Ethan Rutledge

ECO 602

Week 12 Reading Questions

Q1: In the context of a dataset (real or made up), describe the inherent conflict between using a complicated model that minimizes the unexplained variation and using a simple model that is easy to communicate.

I'm actually going to use the exact same situation I ran into during my masters. We wanted to pair habitat data with population dynamic data of bluegill sunfish in the Upper Mississippi River. We wanted to compare the habitat in the upper reaches of the river to the habitat in the lower reaches of the river to try and explain why there might be differences in population dynamics for bluegill. We were using a long term data set where 30 environmental variables were measured over the past 30 years at 6 sites (3 upper and 3 lower). Our hypothesis was that bluegill are using more backwater habitat that is available to them in the upper where it is not in the lower and thus have bigger, better bluegill populations. We compared flow, depth, substrate and water body classification (backwater vs mainstem) between each site using one way anovas, a lot of them, but very simple modeling.

Reviewer after reviewer turned down the paper because they wanted us to use more complicated modeling that used all 30 environmental variables. I stand by our analyses because our results gave us real world answers that managers can actually use with boots on the ground to improve bluegill habitat (more backwater habitat). The other modeling would've given us very specific conditions where most bluegill are found, but our answer would've been too precise to have any applicable management implications and too confusing to get any public backing of the proposed habitat plans.

Q2: Which of the following predictor variables had slope coefficients that were significantly different from zero at a 95% confidence level? Select the correct answer(s)

A and B

Q3: Using the information in the model coefficient table above, calculate the expected biomass for a plant given:

- **0 mL water per week**
- **0 mg nitrogen per week**
- **0 mg phosphorus per week**

Explain how you made the calculation.

-1.7g… to get the answers for 3 and 4, I took the estimate for each predictor variable and multiplied them by the given values and added them together.

Q4: Using the information in the model coefficient table above, what is the expected biomass for a plant given:

- **10 mL water per week**

- **30 mg nitrogen per week**

- **20 mg phosphorus per week**

3.95g

Q5: Describe the key difference between a simple linear regression and a 1-way analysis of variance.

The key difference is that a simple linear regression is the relationship between 2 continuous variables. A 1-way anova tells us if there is a difference at all between 3 or more variables.

Q6: Identify the *deterministic* component(s) of the model equation.

α and B1

Q7: Identify the *stochastic* component(s) of the model equation.

The error ϵ