

11.2.1: Summary of ANOVA Summary Table

Reminder

So far, we've discussed the:

- Variability between groups, showing how effective the IV was; this is shown in the SS_B
- Variability within groups, which masks how effective the IV was; this is shown in the SS_W , and is sometimes called "Error"
- Degrees of Freedom for Between Groups
- Degrees of Freedom for Within Groups
- The Mean Square for Between Groups, which is like an estimated average of the variability between the groups
- The Mean Square for Within Groups (Error), which is like an estimated average of the variability within groups
- The calculated F-score, which is a ratio of the variability between the groups and the variability within the group
- ANOVA Summary Table, which shows all of these!

ANOVA Summary Table

All of our sources of variability fit together in meaningful, interpretable ways as we saw, and the easiest way to do this is to organize them in the ANOVA Summary Table (Table 11.2.1.1), which shows the formulas for everything other than the sum of squares, to calculate our test statistic.

Table 11.2.1.1: ANOVA Table

Source	SS	df	MS	F
Between Groups	SS_B	$k - 1$	$\frac{SS_B}{df_B}$	$\frac{MS_B}{MS_W}$
Within Groups (Error)	SS_W	$N - k$	$\frac{SS_W}{df_W}$	N/A
Total	SS_T	$N - 1$	N/A	N/A

The first column of the ANOVA table, labeled "Source", indicates which of our sources of variability we are using: between groups, within groups, or total. The second column, labeled "SS", contains our values for the sums of squares that we learned to calculate above. As noted previously, calculating these by hand takes too long, and so the formulas are not presented in Table 11.2.1.1. However, remember that the Total is the sum of the other two, in case you are only given two SS values and need to calculate the third.

The next column, labeled " df ", is our degrees of freedom. As with the sums of squares, there is a different df for each group, and the formulas are presented in the table. Notice that the total degrees of freedom, $N - 1$, is the same as it was for our regular variance. This matches the SS_T formulation to again indicate that we are simply taking our familiar variance term and breaking it up into difference sources. Also remember that the capital N in the df calculations usually refers to the overall sample size, not a specific group sample size. Notice that the total row for degrees of freedom, just like for sums of squares, is just the Between and Within rows added together. If you take $N - k + k - 1$, then the " $-k$ " and " $+k$ " portions will cancel out, and you are left with $N - 1$. This is another convenient way to quickly check your calculations.

The third column, labeled " MS ", is our Mean Squares for each source of variance. A "mean square" is just another way to say variability. Each mean square is calculated by dividing the sum of squares by its corresponding degrees of freedom. Notice that we do this for the Between row and the Within row, but not for the Total row. There are two reasons for this. First, our Total Mean Square would just be the variance in the full dataset (put together the formulas to see this for yourself), so it would not be new information. Second, the Mean Square values for Between and Within would not add up to equal the Mean Square Total because they are divided by different denominators. This is in contrast to the first two columns, where the Total row was both the conceptual total (i.e. the overall variance and degrees of freedom) and the literal total of the other two rows.

The final column in the ANOVA table, labeled " F ", is our test statistic for ANOVA. The F statistic, just like a t - or z -statistic, is compared to a critical value to see whether we can reject for fail to reject a null hypothesis. Thus, although the calculations look different for ANOVA, we are still doing the same thing that we did in all of Unit 2. We are simply using a new type of data to test our hypotheses. Let's look at hypotheses when we have three or more levels in our IV next.

You might notice a few cells in Table 11.2.1.1 that say "N/A" for "Not Applicable". In reality, these cells should be kept blank. However, to the table as accessible as possible, Dr. MO include "N/A" to indicate that you don't have to do anything for those cells. In your own ANOVA Summary Tables, you can leave them blank.

This page titled [11.2.1: Summary of ANOVA Summary Table](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Michelle Oja](#).

- [11.3: ANOVA Table](#) by [Foster et al.](#) is licensed [CC BY-NC-SA 4.0](#). Original source: <https://irl.umsl.edu/oer/4>.