

Kaitlyn Valentinetti

November 30, 2022

ECo 602: Michael France Nelson

Reading Questions Week 12

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.

In the context of the Palmer penguin's dataset, a simple linear regression model may only look at two coefficients such as flipper length and body mass of the data. It's easy to see if there is a relationship between these two parameters because there either is or there is not one. However, in a more complicated model, you may look at species (a categorical predictor), flipper length, body mass, and flipper depth. This will give your model more accuracy because there is more data involved to demonstrate the absolute relationship; however, the relationship may be very difficult to interpret. For a theoretical example, we might see that penguin species with the lowest body mass have the longest flippers but the lowest flipper depth. This can be harder to communicate with an audience than a simple model with two parameters.

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)

- Water and nitrogen

Q3 (2 pts.): 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
- $-1.7 + (43 \times 0) + (192 \times 0) + (-27 \times 0) = -1.7$ grams (expected biomass)

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

- 10 mL water per week: $-1.7 + (43 \cdot 10) = 428.3$
- 30 mg nitrogen per week: $-1.7 + (192 \cdot 30) = 5,758.3$
- 20 mg phosphorus per week: $-1.7 + (-27 \cdot 20) = -541.7$
- $428.3 + 5,758.3 - 541.7 = 5644.9\text{mg} \rightarrow 5.64\text{g}$ (expected biomass)

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. The key difference between a simple linear regression and a 1-way ANOVA is the type and number of parameters. In a simple linear regression, there is a single continuous predictor variable and single continuous response variable. In a 1-way ANOVA, there is a categorical predictor that has 3 or more levels and a continuous response variable.

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

β and α are the deterministic components.

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

ϵ is the stochastic component.