Stats: Two-Way ANOVA

The two-way analysis of variance is an extension to the one-way analysis of variance. There are two independent variables (hence the name two-way).

Assumptions

- The populations from which the samples were obtained must be normally or approximately normally distributed.
- The samples must be independent.
- The variances of the populations must be equal.
- The groups must have the same sample size.

Hypotheses

There are three sets of hypothesis with the two-way ANOVA.

The null hypotheses for each of the sets are given below.

1. The population means of the first factor are equal. This is like the one-way ANOVA for the row factor.
2. The population means of the second factor are equal. This is like the one-way ANOVA for the column factor.
3. There is no interaction between the two factors. This is similar to performing a test for independence with contingency tables.

Factors

The two independent variables in a two-way ANOVA are called factors. The idea is that there are two variables, factors, which affect the dependent variable. Each factor will have two or more levels within it, and the degrees of freedom for each factor is one less than the number of levels.

Treatment Groups

Treatement Groups are formed by making all possible combinations of the two factors. For example, if the first factor has 3 levels and the second factor has 2 levels, then there will be 3x2=6 different treatment groups.

As an example, let's assume we're planting corn. The type of seed and type of fertilizer are the two factors we're considering in this example. This example has 15 treatment groups. There are 3-1=2 degrees of freedom for the type of seed, and 5-1=4 degrees of freedom for the type of fertilizer. There are 2*4 = 8 degrees of freedom for the interaction between the type of seed and type of fertilizer.

The data that actually appears in the table are samples. In this case, 2 samples from each treatment group were taken.

	Fert I	Fert II	Fert III	Fert IV	Fert V
Seed A-402	106, 110	95, 100	94, 107	103, 104	100, 102
Seed B-894	110, 112	98, 99	100, 101	108, 112	105, 107

Seed C-952	94, 97	86, 87	98, 99	99, 101	94, 98

Main Effect

The main effect involves the independent variables one at a time. The interaction is ignored for this part. Just the rows or just the columns are used, not mixed. This is the part which is similar to the one-way analysis of variance. Each of the variances calculated to analyze the main effects are like the between variances

Interaction Effect

The interaction effect is the effect that one factor has on the other factor. The degrees of freedom here is the product of the two degrees of freedom for each factor.

Within Variation

The Within variation is the sum of squares within each treatment group. You have one less than the sample size (remember all treatment groups must have the same sample size for a two-way ANOVA) for each treatment group. The total number of treatment groups is the product of the number of levels for each factor. The within variance is the within variation divided by its degrees of freedom.

The within group is also called the error.

F-Tests

There is an F-test for each of the hypotheses, and the F-test is the mean square for each main effect and the interaction effect divided by the within variance. The numerator degrees of freedom come from each effect, and the denominator degrees of freedom is the degrees of freedom for the within variance in each case.

Two-Way ANOVA Table

It is assumed that main effect A has a levels (and A = a-1 df), main effect B has b levels (and B = b-1 df), n is the sample size of each treatment, and N = abn is the total sample size. Notice the overall degrees of freedom is once again one less than the total sample size.

Source	SS	df	MS	F
Main Effect A	*given*	A, a-1	SS / df	MS(A) / MS(W)
Main Effect B	*given*	B, b-1	SS / df	MS(B) / MS(W)
Interaction Effect	*given*	A*B, (a-1)(b-1)	SS / df	MS(A*B) / MS(W)
Within	*given*	N - ab, ab(n-1)	SS / df	
Total	sum of others	N - 1, abn - 1		

Summary

The following results are calculated using the Quattro Pro spreadsheet. It provides the p-value and the critical values are for alpha = 0.05.

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F-crit*
Seed	512.8667	2	256.4333	28.283	0.000008	3.682
Fertilizer	449.4667	4	112.3667	12.393	0.000119	3.056
Interaction	143.1333	8	17.8917	1.973	0.122090	2.641
Within	136.0000	15	9.0667			
Total	**1241.4667**	**29**				

From the above results, we can see that the main effects are both significant, but the interaction between them isn't. That is, the types of seed aren't all equal, and the types of fertilizer aren't all equal, but the type of seed doesn't interact with the type of fertilizer.

Error in Bluman Textbook

The two-way ANOVA, Example 13-9, in the Bluman text has the incorrect values in it. The student would have no way of knowing this because the book doesn't explain how to calculate the values.

Here is the correct table:

Source of Variation	*SS*	*df*	*MS*	*F*
Sample	3.920	1	3.920	4.752
Column	9.680	1	9.680	11.733
Interaction	54.080	1	54.080	65.552
Within	3.300	4	0.825	
Total	70.980	7		

The student will be responsible for finishing the table, not for coming up with the sum of squares which go into the table in the first place.

Table of Contents