Week 10 Reading Questions

Worked with Alex Fink and JT Larkin

1. Why would we want a model selection criterion to penalize the number of parameters in a model?

We want a model selection model selection criterion to penalize the number of parameters in a model so we can simplify the model. The more parameters the more complex the model will become, and this can lead to difficulty reading and understanding the model.

1. Consider the regression equation for a simple linear regression:

yi=α+β1xi+ϵyi=α+β1xi+ϵ

In 2 - 3 paragraphs, describe the meaning of the slope parameter β1 in the context of the relationship between the predictor variable, x, and the response variable y.

Your answer must be in plain non-technical language. Your explanation will be most effective if you use a narrative approach, using a concrete example to illustrate the concept

A simple linear regression helps you to study one predictor variable at a time. You use the equations to determine the relationship between the predictor (x) and response variable (y). These x and y values can be used to create a line by plotting the points as (x,y). To find the slope of the line you can use the β1 variable. This variable is defined as the population slope. The slope will help us to see what the how the responses variable changes as a result of the predictor variable.

We can look at lung capacity and smoking cigarettes as an example. The response variable would be the persons lung capacity and the predictor variable would be the amount of cigarettes a person smokes a day. If we looked at the slope it would tell us how the lung capacity of a person changes in response to the number of cigarettes a person smokes each day. The slope can show that the more cigarettes a person smokes then the lung capacity increases, decreases or stays the same. This information can help us determine the relationship between the two variables.

1. What is the *base case* water treatment?

The base case is the low water treatment.

1. What is the average plant mass, in grams, for the **low** water treatment?

The average plant mass is 2.4. This value is calculated using the estimated value.

1. What is the average plant mass, in grams, for the **medium** water treatment? How did you calculate this quantity?

The average plant mass is 1.3. This value is also calculated using the estimated value.

1. Which of the following questions cannot be addressed with the model coefficient table? Select the correct answer or answers:
2. Is there a positive relationship between increased water availability and plant biomass accumulation?
3. Is water availability a significant predictor for plant biomass accumulation?
4. What is the average biomass of plants in the highwater treatment?

Option letter A cannot be addressed with the table. It shows that there is a relationship between water and plant biomass, but we cannot assume that it is positive.