Quiz 6

1. (Short answer) We fit a multiple linear regression. We can assess two assumptions using the plot below. What are the assumptions? Evaluate whether or not they are met.
2. (Short answer) Taylor is building a prediction model to predict the number of hours her cat Alice will sleep on a given night. In her model, she includes the following predictors:
   1. Number of minutes Alice played during the day
   2. Number of hours of daylight
   3. Outside temperature
   4. Moon phase (full, gibbous, quarter, crescent, new)

Charlie says that Taylor must also include a variable indicating whether or not Alice received a bedtime snack, since this (a) is associated with the number of minutes Alice played, and (b) may causally affect the number of hours Alice sleeps. He claims that this makes the bedtime snack variable a potential confounder, and that it must be included in Taylor's prediction model.

Is Charlie correct that Taylor should include the bedtime snack variable because it is a potential confounder? Explain your rationale.

1. (Multiple choice) Which of the following characteristics of an estimate is not related to the others?
2. Precision of the estimate
3. Standard error of the estimate
4. Width of the confidence interval around the estimate
5. Accuracy of the estimate
6. (Multiple choice) Which is not a step in the process of fitting and evaluating a linear regression prediction model using training and testing data?
   1. Fitting the model on the test data
   2. Generating model predictions for the test data
   3. Randomly splitting the observations into testing and training
   4. Calculating the MSE based on the test data predictions
7. (Multiple choice) Which of the following best describes the major pitfall of using training data to evaluate the predictive ability of a model?
8. We risk overfitting.
9. We risk violating the classical linear regression assumptions.
10. We risk reducing the accuracy of our predictions.
11. We risk reducing the accuracy of our regression coefficient estimates.
12. Which of the following is a scientific question that we would use prediction to answer?
    1. Who is most at risk of being hospitalized if they get the flu?
    2. What is the mechanism which causes Alzheimer's disease to develop?
    3. Is there an association between ambient noise level and stress?
    4. Does the effect of drug A depend on binarized sex at birth?
13. (True/False) The regression equation for adding potential confounders is the same as the regression equation for adding potential effect modifiers.
14. (True/False) Mean-squared error can take on any value between - and .
15. (True/False) Adjusted R2 will always increase when you add additional covariates to your model.
16. (True/False) The mean-squared error of a prediction model is determined by bias and variance.