FES 524: Natural Resources Data Analysis

Reading 4.1

**Contents**

[Other readings to do before class 1](#_Toc76116417)

[Multiple factors of interest 1](#_Toc76116418)

[Crossed factors 1](#_Toc76116419)

[Simple effects 2](#_Toc76116420)

[Factorial design 2](#_Toc76116421)

[Interaction effects 3](#_Toc76116422)

[Overall (main) effects 4](#_Toc76116423)

[Language when working with multiple factors 5](#_Toc76116424)

[Interaction or not? 5](#_Toc76116425)

[Main effects in the presence of an interaction 5](#_Toc76116426)

[Statistical model and analysis 6](#_Toc76116427)

# Other readings to do before class

Read the motivating example in Week 4 Handout 1: “Study on transplanted seedling growth in the Willamette Valley, Oregon”. Come prepared to use this example to talk about the topics covered in reading 4.1 as well as topics we’ve already covered. Make sure you’ve had time to think about the answers to the questions in that handout.

# Multiple factors of interest

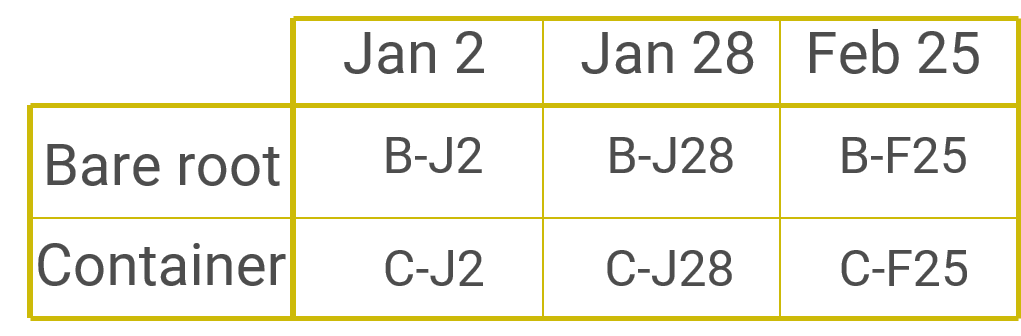
Many studies involve interest in the effect of more than one factor on the mean response variable. There are several reasons we might want to study multiple factors at once. First, it is often practically efficient in terms of costs to investigate multiple factors at once rather than doing separate studies 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.

While you may see research involving many factors of interest, in this class we focus on studies where there are only two factors of interest.

## 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*.

Below is an example of crossed factors from this week’s motivating example. The columns show the three levels of one factor. The rows show the two levels of a second factor. The cells of the table show the combined factor levels.



Here is 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.

1: Fertilizer A at 0.5 kg/ha

2: Fertilizer B at 0.5 kg/ha

3: Fertilizer A at 1.0 kg/ha

4: Fertilizer B at 1.0 kg/ha

## Simple effects

When there are multiple factors of interest, 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. In the fertilizer example above, you can see that this new factor would have four levels.

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 fertilizer-concentration example is a situation where this could come up. A third level of the fertilizer factor might be to apply no fertilizer. But you can’t apply different concentrations of no fertilizer, so the fertilizer levels are not fully crossed with the levels of concentration.

1: Fertilizer A at 0.5 kg/ha

2: Fertilizer B at 0.5 kg/ha

3: Fertilizer A at 1.0 kg/ha

4: Fertilizer B at 1.0 kg/ha

5: No fertilizer

To analyze these data we could use a single factor with five levels as the factor of interest in a simple effects analysis.

# Factorial design

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.

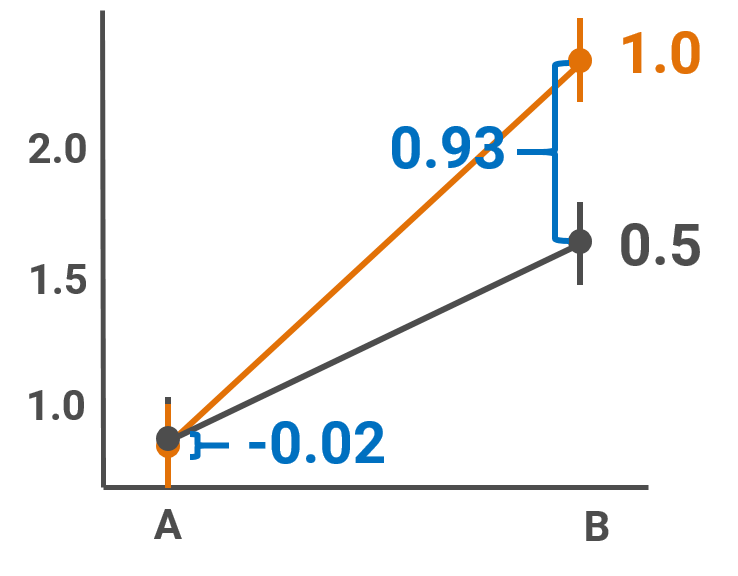
A fully factorial design is one with fully crossed factors. Such designs are often described by the number of levels of the factors. For example, the fertilize-concentration study example without the control would be called a 2 by 2 factorial, since the fertilizer factor has 2 levels and the concentration factor has 2 levels. You’ll see this written as 2x2. The example above from this week’s motivating example would be called a 2x3 factorial.

## Interaction effects

Over the next few weeks we are going to be focusing on *interactions,* which is one of the important concepts that arises when talking about factorials 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.

We’ll use the fertilizer and concentration example to explore what an interaction might look like in this case.

Below is 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.



This interaction plot above 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’s a couple of ways we could describe the interaction:

The effect of concentration depends on fertilizer type.

The difference in mean response among concentrations differs between fertilizers.

I wrote the interaction statements with a focus on the concentration variable. However, I 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 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).

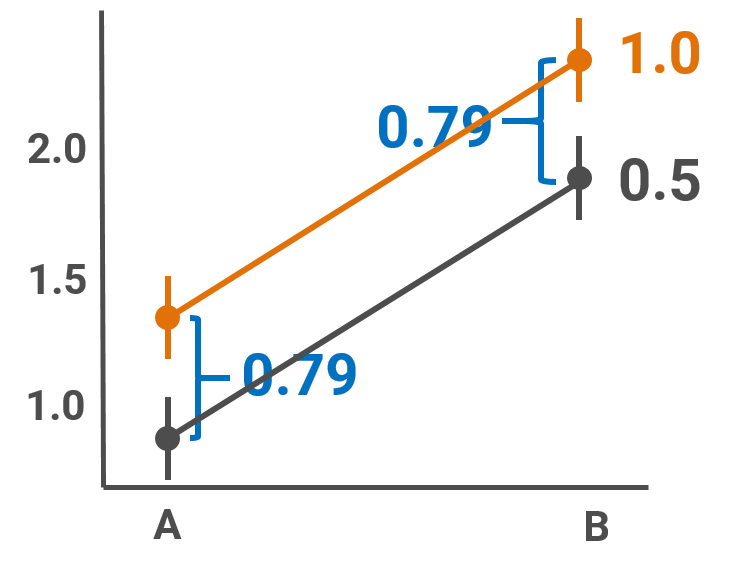
In some fields you may see comparisons made for factor combinations due to interactions called *simple effects* or *cell means*.

## Overall (main) effects

When two factors are applied at once 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’ll again use the fertilizer and concentration example, this time to explore what it means to be interested in the overall effect of the factors.

Below is an interaction plot like the one I made earlier. This time I’m showing 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.



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).

In some fields you may see comparisons made for the overall effects of factors called *main effects* or *marginal means*.

## 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. In the motivating example this week I carefully use “stock type” and “transplanting date” to refer to the factors, not “stock treatment” and “date treatment”.

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

## Interaction or not?

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’ve now discussed multiple times, 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.

## 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 next week.

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.

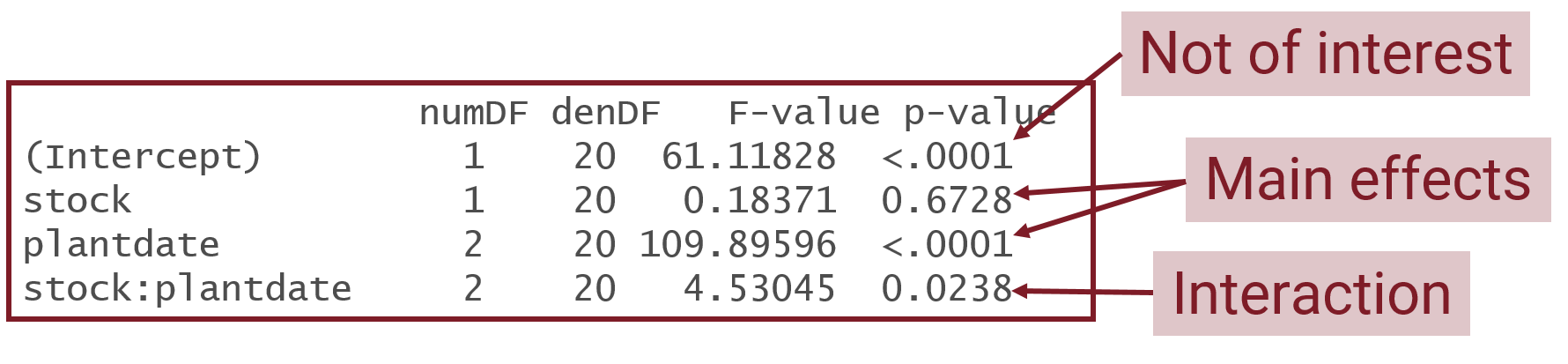
Note that there is no need to report overall slopes if the original question only involved an interaction. On the other hand, if the research question was specifically about overall slopes and the interaction was a check of model fit, leaving it out of the model in the end can make scientific sense (since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

# Statistical model and analysis

The only new part of the statistical model this week is in how the two factors are written. Since we will be using the emmeans package in R to do comparisons it is not as important to understand the summary output the model. We will not cover the mathematical notation in class.

New this week, the analysis in R will include an interaction term in the linear mixed model. As you will see in lab, there will be three possible overall F tests that may be of interest. Ignore the top row that is output when using the anova.lme function.

See the overall F tests below from the example analysis you will do in R in lab. While the anova.lme function gives three tests, you should only plan on reporting those that are actually of interest based on the research question. If the research question is one that involves an interaction, only the F test for the interaction should be reported as evidence against the null hypothesis of no interaction. Tests of main effects should be ignored unless there was specific research questions about main effects.



As long as the data are perfectly balanced (i.e., each factor combination has the same number of observations), the overall F tests for the main effects output by anova.lme for a mixed model are valid. If tests for main effects are of interest those can be reported in your results section. If model assumptions are relaxed or there isn’t perfect balance among groups then the default *sequential* main effects tests from anova.lme in the presence of an interaction are not useful. We will touch on this again and I will provide you a bit more information on what to do in these cases later this term.