Unity Environmental University - Math Hub
ANOVA Notes

ANOVA (Analysis of Variance) is a statistical method used to compare means among two or more groups to determine if there are significant differences. It works by analyzing the variance within groups and between groups, partitioning the total variance into components attributable to different sources.

The most common types are one-way ANOVA, used for one independent variable, and two-way ANOVA, used when there are two factors or independent variables. ANOVA helps determine whether the observed differences in means are due to the independent variables or simply random chance.

If you were only comparing the means of two groups, then it would be appropriate to employ a t-Test. When you have more than one group means to compare, it might be tempting to employ a t-Test between each group. However, this approach will lead to incorrectly rejecting the null hypothesis (called a Type 1 Error). When you are comparing the means of several groups, the null hypothesis to be tested will be that the means of all groups are equal. Only the ANOVA test provides you with the tools you need to compare means across multiple groups without making a Type 1 Error.

Hypothesis Testing via ANOVA

The **null hypothesis (H_0)** for an ANOVA states that there is **no significant difference** between the means of the groups being compared. In other words, all group means are equal. Conversely, the **alternative hypothesis (H_a)** for an ANOVA will be that all group means are not equal. Here is how these hypothesis can represented symbolically:

H_o: $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$
H_a: $\mu_1 \neq \mu_2 \neq \mu_3 \neq \ldots \neq \mu_k$

Where μ represents mean of a sample and there k sample (or treatments). The number of groups can also be referred to as "levels."

One-Way ANOVA

A **one-way ANOVA** tests the effect of a single independent variable (factor) on a dependent variable, comparing means across different levels of that factor to see if they are significantly different. For example, comparing test scores between three teaching methods.

Consider the following example data set. An experimenter wants to determine which of four diets should be fed to chickens being raised for market. Each of the four diets is fed to six chickens from age 2 weeks to age 8 weeks. The four samples represent the weight in ounces gained by the chickens during the feeding period.

Sample 1	Sample 2	Sample 3	Sample 4
37	49	33	41
42	38	34	48
45	40	40	40
49	39	38	42
50	50	47	38
45	41	36	41

Since you are only testing the potential effect of one treatment here, diet, then you can subject this data set to a one-way ANOVA to see if any of the diets has a greater impact than the others. The following represents the output from a one-way ANOVA performed on this dataset.

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	142.458333	3	47.4861111	2.16091518	0.12445576	3.09839122
Within Groups	439.5	20	21.975			
Total	581.958333	23				

Just like a t-score in a t-test, the ANOVA uses what is called an F distribution. When performing an ANOVA, an F critical is established based upon the degrees of freedom and alpha chosen. By default, our level of significance is set to α = .05. For this particular dataset, the F critical value was established to be 3.09 while the F calculated was determined to be 2.16, based upon the comparison of F values, the null hypothesis will not be rejected. However, similar to a t-test, the ANOVA provides a p-value as output. In this case, the p-value = 0.124 which is larger than α = .05, which indicates no statistical significance, so you will fail to reject the null hypothesis that the means are equal from this result.

Recall the following:
p - value > α   *Not statistically significant, fail to reject the null hypothesis.*
p - value < α   *Statistically significant, reject the null hypothesis.*

It is also worth noting the following:

- Large F means small p-value.
- Small F means large p-value.

The F statistic can be calculated from the ANOVA table as follows:

$$F = \frac{MS\ between\ groups}{MS\ within\ groups}$$

Using our ANOVA results, the F statistics can be found to be:

$$F = \frac{MS\ between\ groups}{MS\ within\ groups} = \frac{47.4861111}{21.975} = 2.16091518$$

As mentioned above, the F critical value is determined from the degrees of freedom and the established level of statistical significance (typically .05 meaning that the researcher accepts that there is a 5% chance the null hypothesis will be rejected when it shouldn't).

In an ANOVA, the *degrees of freedom between groups* is calculated based on the number of groups (k; also called levels or treatments) being compared: df = k - 1
The degrees of freedom between groups represents how much the group means are free to vary relative to the grand mean. By subtracting 1 from the total number of groups, you're accounting for the constraint that the sum of deviations of group means from the overall mean must equal zero. For example, if you have 4 groups (e.g., 4 different treatments), the degrees of freedom between groups would be: df = 4 - 1 = 3. That's what we have for the *df between groups* with this example.

The ANOVA table also typically includes the *degrees of freedom within groups*. Also known as "error" it is determined from N - k (total number of observations minus the number of groups). In our example, there are 24 observations, so N = 24, then N - k = 24 - 4 = 20 is the df within groups. Additionally, the table will display a df Total which equals N - 1 = 24 - 1 = 23.
- df total:N - 1, where N is the total number of observations.

Since the results of our one way ANOVA indicate that there is not significant difference between the sample means, there is one more tool we can use to determine if the different diets may have had some impact on chicken growth. In research today, it is important to report **practical significance** in addition to statistical significance. To report practical significance, you will calculate an **effect size** known as **eta-squared**. Mathematically, it is calculated as:

$$\eta^2 = \frac{SS\ between}{SS\ total}$$

where:

- *SSbetween* is the sum of squares between groups (the variance due to the independent variable),
- *SStotal* is the total sum of squares (the total variance in the data).

The value of η^2 ranges from 0 to 1:

- A value of 0 means the independent variable does not explain any of the variance in the dependent variable.
- A value closer to 1 indicates that a larger proportion of the variance is explained by the independent variable.

The eta-squared effect size calculations helps researchers understand the practical significance of the ANOVA results beyond just statistical significance. Typically, eta-squared is interpreted as follows:

$\eta^2 \sim 0.01$: small effect

$\eta^2 \sim 0.06$: medium effect

$\eta^2 \sim 0.14$: large effect

Here is the eta-squared calculation for our example:

$$\eta^2 = \frac{SS\ between}{SS\ total} = \frac{142.458333}{581.958333} = 0.2447912938$$

This can be reported as $\eta^2 = 0.24$, which can be interpreted as a large effect. This result can also be interpreted as diet explained 24% of the variability in chicken weight gain. However, the interpretation of the results of an ANOVA need to take into account the context of the study and representative nature of the samples collected.

Two-Way ANOVA

An experimenter may sometimes be interested in comparing two types of treatments and testing whether there are differences in the means for each treatment. A two-way ANOVA assesses the effects of two independent variables (factors) on a dependent variable, while also examining the interaction between these factors. For example, it can analyze the impact of teaching method and study time on test scores, considering both individual effects and how the factors interact. This analysis provides insights into the main effects of each factor as well as their combined interaction effects, offering a more comprehensive understanding of the relationships.

Example:

	Method 1	Method 2	Method 3
Work Pattern 1	58	56	63
Work Pattern 2	49	54	52
Work Pattern 3	60	71	39
Work Pattern 4	76	58	49

This data represents three different teaching methods for a math subject. For each teaching method, there are 4 different work patterns. The scores in the table represent student scores on an assessment. For example, the student who was taught method 1 using work pattern 1 earned 58 on the assessment. The student who was taught with teaching method 3 work pattern 4 earned a 49. We want to know if there is a difference between teaching methods, between the different work patterns, or between combinations of the two.

Since you are testing the impact of two treatments, this will require the application of a two-way ANOVA. There are 'k' levels of Treatment 1, in this case they are the three methods (columns). There are 'r' levels of Treatment 2, in this case they are the four work methods (rows). The degrees of freedom in a two-way ANOVA are found as follows:

degrees of freedom	
df among rows	r - 1
df among columns	k - 1
within	(r -1)(k-1)
total	rk - 1

Here's the two-way ANOVA output:

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows	145.583333	3	48.5277777	0.41437381	0.74906803	4.75706266
Columns	242.666666	2	121.333333	1.03605313	0.41067015	5.14325285
Error	702.666666	6	117.1111111			
Total	1090.91666	11				

In this table, "rows" refers to the different *work patterns* while "columns" refers to the different *teaching methods*. You can see that for "rows" the p-value is 0.7490 which is larger than .05, so you will fail to reject the null hypothesis that the means across all work patterns is the same. Similarly, for "columns" the p-value is 0.4106 which is larger than .05 meaning that the null hypothesis that the means across all teaching methods is the same can not be rejected.

It is possible to calculate an eta-squared effect size for each "treatment" or "factor" in a two-way ANOVA table. Here's how that would work out:

Effect Size of Teaching Methods: $\frac{SS\ columns}{SS\ total} = \frac{242.66666}{1090.9166} = 0.2224428$

Effect Size for Work Patterns: $\frac{SS\ rows}{SS\ total} = \frac{145.58333}{1090.9166} = 0.13345046$

You can also calculate an eta-squared for the combined effect of teaching method and work pattern as follows:

$$\frac{SS\ rows + SS\ columns}{SS\ total} = \frac{145.58333 + 242.66666}{1090.9166} = 0.3558933$$

Here's how you can interpret these effect sizes:

For the effect of teaching method on results, an η^2 = 0.22 was calculated which can be interpreted as a large effect. Similarly, a large effect size η^2 = 0.35 was noted for impact of both teaching method and work pattern. Interestingly, an η^2 = 0.13 was determined for work pattern which is interpreted as a medium effect. A comparison of these results may indicate that work pattern only explained 13% of the variability in scores as compared to the 22% due to teaching method. It is interesting to note that 35% of the variability can be attributed to the combination of teaching method and work pattern.