Stella Dee

Week 10 Reading Questions

I did not work with any other students.

"The first part of the AIC definition is a measure of goodness of fit. The second part is a penalty for the number of parameters in the model." - Zuur (2007)

**Q1:** Why would we want a model selection criterion to penalize the number of parameters in a model?

We would want model selection to penalize the number of parameters because the more parameters there are the more imprecise the predicted intervals will be and the wider the confidence bands (the less confident the predictions will be). Therefore, the more parameters the less useful the model (broadly speaking.)

Consider the regression equation for a simple linear regression:

$$y_i = \alpha + \beta_1 x_i + \epsilon$$

**Q2:** In 2 - 3 short paragraphs, describe the meaning of the slope parameter $\beta_1$

The slope parameter is a quantitative picture of the change in the response variable compared to the change in the predictor variable. For example, if my predictor variable is number of cookies and my response variable is number of chocolate chips (contained in the cookies), if for every 3 cookies (x)I would expect 27 chocolate chips (y) to be on the table (for an average of 9 chips per cookie), the slope of that relationship as expressed above would be .11 repeating.

The slope multiplies the predictor variable to model the relationship between the predictor and the response variable. This may not accurately capture the stochastic model, which is why $\epsilon$ is included in the final model. A linear slope means the relationship between the change in cookies and the change in chocolate chips will remain the same over time—it is multiplicative and linear.

| | Estimate | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | 2.4 | 2.19 | 98.371 | 0.001 |
| waterMed | 1.3 | 5.12 | 0.480 | 0.231 |
| waterHigh | 13.6 | 3.48 | 24.495 | 0.001 |

**Q3:** Based on the model table, what is the *base case* water treatment?

The base case is low water.

**Q4:** What is the average plant mass, in grams, for the **low** water treatment?

2.4 grams. It is the intercept for the base case (low water).

**Q5:** What is the average plant mass, in grams, for the **medium** water treatment?

3.7 grams. It is $2.4 + 1* 1.3 + 0 *13.6$ or 3.7 grams.

**Q6 :** Which of the following questions cannot be addressed with the model coefficient table? Select the correct answer or answers:

A. Is there a positive relationship between increased water availability and plant biomass accumulation?
B. Is water availability a significant predictor for plant biomass accumulation?
C. What is the average biomass of plants in the high water treatment?

It **cannot address** B—overall significance.