (weighted means) to reflect the \_\_\_\_\_.
  - The means assumed to have \_\_\_\_\_ values are assigned the \_\_\_\_\_ weights
- The weights are somewhat \_\_\_\_\_, but they have to satisfy certain conditions:
  - The weights used in a comparison must sum to \_\_\_\_\_. Put another way, they are all \_\_\_\_\_ to one another.
  - Every mean \_\_\_\_\_ to have a weight even if it is zero
  - Larger weights represent \_\_\_\_\_ hypothesized means, and means that are hypothesized to be equal should be assigned the weights \_\_\_\_\_

- Example: All of the following weights are for the same hypothesis where groups 2 and 3 are equal, and group 1 is more than each of them

G1	G2	G3
2	-1.0	-1.0
1	-0.5	-0.5
4	-2.0	-2.0

### ! Important

With all of this said, contrast testing, just like normal inferential testing, should never be careless!

Discuss: Write a set of weight that would suffice the conditions that G1 is greater than G2 is greater than G3.

## 6 Conclusion

### 6.1 Recap

- We can think of a one-way ANOVA as being similar to an independent-samples t-test, but for more than 2 groups. Just note the differences in the null/alternative hypotheses setup
- There are some important extensions and other applications of the ANOVA that we did not comprehensively cover here. However, they do still employ the F-distribution and have the underlying focus on comparing variances to make conclusions about differences between groups. We will cover those later on!
- The F-distribution adds to the other practical distributions employed as part of hypothesis testing, such as the t-distribution and chi-squared distribution (which you learned about in EDPS-641). Just like those, it allows us to conclude whether a null hypothesis can be rejected or retained.

- On top of performing a One-way ANOVA, we can also follow it up with various **Mean Comparisons in One-way ANOVA!** These help us determine where our differences exists between our groups, and can offer important insights on top of our one-way ANOVA results. We continue working through these in next lecture.

## Key Terms

### A

**Analysis of Variance (ANOVA)** A family of statistical tests used to compare variances as a way to detect differences between more than two groups [2](#)

### C

**contrast/comparison tests** A method to compares means against means in ANOVA [13](#)

### E

**effect size** A measure of the magnitude of an effect [12](#)

**eta-squared** A measure of effect size for one-way ANOVA to describe proportion of variance in the dependent variable explained by the independent variable in this sample [12](#)

### F

**F-distribution** A distribution developed by Ronald Fisher, expressed as a ratio, used commonly for the ANOVA family of tests [3](#)

**F-ratio** Another name for the F-statistic derived from an ANOVA test, based upon the ratio of the variances between the samples to the variances within the samples [5](#)

### M

**mean square** The computational representation of variation between groups [6](#)

### O

**omega-squared** A measure of effect size for one-way ANOVA to describe proportion of variance in the dependent variable explained by independent variable in the target population; unbiased effect size estimate [13](#)

### P

**post-hoc testing** Statistical testing done after a significant ANOVA to tease apart where exactly differences lie between the groups [11](#)

**R**

**repeated-measures ANOVA** A type of ANOVA used when measurements are taken on the same sample 3 or more times, to compare against each other at different time points [12](#)

**T**

**two-way ANOVA** A type of ANOVA that allows for testing two independent (grouping) variables impact on a single dependent (outcome) variable [12](#)

**V**

**Variances between samples** How much differences there are between the 3 or more groups that we are comparing [6](#)

**Variances within samples** How much variances occur within each of the groups we are comparing against one another [6](#)

**W**

**weights** Relative amounts assigned to describe the relationships between variables in contrast testing [14](#)

*The instructor-provided glossary may not include all terms worth memorizing, make sure you consider using the vocabulary list in your book and your own judgment to make sure you have all relevant terms*