# 6.1: Main Effects and Interaction Effect

In the previous chapter we used one-way ANOVA to analyze data from three or more populations using the null hypothesis that all means were the same (no treatment effect). For example, a biologist wants to compare mean growth for three different levels of fertilizer. A one-way ANOVA tests to see if at least one of the treatment means is significantly different from the others. If the null hypothesis is rejected, a multiple comparison method, such as Tukey's, can be used to identify which means are different, and the confidence interval can be used to estimate the difference between the different means.

Suppose the biologist wants to ask this same question but with two different species of plants while still testing the three different levels of fertilizer. The biologist needs to investigate not only the average growth between the two species (main effect A) and the average growth for the three levels of fertilizer (main effect B), but also the **interaction** or relationship between the two factors of species and fertilizer. Two-way analysis of variance allows the biologist to answer the question about growth affected by species and levels of fertilizer, and to account for the variation due to both factors simultaneously.

Our examination of one-way ANOVA was done in the context of a completely randomized design where the treatments are assigned randomly to each subject (or experimental unit). We now consider analysis in which two factors can explain variability in the response variable. Remember that we can deal with factors by controlling them, by fixing them at specific levels, and randomly applying the treatments so the effect of uncontrolled variables on the response variable is minimized. With two factors, we need a factorial experiment.

|  | | Factor B (fertilizer) | | |
| --- | --- | --- | --- | --- |
|  |  | Level 1 | Level 2 | Level 3 |
| Factor A (species) | Species 1 | 1.2, 2.4, 2.6, 2.2 | 2.4, 2.7, 2.7, 2.9 | 3.1, 3.0, 3.2, 3.4 |
|  | Species 2 | 0.6, 0.9, 1.0, 0.9 | 2.1, 2.3, 2.0, 1.9 | 0.7, 0.5, 0.6, 0.5 |

*Table 1. Observed data for two species at three levels of fertilizer.*

This is an example of a factorial experiment in which there are a total of 2 x 3 = 6 possible combinations of the levels for the two different factors (species and level of fertilizer). These six combinations are referred to as treatments and the experiment is called a **2 x 3 factorial experiment**. We use this type of experiment to investigate the effect of multiple factors on a response and the interaction between the factors. Each of the *n* observations of the response variable for the different levels of the factors exists within a cell. In this example, there are six cells and each cell corresponds to a specific treatment.

When you compare treatment means for a factorial experiment (or for any other experiment), multiple observations are required for each treatment. These are called replicates. For example, if you have four observations for each of the six treatments, you have four **replications** of the experiment. Replication demonstrates the results to be reproducible and provides the means to estimate experimental error variance. Replication also provides the capacity to increase the precision for estimates of treatment means. Increasing replication decreases $s_{\bar{y}}^2 = \frac{s^2}{r}$ thereby increasing the precision of $\bar{y}$.

Notation

- $k$ = number of levels of factor A
- $l$ = number of levels of factor B
- $kl$ = number of treatments (each one a combination of a factor A level and a factor B level)
- $m$ = number of observations on each treatment

## Main Effects and Interaction Effect

Main effects deal with each factor separately. In the previous example we have two factors, A and B. The main effect of Factor A (species) is the difference between the mean growth for Species 1 and Species 2, averaged across the three levels of fertilizer. The main effect of Factor B (fertilizer) is the difference in mean growth for levels 1, 2, and 3 averaged across the two species. The interaction is the simultaneous changes in the levels of both factors. If the changes in the level of Factor A result in different changes in the value of the response variable for the different levels of Factor B, we say that there is an interaction effect between the factors. Consider the following example to help clarify this idea of interaction.

> Example 6.1.1:
>
> Factor A has two levels and Factor B has two levels. In the left box, when Factor A is at level 1, Factor B changes by 3 units. When Factor A is at level 2, Factor B again changes by 3 units. Similarly, when Factor B is at level 1, Factor A changes by 2 units. When Factor B is at level 2, Factor A again changes by 2 units. There is no interaction. The change in the true average response when the level of either factor changes from 1 to 2 is the same for each level of the other factor. In this case, changes in levels of the two factors affect the true average response separately, or in an additive manner.
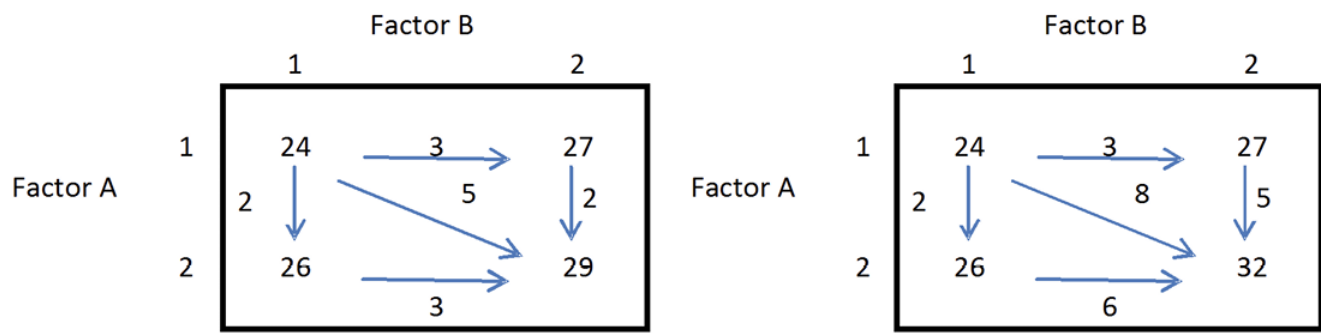
*Figure 1. Illustration of interaction effect.*

### Solution

The right box illustrates the idea of interaction. When Factor A is at level 1, Factor B changes by 3 units but when Factor A is at level 2, Factor B changes by 6 units. When Factor B is at level 1, Factor A changes by 2 units but when Factor B is at level 2, Factor A changes by 5 units. The change in the true average response when the levels of both factors change simultaneously from level 1 to level 2 is 8 units, which is much larger than the separate changes suggest. In this case, there is an interaction between the two factors, so the effect of simultaneous changes cannot be determined from the individual effects of the separate changes. Change in the true average response when the level of one factor changes depends on the level of the other factor. You cannot determine the separate effect of Factor A or Factor B on the response because of the interaction.

## Assumptions

> Note: Basic Assumption
>
> The observations on any particular treatment are independently selected from a normal distribution with variance σ2 (the same variance for each treatment), and samples from different treatments are independent of one another.

We can use normal probability plots to satisfy the assumption of normality for each treatment. The requirement for equal variances is more difficult to confirm, but we can generally check by making sure that the largest sample standard deviation is no more than twice the smallest sample standard deviation.

Although not a requirement for two-way ANOVA, having an equal number of observations in each treatment, referred to as a balance design, increases the power of the test. However, unequal replications (an unbalanced design), are very common. Some statistical software packages (such as Excel) will only work with balanced designs. Minitab will provide the correct analysis for both balanced and unbalanced designs in the General Linear Model component under ANOVA statistical analysis. However, for the sake of simplicity, we will focus on balanced designs in this chapter.

## Sums of Squares and the ANOVA Table

In the previous chapter, the idea of sums of squares was introduced to partition the variation due to treatment and random variation. The relationship is as follows:

$$SSTo = SSTr + SSE \tag{6.1.1}$$

We now partition the variation even more to reflect the main effects (Factor A and Factor B) and the interaction term:

$$SSTo = SSA + SSB + SSAB + SSE \tag{6.1.2}$$

where

1. SSTo is the total sums of squares, with the associated degrees of freedom $klm – 1$
2. SSA is the factor A main effect sums of squares, with associated degrees of freedom $k – 1$
3. SSB is the factor B main effect sums of squares, with associated degrees of freedom $l – 1$
4. SSAB is the interaction sum of squares, with associated degrees of freedom $(k – 1)(l – 1)$
5. SSE is the error sum of squares, with associated degrees of freedom $kl(m – 1)$

As we saw in the previous chapter, the magnitude of the SSE is related entirely to the amount of underlying variability in the distributions being sampled. It has nothing to do with values of the various true average responses. SSAB reflects in part underlying variability, but its value is also affected by whether or not there is an interaction between the factors; the greater the interaction, the greater the value of SSAB.

The following ANOVA table illustrates the relationship between the sums of squares for each component and the resulting F-statistic for testing the three null and alternative hypotheses for a two-way ANOVA.

1. $H_0$: There is no interaction between factors
   $H_1$: There is a significant interaction between factors
2. $H_0$: There is no effect of Factor A on the response variable
   $H_1$: There is an effect of Factor A on the response variable
3. $H_0$: There is no effect of Factor B on the response variable
   $H_1$: There is an effect of Factor B on the response variable

If there is a significant interaction, then ignore the following two sets of hypotheses for the main effects. A significant interaction tells you that the change in the true average response for a level of Factor A depends on the level of Factor B. The effect of simultaneous changes cannot be determined by examining the main effects separately. If there is NOT a significant interaction, then proceed to test the main effects. The Factor A sums of squares will reflect random variation and any differences between the true average responses for different levels of Factor A. Similarly, Factor B sums of squares will reflect random variation and the true average responses for the different levels of Factor B.

| Source of variation | df | Sums of squares | Mean square | F |
|---|---|---|---|---|
| Factor A | $k - 1$ | SSA | $MSA = \dfrac{SSA}{k-1}$ | $F_A = \dfrac{MSA}{MSE}$ |
| Factor B | $l - 1$ | SSB | $MSB = \dfrac{SSB}{l-1}$ | $F_B = \dfrac{MSB}{MSE}$ |
| Interaction AB | $(k-1)(l-1)$ | SSAB | $MSAB = \dfrac{SSAB}{(k-1)(l-1)}$ | $F_{AB} = \dfrac{MSAB}{MSE}$ |
| Error | $kl(m-1)$ | SSE | $MSE = \dfrac{SSE}{kl(m-1)}$ | |
| Total | $klm - 1$ | SSTo | | |

Table 2. Two-way ANOVA table.

Each of the five sources of variation, when divided by the appropriate degrees of freedom (df), provides an estimate of the variation in the experiment. The estimates are called **mean squares** and are displayed along with their respective sums of squares and df in the analysis of variance table. In one-way ANOVA, the mean square error (MSE) is the best estimate of $\sigma^2$ (the population variance) and is the denominator in the F-statistic. In a two-way ANOVA, it is still the best estimate of $\sigma^2$. Notice that in each case, the MSE is the denominator in the test statistic and the numerator is the mean sum of squares for each main factor and interaction term. The F-statistic is found in the final column of this table and is used to answer the three alternative hypotheses. Typically, the p-values associated with each F-statistic are also presented in an ANOVA table. You will use the **Decision Rule** to determine the outcome for each of the three pairs of hypotheses.

**If the p-value is smaller than α (level of significance), you will reject the null hypothesis.**

When we conduct a two-way ANOVA, we always first test the hypothesis regarding the interaction effect. If the null hypothesis of no interaction is rejected, we do NOT interpret the results of the hypotheses involving the main effects. If the interaction term is NOT significant, then we examine the two main effects separately. Let's look at an example.

Example 6.1.2:

An experiment was carried out to assess the effects of soy plant variety (factor A, with $k = 3$ levels) and planting density (factor B, with $l = 4$ levels – 5, 10, 15, and 20 thousand plants per hectare) on yield. Each of the 12 treatments ($k * l$) was randomly applied to $m = 3$ plots ($klm = 36$ total observations). Use a two-way ANOVA to assess the effects at a 5% level of significance.

| Variety (A) | 5 (k/ha) | Density (B) 10 (k/ha) | 15 (k/ha) | 20 (k/ha) |
|---|---|---|---|---|
| 1 | 7.8, 9.1, 10.6 | 11.2, 12.7, 13.3 | 12.1, 12.5, 14.1 | 9.1, 10.7, 12.6 |
| 2 | 8.0, 8.7, 10.0 | 11.3, 12.9, 13.8 | 13.8, 14.3, 15.4 | 11.3, 12.7, 14.3 |
| 3 | 15.3, 16.0, 17.6 | 16.8, 18.3, 19.2, | 17.9, 21.0, 20.7 | 17.2, 18.3, 19.1 |

Table 3. Observed data for three varieties of soy plants at four densities.

It is always important to look at the sample average yields for each treatment, each level of factor A, and each level of factor B.

Table 4. Summary table.

| Density |
|---|
| |

| | Density | | | | Sample average yield for each level of factor A |
|---|---|---|---|---|---|
| Variety | 5 | 10 | 15 | 20 | Sample average yield for each level of factor A |
| Variety | 5 | 10 | 15 | 20 | Sample average yield for each level of factor A |
| 1 | 9.17 | 12.40 | 12.90 | 10.80 | 11.32 |
| 2 | 8.90 | 12.67 | 14.50 | 12.77 | 12.21 |
| 3 | 16.30 | 18.10 | 19.87 | 18.20 | 18.12 |
| Sample average yield for each level of factor B | 11.46 | 14.39 | 15.77 | 13.92 | 13.88 |

For example, 11.32 is the average yield for variety #1 over all levels of planting densities. The value 11.46 is the average yield for plots planted with 5,000 plants across all varieties. The grand mean is 13.88. The ANOVA table is presented next.

Table 5. Two-way ANOVA table.

| Source | DF | SS | MSS | F | P |
|---|---|---|---|---|---|
| variety | 2 | 327.774 | 163.887 | 100.48 | <0.001 |
| density | 3 | 86.908 | 28.969 | 17.76 | <0.001 |
| variety*density | 6 | 8.068 | 1.345 | 0.82 | 0.562 |
| error | 24 | 39.147 | 1.631 | | |
| total | 35 | | | | |

You begin with the following null and alternative hypotheses:

- $H_0$ : There is no interaction between factors
- $H_1$ : There is a significant interaction between factors

The F-statistic:

$$F_{AB} = \frac{MSAB}{MSE} = \frac{1.345}{1.631} = 0.82 \qquad (6.1.3)$$

The p-value for the test for a significant interaction between factors is 0.562. This p-value is greater than 5% (α), therefore we fail to reject the null hypothesis. There is no evidence of a significant interaction between variety and density. So it is appropriate to carry out further tests concerning the presence of the main effects.

$H_0$ : There is no effect of Factor A (variety) on the response variable

$H_1$ : There is an effect of Factor A on the response variable

The F-statistic:

$$F_A = \frac{MSA}{MSE} = \frac{163.887}{1.631} = 100.48 \qquad (6.1.4)$$

The p-value (<0.001) is less than 0.05 so we will reject the null hypothesis. There is a significant difference in yield between the three varieties.

- $H_0$ : There is no effect of Factor B (density) on the response variable
- $H_1$ : There is an effect of Factor B on the response variable

The F-statistic:

$$F_A = \frac{MSB}{MSE} = \frac{28.969}{1.631} = 17.76 \qquad (6.1.5)$$

The p-value (<0.001) is less than 0.05 so we will reject the null hypothesis. There is a significant difference in yield between the four planting densities.