Simple Linear Models

Lecture 04.2: Interactions

Lauren Sullivan

Module: Linear, Non-linear, and Mixed Effects Models

Readings

Required for class:

- NA

Optional:

- Crawley, M. *Statistics: An Introduction Using R*
- Bolker, B. *Ecological Models and Data in R - Ebook version*
- Harpole et al. (2011) Nutrient co-limitation of primary producer communities. *Ecology Letters*. 14(9):852-862.

Factorial Experiments

	+ Trt 2	- Trt 2
+ Trt 1	+ Trt 2 + Trt 1	- Trt 2 + Trt 1
- Trt 1	+ Trt 2 - Trt 1	- Trt 2 - Trt 1

Factorial Experiments

	+ Trt 2	- Trt 2
+ Trt 1	+ Trt 2 + Trt 1 both trt factors	- Trt 2 + Trt 1
- Trt 1	+ Trt 2 - Trt 1	- Trt 2 - Trt 1 control

Factorial ANOVA - Experimental Design

Dependent variable (Y) is continuous, independent variables (X) are categorial and interacting

Interactions measure the *joint* effect of main effects A and B

- Identifies if response of A is dependent on level of B

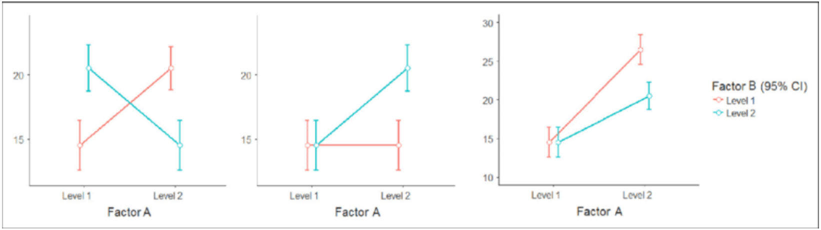
$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$

Significant interactions: main effects not interpertable **without** clarification.

Very common in biology

You can add more factors, but it requires a larger sample size (N)

What do Interactions Look Like?



How do Factors Add?

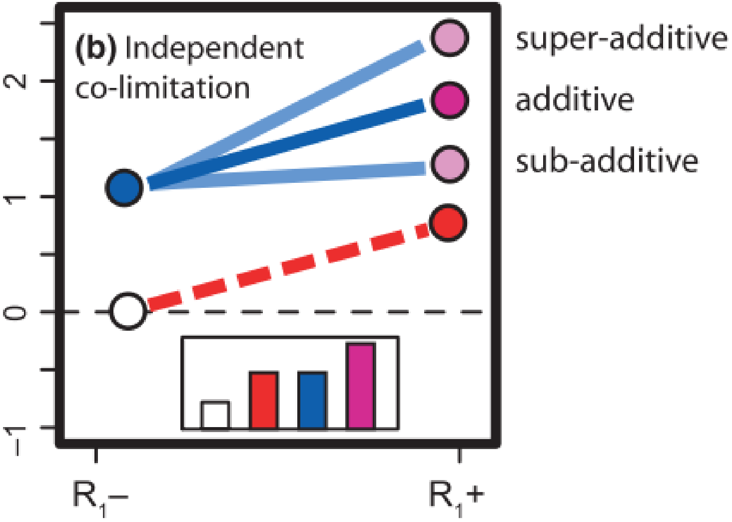


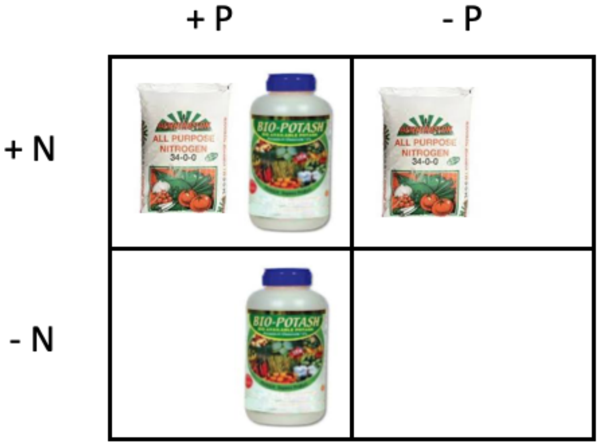
Figure 1: Harpole et al. 2011. Ecology Letters

Data

Biomass differences for plants with either nitrogen (N), phosphorus (P), both (NP), or neither (C) treatments.

Experimental Design

Biomass differences for plants with either nitrogen (N), phosphorus (P), both (NP), or neither (C) treatments.



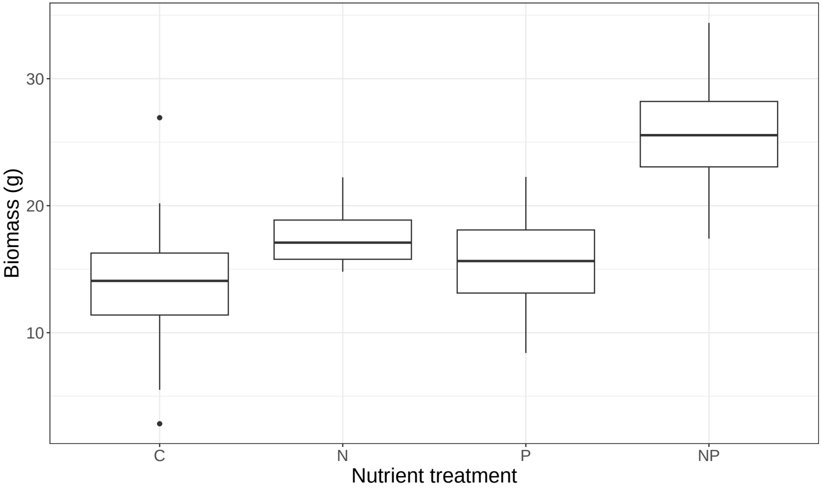
Data

You can see that the columns `n` and `p` mirror the experimental design in the previous slide. 0's indicate that the treatment is not in that plot, and 1's indicate that the treatment is applied. So control plots are `0-0`, nitrogen plots are `1-0`, phosphorus plots are `0-1`, and nitrogen and phophorous plots are `1-1`.

```
## # A tibble: 80 x 5
##         n     p trt   block biomass
##     <dbl> <dbl> <chr> <dbl>   <dbl>
##  1      1     1 NP        1   23.1
##  2      1     0 N         1   15.5
##  3      0     1 P         1   15.0
##  4      0     0 C         1   20.2
##  5      1     1 NP        2   34.4
##  6      1     0 N         2   14.8
##  7      0     1 P         2   18.4
##  8      0     0 C         2    2.83
##  9      1     1 NP        3   25.6
## 10      1     0 N         3   14.9
## # ... with 70 more rows
```

ANOVA

How does the addition of Nitrogen and/or Phosphorous influence the biomass of plants?



ANOVA

Does the addition of Nitrogen, Phosphorous, or Nitrogen and Phosphorous influence the biomass of plants?

```
test_aov <- lm(biomass ~ trt , data = biomass)
summary(test_aov)
```

```
##
## Call:
## lm(formula = biomass ~ trt, data = biomass)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -10.726  -2.241  -0.107   2.482  13.377
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  13.5526     0.9317  14.546  < 2e-16 ***
## trtN          3.9901     1.3177   3.028  0.00336 **
## trtP          1.9610     1.3177   1.488  0.14083
## trtNP        12.2790     1.3177   9.319 3.28e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.167 on 76 degrees of freedom
## Multiple R-squared:   0.57,  Adjusted R-squared:  0.553
## F-statistic: 33.58 on 3 and 76 DF,  p-value: 6.316e-14
```

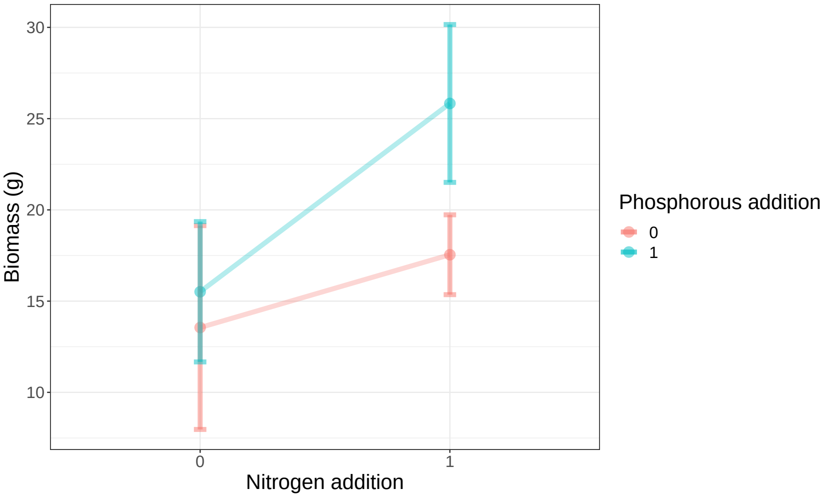
Factorial ANOVA

Does the interaction of Nitrogen and Phosphorous influence the biomass of plants?

```
test_aovf <- lm(biomass ~ n * p , data = biomass)
summary(test_aovf)
```

```
##
## Call:
## lm(formula = biomass ~ n * p, data = biomass)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -10.726  -2.241  -0.107   2.482  13.377
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  13.5526     0.9317  14.546  < 2e-16 ***
## n             3.9901     1.3177   3.028  0.00336 **
## p             1.9610     1.3177   1.488  0.14083
## n:p           6.3280     1.8634   3.396  0.00109 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.167 on 76 degrees of freedom
## Multiple R-squared:   0.57,  Adjusted R-squared:  0.553
## F-statistic: 33.58 on 3 and 76 DF,  p-value: 6.316e-14
```

Interaction Plot



Higher Order Interactions

`lm(biomass ~ n * p , data = biomass)` is the same thing as `lm(biomass ~ n + p + n:p , data = biomass)`

This matters when you start to get higher order interactions (e.g. more than 2 factors. Because you can start to eliminate the highest interactions if they are not significant via reverse elimination.)

Start with: `lm(biomass ~ n * p * k , data = biomass)`

then: `lm(biomass ~ n * p * k - n:p:k , data = biomass)`

then: `lm(biomass ~ n * p * k - n:p:k - n:p , data = biomass)`

etc.