Lina Clifford (worked with Olivia Dinkelacker, Liz Clark, Laura Haynes, Jessica Martinez)
ECO 602 – Analysis of Environmental Data
Week 12 Reading Questions
Due 12/3/2022

Q1 (2 pts.): 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. Consider the tradeoff between model complexity and interpretability.

Since your answer is targeted to a non-scientist audience, you should use narrative style using a concrete example.

The benefit of using a complicated model is that it can minimize the amount of unexplained variation or error, but the drawback is that complex models are often not intuitive to understand. Conversely, simple models are more intuitive/easy to interpret, but have more unexplained error or variation. For example, when modeling x tree species distribution, a simple model might only use aspect and temperature to model where the species is likely found. A more complex model might include more predictor variables like slope, precipitation, soil type, and percent tree cover. This second model could likely explain more of the variation in tree species x's distribution, but it isn't as simple to interpret how the predictor variables are effecting the response variable—the tree species x's distribution.

Q2 (1 pt.): 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)

None.

Q3 (2 pts.): Using the information in the model coefficient table above, calculate the expected biomass for a plant given: Explain how you made the calculation.

- **0 mL water per week**

-1.7 + (0.043 * 0) = -1.7g

Intercept plus slope for water times units of water, which ultimately just equals the intercept value.

- **0 mg nitrogen per week**

-1.7 + (0.192 * 0) = -1.7g

Intercept plus slope for nitrogen times units of water, which ultimately just equals the intercept value.

- **0 mg phosphorus per week**

-1.7 + (-0.027 * 0) = -1.7g

Intercept plus slope for phosphorus times units of water, which ultimately just equals the intercept value.

Q4 (2 pts.): Using the information in the model coefficient table above, what is the expected biomass for a plant given: Explain how you made the calculation.

- **10 mL water per week**

-1.7 + (0.043 * 10) = -1.27g

Take the intercept plus the product of the number of units of water and the slope/ “estimate” of water which is 0.043.

- **30 mg nitrogen per week**

-1.7 + (0.192 *30) = 4.06g

Take the intercept plus the product of the number of units of nitrogen and the slope / “estimate” of nitrogen which is 0.192.

- **20 mg phosphorus per week**

-1.7 + (-0.027 * 20) = 2.24g

Take the intercept plus the product of number of units of phosphorus and the slope / “estimate” of phosphorus which is -0.027.

Q5 (1 pt.): Describe the key difference between a simple linear regression and a 1-way analysis of variance. Consider the data types/scales of the predictor and response variables.

A simple linear regression has to have continuous response and predictor variables. 1-way ANOVA requires a categorical predictor variable.

Q6 (1 pt.): Identify the deterministic component(s) of the model equation.

The deterministic components of the model equation are alpha and beta.

Q7 (1 pt.): Identify the stochastic component(s) of the model equation.

The stochastic component of the model equation is epsilon.