FES 524: Natural Resources Data Analysis

Reading 4.1: Factorial designs

Contents

1 Multiple factors of interest

Up to this point, we have been studying relatively simple study designs with a single factor of interest, but many studies involve interest in the effect of more than one factor on the mean of the response variable. There are several reasons we might want to study multiple factors at once. First, it is often practically efficient in terms of costs to investigate multiple factors at once rather than doing separate studies to investigate separate factors. In addition, the analysis can often be done as a single analysis instead of separate analyses, which is also efficient. Finally, investigators may expect the factors to act together, in which case they can explore the interaction between the two factors. You will learn more about interactions later in this reading.

1.1 Crossed factors

Factors are crossed if all combinations of factor levels are present. This means every level of one factor is present with every level of another factor. To check for crossing we check that all combinations are present in our data. If they are, we will describe the two factors as crossed or fully crossed.

Table 1 shows an example of crossed factors, A and B. The columns show the three levels of factor A (for example three levels of herbicide concentrations) while the rows show the two levels of a second factor, B (for example, ungulate exclosure or not). The cells of the table show the combined factor levels.

Table 2 another example. This example involves two different fertilizers (A and B) and two different fertilizer concentrations (0.5 kg/ha and 1.0 kg/ha). All combinations of the two factors are present.

Table 1: Example of a factorial design crossing the three levels of factor *A* with the two levels of factor *B*.

	A1	A2	A3
B1	A1-B1	A2-B1	A3-B1
B2	A1-B2	A2-B2	A3-B2

Table 2: Example of a factorial design crossing the two fertilizers with the two concentrations of interest.

	Fertilizer *A*	Fertilizer *B*
0.5 kg/ha	0.5 kg/ha A	0.5 kg/ha B
1.0 kg/ha	1.0 kg/ha A	1.0 kg/ha B

1.2 Simple effects

When there are multiple factors of interest and we are interested in the mean response for each combination of the levels of each, one analysis option is to continue to use the separate means model approach we used in weeks 2 and 3. This involves making a new, *combined* factor variable. The levels of this new factor are made by combining the levels of the original factors. Using Table 1 as an example, the new factor would have six levels, one for each cell of the table.

Doing an analysis like this can be referred to as a *simple effects* analysis. There is nothing wrong with such an analysis, and in some cases this may actually be the preferred analysis. The usefulness of this approach can depend on the research question. Setting up the comparisons of interest can be more complicated in a simple effects analysis unless there is interest in all pairwise comparisons.

The simple effects analysis approach is especially useful when factors are only partially crossed rather than fully crossed. The most common example of partial crossing that I see is when there is a control group for one of the factors that involves doing nothing. The example in Table 2 could be a situation where this could come up. For example, perhaps there is a third level of the fertilizer factor that is a "do nothing" treatment where neither fertilizer is applied. Obviously we can't have 0.5 kg/ha of no fertilizer, so the last level of no fertilizer would not be crossed with the concentration factor.

2 Factorial designs

It is important to know about the simple effects analysis as an option. However, the new concept we are learning about this week is factorial designs and analyzing multiple factors simultaneously rather than combining multiple factors into a single factor.

A fully factorial design is one with fully crossed factors. Such designs are often described by the number of levels of the factors. For example, the fertilizer × concentration study example without the control would be called a 2 by 2 factorial, since the fertilizer factor has 2 levels and the concentration factor has 2 levels. You may see this written as 2×2. The example in Table 1 from this week's motivating example would be called a 2x3 factorial.

2.1 Interaction effects

Over the next few weeks we are going to be focusing a lot on *interactions*, which is one of the important concepts that arises when talking about factorial designs. An interaction is when the effect of one variable depends on the value of another variable. Interactions are often specifically of interest when working with data collected from a factorial design.

Let's use the fertilizer and concentration example to explore what an interaction might look like in this case.

Figure 1 shows what is called an interaction plot for a hypothetical analysis using data collected for the fertilizer-concentration example. An interaction plot shows the estimated means for every factor combination.

Such plots are most often used as exploratory plots and so should not be used to make statistical inference. In this case I have included error bars from a fitted model to show there is little variation around the estimated means. In lab we will add the raw data to our interaction plots to show variation and to make sure we recognize the plot as an exploratory plot.

```
## Warning in knitr::include_graphics(here("images/interaction_plot.png")): It is
## highly recommended to use relative paths for images. You had absolute paths: "/
## Users/dusty/Documents/Teaching/FES524/readings/images/interaction_plot.png"
```
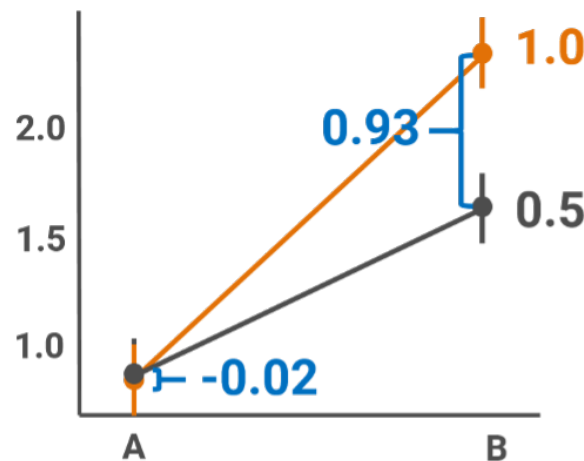

Figure 1: Example interaction plot. The orange colors represent the 1.0 kg/ha concentration treatment, while grey shows the 0.5 kg/ha concentration. The points represent the estimated mean response for each combination of the two factors of interest.

The example in Figure 1 shows evidence of an interaction. The difference in mean response between the 1.0 kg/ha concentration and the 0.5 kg/ha concentration is small and negative for the A fertilizer but large and positive for the B fertilizer. The effect of one factor variable clearly depends on the value of the other factor variable.

Here are a couple of ways you could describe the interaction:

- The effect of concentration depends on fertilizer type.

- The difference in mean response among concentrations differs between fertilizers.

The above examples are focused on the concentration variable. However, we could make the same statements using the fertilizer factor as the focus. The way you describe an interaction will depend on how you worded your research question and what you want to focus on in the results.

Alternative ways to describe the interaction:

- The effect of fertilizer depends on the concentration.

- The difference in mean response among fertilizers differs between concentrations.

Below is language we can use to describe results from an analysis that focuses on the interaction. We would report estimated differences in means among concentrations (fertilizers) for each fertilizer (concentration). It wouldn't make sense to talk about some overall effect of concentration when we believed that the effect depended on which fertilizer was applied.

Effect of concentration for Fertilizer A:

Increasing the amount of Fertilizer A from 0.5 to 1.0 kg/ha is estimated to decrease mean biomass by 0.02 kg/ha (95% CI: 0.08 decrease to 0.04 increase).

Effect of concentration for Fertilizer B:

Increasing the amount of Fertilizer B from 0.5 to 1.0 kg/ha is estimated to increase the mean biomass 0.93 kg/ha (95% CI: 0.86 to 0.99 increase).

2.2 Main (overall) effects

When two factors are applied at once simply for efficiency of a study design, the investigator may be solely interested in overall effects. In such a case, the factors are not expected to interact.

We will again use the fertilizer and concentration example, this time to explore what it means to be interested in the overall effect of the factors.

Figure 2 shows another interaction plot, however, Figure 2 shows the results of a hypothetical analysis where there was no interaction. You can see in this idealized example that the difference in mean response for the two concentrations is identical for the A fertilizer and the B fertilizer.

```
## Warning in knitr::include_graphics(here("images/no_interaction_eg.png")): It is
## highly recommended to use relative paths for images. You had absolute paths: "/
## Users/dusty/Documents/Teaching/FES524/readings/images/no_interaction_eg.png"
```
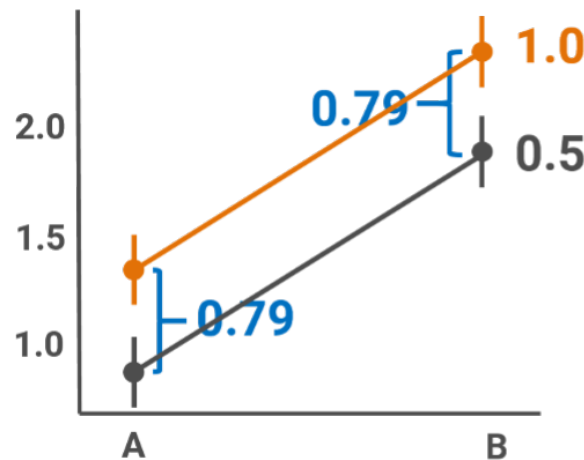

Figure 2: Example interaction plot in which there is no interaction between the fertilizer and concentration factors.

If needed, here is how we could describe the lack of interaction:

- The effect of concentration does not depend on fertilizer type.
- The difference in mean response among concentrations does not differ between fertilizers.

If the research question was clearly about overall effects, only comparisons of main effects (ignoring the other factor) should be reported. Below is some example language about what this would look like.

Overall effect of concentration:

Increasing the amount of fertilizer from 0.5 to 1.0 kg/ha is estimated to increase the mean biomass by 0.79 kg/ha (95% CI: 0.74 to 0.84 increase).

Overall effect of fertilizer:

Fertilizer B is estimated to increase mean biomass by 1.25 kg/ha compared to Fertilizer A (95% CI: 1.20 to 1.30 increase).

3   Interaction or no interaction? That is the question.

Working with multiple explanatory variables should not automatically mean the analysis has to include an interaction, although many classes that discuss factorial designs focus on interactions and main effects from a purely statistical and not scientific perspective. When should the investigator think about and include interactions?

This is actually a question about good research practices as well as appropriate statistical practice. An investigator with a well-fleshed out research question and study design in an established field should have an idea if they expect the factors to interact or not. In confirmatory work, this expectation should be made clear in the introduction section that describes the study. The examples we are using in this class will demonstrate different ways a researcher can do this.

As we have and will continue to emphasize in this class, a p-value from an overall test about an interaction cannot by itself indicate if there "was" or "was not" an interaction. How the analysis proceeds and what results should be reported needs to be driven by scientific expertise and the research question. Regardless of the size of a p-value as evidence against the null hypothesis of no interaction, if the research question was explicitly about how two factors together influence the mean response, then comparisons should be reported using both factors (i.e., at the interaction level) to assess practical importance.

There are of course situations where research is exploratory. In such cases, there are no previous studies and no scientific theories to guide the research. The investigator may then have no idea if two factors should interact or not. Remember that the goal of exploratory work is to gather information that can be used to inform future research. In exploratory work with multiple factors of interest, the investigator will most often plan on reporting all possible comparisons, both main effects and all interaction effects, in order to best inform future research. p-values can certainly be used as evidence against the null hypothesis of no interaction with all the usual caveats, but it is the estimated differences among groups of interest, both simple effects and main effects, that will be most useful in forming future research questions.

4   Main effects in the presence of an interaction

Another common discussion point when working with factorial designs is when and if it makes sense to talk about main effects when the effect of one variable actually depends on the value of another variable. You will see some sweeping rules on this in some textbooks, such as "never report a main effect in the presence of an interaction". Like so many things in statistics, things are not really so cut and dry. Whether or not you want to report a main effect in the presence of an interaction depends on the situation you are in.

To estimate the main effect of one factor in the presence of an interaction with another factor, we average over the levels of the second factor. Sometimes it makes sense to do this, but other times it may not. We will walk through examples of this in class.

Much of this question about "main effects in the presence of an interaction" comes down to the research question and the kind of research being done. If an investigator wants to estimate main effects based on some a priori reason, even though they also believed a factor would interact with another, they are free to

do so. They should be clear why the main effects are of interest. If a study involves research that is totally exploratory, the investigator will most likely want to report all effects to best inform future research.

Note that with categorical explanatory variables, we never have to remove an interaction from the model to talk about main effects. Model selection is not a standard part of statistical analyses and working with factors makes this easy. From the statistical point of view, taking things in and out of the model is not a standard modeling approach. This may come as a surprise to some of you based on the statistical practice you may have seen reported in the literature in your field. While we don't have a lot of time to discuss this, we will touch on the topic of model selection in the coming weeks.

Working with continuous variables is a little bit different. We can still estimate overall slopes even if we allow separate slopes per group from a model (i.e., an interaction between a continuous and a categorical variable). However, we don't have any easy way to calculate standard errors for that overall slope. We are often forced to remove continuous by categorical interactions if the focus is on overall slopes instead of group specific slopes unless we bootstrap confidence intervals.

Note that there is no need to report overall slopes if the original question only involved an interaction. On the other hand, if the research question was specifically about overall slopes and the interaction was a check of model fit, leaving it out of the model in the end can make scientific sense (since the model without it was the one the investigator believed was "true" based on *a priori* knowledge).

5   A note on language when working with multiple factors

Once we are working with multiple factors in the same model we must be more careful with the language we use. It is not unusual to see people refer to each individual factor as a "treatment", but also use the word "treatment" to describe applying the combination of multiple factors. This is extremely confusing for the reader. Choose your language carefully to avoid this problem. For example, we could use "concentration" and "fertilizer type" to refer to the factors in the examples above.

If you must use the word "treatment", one option I have seen that makes sense is to only use it to refer to the combination of factor levels applied to a study/experimental unit. Individual factors should then be referred to without that language, reserving "treatment" to mean the factor combinations.

6   Statistical model

The statistical model for a 2×2 factorial design using the regression parameterization can be written as

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 z_{i1} + \beta_3 x_{i1} z_{i1} + \epsilon_i$$

where the variables x_{i1} and z_{i1} are "dummy" variables (0's and 1's) taking the value 1 if observation i comes from factor level 2 for each factor, respectively. As usual, we assume $\epsilon_1, \epsilon_2, ..., \epsilon_n \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$. Thus, β_0 is the mean response when in the reference level, with is the treatment of the first level of each factor. Using the concentration-fertilizer example above, β_0 is the mean response for study units that get fertilizer A at a concentration of 0.5 kg/ha (level 1 of each factor).

The coefficients β_1 and β_2 are *differences* between the mean of other treatments and that of the reference level. Finally, β_3 represents a difference of differences. Don't worry if this is unclear now, we will discuss this further in class and use `emmeans` to construct estimates of interest in practice.