

Comparing more than Two Means: Analysis of Variance



Using Analysis of Variance to Compare More than Two Means.

Researchers are frequently required to test whether there is any significant difference in the performance of a number of different types of “treatments”. For example, to compare the effects of: medicines, fertilizers or lubricants on “units” such as patients, land plots or machines. An exhaustive (pairing all combinations) series of two sample tests is possible, but very time consuming. A more efficient method is to begin by testing if there *is* any significant difference in *any* of them to start with. Once that is established, a further process is used to distinguish which of the treatments is significantly different.

For example:

Does where students sit in lectures make any difference to their exam results?

Students were asked where they typically sat – front, middle or back, and their respective final grades (GPA’s) were noted.

*In this example, the mean GPA’s of each group of students could have been tested for a difference that was significant. This would require three Hypothesis tests for comparison between students who sit in the (1) front and middle, (2) front and back and (3) middle and back. That is, three separate tests. Each test requires calculation of *sed*’s, *t* tests and comparison with *t* tables. This process becomes unwieldy when comparing larger numbers of treatments. (4 treatments would require 12 comparisons, 5 “treatments” would require 60.)*

The following are the first steps in the ANOVA process.

State the null and alternative hypotheses:

The null hypothesis for this case would be that there would be no difference in the mean grades for students in all three locations. That is, they would be equal:

$$H_0 : \mu_{front} = \mu_{middle} = \mu_{back}$$

And the alternative hypothesis is that the mean grades for students in the different locations are not equal.

Draw up and complete an ANOVA (Analysis of Variance) table:

Source of variation	Degrees of freedom (df)	Sum of Squares (SS)	Mean Square (MS)	F	P
Treatment					
Error					
Total					

This table is almost identical to the ANOVA table for Simple Linear Regression.

The difference is that source of variation due to Regression is replaced by the term “treatment”.

“Treatment” is the general term for the variable being compared. In the worked example above, “treatment” = location

A summary of the analysis of data from 384 students is given:

Source of variation	Degrees of freedom (df)	Sum of Squares (SS)	Mean Square (MS)	F	P
Location	2	3.994			
Error	381	113.775			
Total	383	117.769			

Before proceeding, assumptions about using an ANOVA F test are made:

1. Samples must be independent and random
2. Populations are normally distributed.
3. Populations have equal variances.

Degrees of freedom

For "Location" this is 2 (one less than the number of locations)

For Total, this is 383 (one less than the number of students)

And from this, df for error is calculated: $383 - 2 = 381$

Mean Square (MS)

To complete this column, divide the Sum of Squares (SS) by the respective df:

$$\frac{3.994}{2} = 1.997 \text{ and } \frac{113.775}{381} = 0.299$$

Enter these in the appropriate position:

Source of variation	df	SS	MS	F	P
Location	2	3.994	1.997		
Error	381	113.775	0.299		
Total	383	117.769			

F statistic

Divide MS Location by MS Error: $\frac{1.996}{0.229} = 6.69$ and enter this.

Source of variation	df	SS	MS	F	P
Location	2	3.994	1.997	6.69	
Error	381	113.775	0.299		
Total	383	117.769			

Note that these formulae are NOT on your formula sheet, .

(MS_{Location} represents the mean variation *AMONG* the locations and MS_{Error} represents the meant variation *WITHIN* each of the locations – ie the variability within each location sample. This is equivalent to the F calculation for a Simple Liner Regression:

$$F = \frac{MS_{\text{regression}}}{MS_{\text{error}}} = \frac{MSR}{MSE}$$

Using the F table to test the hypothesis:

The table is completed by deciding on the p value. You, however, will most likely be given the value of p as a basis on which to make your decision.

You will see that the F table has degrees of freedom along the top (numerator) and down the left hand side (denominator). This corresponds to the degrees of freedom in the two values used to calculate F for the table. That is, the treatment degrees of freedom are the numerator df and the error degrees of freedom are the denominator df.

There are various formats for the F table. The ones you are likely to be given will be for testing the hypothesis at the 5% and 1% level. That is $\alpha = 0.05$ or 0.01 . A section of the $\alpha = 0.01$ F table is below.

There are further tables for $\alpha = 0.025$ and for $\alpha = 0.05$. Obviously for research, the smaller the value of α (or p), the more significant the result. Recall that this is the probability of a Type I error – the probability that this is a chance result.

Note that after 29 $df_{denominator}$, values go from 30, 40 60 120 and ∞ (infinity). If the denominator df are in between any of the values given, use the df below. For example $df = 36$, use $df = \underline{30}$. For values over 120, use ∞ .

Denominator df_2	Numerator, df_1										
	1	2	3	4	5	6	7	8	9	10	12
1	4,052.00	4,999.50	5,403.00	5,625.00	5,764.00	5,859.00	5,928.00	5,982.00	6,022.00	6,056.00	6,106.00
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18

In this example, the numerator (top line = MS_{location}) has 2 df, and the denominator (bottom line = MS_{error}) has 381 df.

Hence df for this example are 2 and ∞ respectively and the F value from the table is 4.61.

The decision as to whether to reject the null hypothesis or not hinges on a comparison of the test and table values. **If the test value is greater than the table value, reject H_0 .**

In the example, test value is 6.69 and table value is 4.61.

That is, there is enough evidence at the 1% level to reject H_0 in favour of the alternative, that there is a difference in the GPA's of students, dependent on where they sit in a lecture room.

$p = 0.01$ can now be entered in the table.

Source	Df	SS	MS	F	P
Location	2	3.994	1.997	6.69	0.01
Error	381	113.775	0.299		
Total	383	117.769			

Practice

(q.1 and q.2) Given following examples of partially completed ANOVA tables:

- calculate the relevant values and complete the table
- test at the $\alpha = 0.01$ level of significance and provide a simple conclusion that could be understood by somebody with no background in statistics.

- ANOVA table for an experiment to compare three different programs for losing weight:

Analysis of Variation for Weight Loss Program

Source	df	SS	MS	F	P
Program		1140.0			
Error	7				
Total		1480.00			

- Five different models of cars were tested for top speeds reached. In each case, a sample size of 6 was used.

Analysis of Variation for top speed

Source	df	SS	MS	F	P
Model			364.1		
Error		362			
Total					

- Cindy Ho, Finance Manager at Red Hill Radiator Supplies (RHRS), theorises that the discount level offered to credit customers affects how long they take to pay. Accordingly, she has designed an experiment to test her theory using 4 sales discount rates (0%, 2%, 4% and 6%). She randomly assigned 5 customers to each sales discount rate and recorded the time each customer took to pay their bill. An analysis of Cindy's data produced the following ANOVA table

Source	df	SS	MS	F	P
Between Groups		25.35			
Error (within)					
Total		182.55			

- a) Complete the blank cells in the above ANOVA table. (4 Marks)
- b) Which of the following statements gives a valid interpretation of what a large F value in the ANOVA table above would mean? (2 Marks)
- Payment times vary much more than discount rates
 - The mean payment time is almost the same for all discount rates
 - There is so much variation in the payment time between customers at the same discount rate that you cannot tell if discounts make a difference
 - The random differences between how long customers take to pay, is small compared to the differences associated with different discount rates.
- c) Does the data provide sufficient evidence to indicate that customers can be influenced to pay sooner by giving them discounts? Test at the $\alpha = 0.01$ level of significance and provide a simple conclusion that could be understood by somebody with no background in statistics. (Show ALL working). (4 Marks)

Answers:

1.

Source	df	SS	MS	F	P
Program	2	1140.0	570.0	11.74	0.01
Error	7	340.0	48.57		
Total		1480.00			

F table value = 9.55. Reject H_0 in favour of H_A that there is a significant difference in the weight loss programmes.

2.

Source	df	SS	MS	F	P
Model	4	1456.6	364.1	25.1	0.01
Error	25	362.00	14.48		
Total	29	1818.5			

F table value = 4.185. Reject H_0 in favour of H_A that there is a significant difference in the top speeds of the different cars.

3. a) number of "groups" = 4 and sample size = 20 (4 x 5)

Source	df	SS	MS	F	P
Between Groups	3	25.35	8.45	0.8600	>0.05
Error (within)	16	157.2	9.825		
Total	19	182.55			

- b) A
- c) From the table with (3,16) df at $\alpha = 0.01$, $F = 5.29$ or 3.24 at $\alpha = 0.05$. Not enough evidence to reject the null hypothesis.