

2.3: Computing the F statistics

Computing the F -value

Fisher's ANOVA is very elegant in my opinion. It starts us off with a big problem we always have with data. We have a lot of numbers, and there is a lot of variation in the numbers, what to do? Wouldn't it be nice to split up the variation into kinds, or sources. If we could know what parts of the variation were being caused by our experimental manipulation (i.e., the independent variable we choose as researchers), and what parts were being caused by sampling error, we would be making really good progress. We would be able to know if our experimental manipulation was causing more change in the data than sampling error, or chance alone. If we could measure those two parts of the total variation, we could make a ratio, and then we would have an F value. This is what the ANOVA does. It splits the total variation in the data into two parts. The formula is:

Total Variation = Variation due to Manipulation + Variation due to sampling error

This is a nice idea, but it is also vague. We haven't specified our measure of variation. What should we use?

Remember the sums of squares that we used to make the variance and the standard deviation? That's what we'll use. Let's take another look at the formula, using sums of squares for the measure of variation:

$$SS_{\text{total}} = SS_{\text{Effect}} + SS_{\text{Error}}$$

SS Total

The total sums of squares, or SS_{Total} is a way of thinking about all of the variation in a set of data. It's pretty straightforward to measure. No tricky business. All we do is find the difference between each score and the grand mean, then we square the differences and add them all up.

Let's imagine we had some data in three groups, A, B, and C. For example, we might have 3 scores in each group. The data could look like this:

groups	scores	diff	diff_squared
A	20	13	169
A	11	4	16
A	2	-5	25
B	6	-1	1
B	2	-5	25
B	7	0	0
C	2	-5	25
C	11	4	16
C	2	-5	25
Sums	63	0	302
Means	7	0	33.5555555555556

The data is organized in long format, so that each row is a single score. There are three scores for the A, B, and C groups. The mean of all of the scores is called the **Grand Mean**. It's calculated in the table, the Grand Mean = 7.

We also calculated all of the difference scores **from the Grand Mean**. The difference scores are in the column titled `diff`. Next, we squared the difference scores, and those are in the next column called `diff_squared`.

Remember, the difference scores are a way of measuring variation. They represent how far each number is from the Grand Mean. If the Grand Mean represents our best guess at summarizing the data, the difference scores represent the error between the guess and each actual data point. The only problem with the difference scores is that they sum to zero (because the mean is the balancing

point in the data). So, it is convenient to square the difference scores, which gets rid of the negative signs (or values) and turns all of them into positive numbers. The size of the squared difference scores still represents error between the mean and each score. And, the squaring operation exacerbates the differences as the error grows larger (squaring a big number makes a really big number, squaring a small number still makes a smallish number).

OK fine! We have the squared deviations from the grand mean, we know that they represent the error between the grand mean and each score. What next? SUM THEM UP!

When you add up all of the individual squared deviations (difference scores) you get the sums of squares. That's why it's called the sums of squares (SS).

Now, we have the first part of our answer:

$$SS_{\text{total}} = SS_{\text{Effect}} + SS_{\text{Error}}$$

$$SS_{\text{total}} = 302$$

and

$$302 = SS_{\text{Effect}} + SS_{\text{Error}}$$

What next? If you think back to what you learned about algebra, and solving for X , you might notice that we don't really need to find the answers to both missing parts of the equation. We only need one, and we can solve for the other. For example, if we found SS_{Effect} , then we could solve for SS_{Error} .

SS Effect

SS_{Total} gave us a number representing all of the change in our data, how all the scores are different from the grand mean.

What we want to do next is estimate how much of the total change in the data might be due to the experimental manipulation. For example, if we ran an experiment that causes change in the measurement, then the means for each group will be different from other. As a result, the manipulation forces change onto the numbers, and this will naturally mean that some part of the total variation in the numbers is caused by the manipulation.

The way to isolate the variation due to the manipulation (also called effect) is to look at the means in each group, and calculate the difference scores between each group mean and the grand mean, and then sum the squared deviations to find SS_{Effect} .

Consider this table, showing the calculations for SS_{Effect} .

groups	scores	means	diff	diff_squared
A	20	11	4	16
A	11	11	4	16
A	2	11	4	16
B	6	5	-2	4
B	2	5	-2	4
B	7	5	-2	4
C	2	5	-2	4
C	11	5	-2	4
C	2	5	-2	4
Sums	63	63	0	72
Means	7	7	0	8

Notice we created a new column called `means`. For example, the mean for group A was 11. You can see there are three 11s, one for each observation in row A. The means for group B and C happen to both be 5. So, the rest of the numbers in the means column

are 5s.

What we are doing here is thinking of each score in the data from the viewpoint of the group means. The group means are our best attempt to summarize the data in those groups. From the point of view of the mean, all of the numbers are treated as the same. The mean doesn't know how far off it is from each score, it just knows that all of the scores are centered on the mean.

Now that we have converted each score to it's mean value we can find the differences between each mean score and the grand mean, then square them, then sum them up. We did that, and found that the $SS_{\text{Effect}} = 72$.

SS_{Effect} represents the amount of variation that is caused by differences between the means. I also refer to this as the amount of variation that the researcher can explain (by the means, which represent differences between groups or conditions that were manipulated by the researcher).

Notice also that $SS_{\text{Effect}} = 72$, and that 72 is smaller than $SS_{\text{total}} = 302$. That is very important. SS_{Effect} by definition can never be larger than SS_{total} .

SS Error

Great, we made it to SS Error. We already found SS Total, and SS Effect, so now we can solve for SS Error just like this:

$$SS_{\text{total}} = SS_{\text{Effect}} + SS_{\text{Error}}$$

switching around:

$$SS_{\text{Error}} = SS_{\text{total}} - SS_{\text{Effect}}$$

$$SS_{\text{Error}} = 302 - 72 = 230$$

We could stop here and show you the rest of the ANOVA, we're almost there. But, the next step might not make sense unless we show you how to calculate SS_{Error} directly from the data, rather than just solving for it. We should do this just to double-check our work anyway.

groups	scores	means	diff	diff_squared
A	20	11	-9	81
A	11	11	0	0
A	2	11	9	81
B	6	5	-1	1
B	2	5	3	9
B	7	5	-2	4
C	2	5	3	9
C	11	5	-6	36
C	2	5	3	9
Sums	63	63	0	230
Means	7	7	0	25.555555555556

Alright, we did almost the same thing as we did to find SS_{Effect} . Can you spot the difference? This time for each score we first found the group mean, then we found the error in the group mean estimate for each score. In other words, the values in the *diff* column are the differences between each score and it's group mean. The values in the *diff_squared* column are the squared deviations. When we sum up the squared deviations, we get another Sums of Squares, this time it's the SS_{Error} . This is an appropriate name, because these deviations are the ones that the group means can't explain!

Degrees of freedom

Degrees of freedom come into play again with ANOVA. This time, their purpose is a little bit more clear. Dfs can be fairly simple when we are doing a relatively simple ANOVA like this one, but they can become complicated when designs get more complicated.

Let's talk about the degrees of freedom for the SS_{Effect} and SS_{Error} .

The formula for the degrees of freedom for SS_{Effect} is

$df_{\text{Effect}} = \text{Groups} - 1$, where Groups is the number of groups in the design.

In our example, there are 3 groups, so the df is $3 - 1 = 2$. You can think of the df for the effect this way. When we estimate the grand mean (the overall mean), we are taking away a degree of freedom for the group means. Two of the group means can be anything they want (they have complete freedom), but in order for all three to be consistent with the Grand Mean, the last group mean has to be fixed.

The formula for the degrees of freedom for SS_{Error} is

$df_{\text{Error}} = \text{scores} - \text{groups}$, or the number of scores minus the number of groups. We have 9 scores and 3 groups, so our df for the error term is $9 - 3 = 6$. Remember, when we computed the difference score between each score and its group mean, we had to compute three means (one for each group) to do that. So, that reduces the degrees of freedom by 3. 6 of the difference scores could be anything they want, but the last 3 have to be fixed to match the means from the groups.

Mean Squares

OK, so we have the degrees of freedom. What's next? There are two steps left. First we divide the SS es by their respective degrees of freedom to create something new called Mean Squared deviations or Mean Squares. Let's talk about why we do this.

First of all, remember we are trying to accomplish this goal:

$$F = \frac{\text{measure of effect}}{\text{measure of error}}$$

We want to build a ratio that divides a measure of an effect by a measure of error. Perhaps you noticed that we already have a measure of an effect and error! How about the SS_{Effect} and SS_{Error} . They both represent the variation due to the effect, and the leftover variation that is unexplained. Why don't we just do this?

$$\frac{SS_{\text{Effect}}}{SS_{\text{Error}}}$$

Well, of course you could do that. What would happen is you can get some really big and small numbers for your inferential statistic. And, the kind of number you would get wouldn't be readily interpretable like a t value or a z score.

The solution is to **normalize** the SS terms. Don't worry, normalize is just a fancy word for taking the average, or finding the mean. Remember, the SS terms are all sums. And, each sum represents a different number of underlying properties.

For example, the SS effect represents the sum of variation for three means in our study. We might ask the question, well, what is the average amount of variation for each mean...You might think to divide SS effect by 3, because there are three means, but because we are estimating this property, we divide by the degrees of freedom instead ($\# \text{ groups} - 1 = 3 - 1 = 2$). Now we have created something new, it's called the MS_{Effect} .

$$MS_{\text{Effect}} = \frac{SS_{\text{Effect}}}{df_{\text{Effect}}}$$

$$MS_{\text{Effect}} = \frac{72}{2} = 36$$

This might look alien and seem a bit complicated. But, it's just another mean. It's the mean of the sums of squares for the effect. It shows the change in the data due to changes in the means (which are tied to the experimental conditions).

The SS_{Error} represents the sum of variation for nine scores in our study. That's a lot more scores, so the SS_{Error} is often way bigger than than SS_{Effect} . If we left our SS error this way and divided them, we would almost always get numbers less than one, because the SS_{Error} is so big. What we need to do is bring it down to the average size. So, we might want to divide our SS_{Error} by

9, after all there were nine scores. However, because we are estimating this property, we divide by the degrees of freedom instead (scores-groups) = $9 - 3 = 6$). Now we have created something new, it's called the MS_{Error} .

$$MS_{\text{Error}} = \frac{SS_{\text{Error}}}{df_{\text{Error}}}$$
$$MS_{\text{Error}} = \frac{230}{6} = 38.33$$

Calculate F

Now that we have done all of the hard work, calculating F is easy:

$$F = \frac{\text{measure of effect}}{\text{measure of error}}$$
$$F = \frac{MS_{\text{Effect}}}{MS_{\text{Error}}}$$
$$F = \frac{36}{38.33} = .939$$

Once we have the F statistics, we can find the corresponding significance or p value (the statistics program you use for calculation will likely present this value in the output automatically), and compare it to the pre-determined p critical value to make a decision about the null hypothesis.

This page titled [2.3: Computing the F statistics](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Yang Lydia Yang](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.

- **7.2: One-factor ANOVA** by [Matthew J. C. Crump](#) is licensed [CC BY-SA 4.0](#). Original source: <https://www.crumplab.com/statistics/>.