BS805 Fall 2022 Week 8

Be sure to follow the *Assessment Guideline 1: Writing up Homework* at the end of the syllabus in preparing the homework for submission.

Homework assignments need to be uploaded to Blackboard by November 9th at 2 PM.

In each homework report, be sure to include an introductory and a summary paragraph.

Use the rehydration data again from assignment 5 to complete this assignment. Recall that the data on this study are available for you in the file *dehydration_f22.xlsx* in the Assignments folder for Class 5. The variables are ID, rehydration score, dose, age, and weight.

1. Run three multiple linear regressions predicting rehydration score.

    a) In model A, predict rehydration score from dose, age, and weight. Also, perform a collinearity diagnostic analysis for this model. Summarize all results.
    b) In model B, predict rehydration score from dose and age and perform collinearity diagnostics. What has changed by removing weight from the model?
    c) In model C, predict rehydration score from dose and weight and perform collinearity diagnostics. What has changed by removing age from the model?
    d) Which analysis do you prefer among models A), B), and C)? Why do you prefer the model that you chose?

2. Construct three new continuous variables called *dose1*, *dose2*, and *dose3* with which you will be able to perform a piecewise linear analysis:

    dose1 – those who received a dose of 0 through < 1;
    dose2 – those who received a dose of 1 through less than 2;
    dose3 – those who received a dose greater than or equal to 2.

    Perform analyses examining the relationship of rehydration score and dose below.

    a) Statistically compare the mean rehydration scores between the dose groups and summarize your results.
    b) Using simple linear regression, perform an analysis of the linear association of rehydration score and dose (continuous). Summarize these results.
    c) Perform a piecewise linear model using *dose1*, *dose2*, and *dose3* to predict rehydration score. Summarize the results. Does the piecewise model generate a higher R^2 than a simple linear regression model? Is the overall model statistically significant at the 0.05 level? What are the estimated individual slopes for *dose1*, *dose2*, and *dose3* and were they statistically significant predictors of rehydration

score? Is the slope for *dose1* equal to the slope for *dose2* and is the slope for *dose2* equal to the slope for *dose3*?

d) Generate a plot to visually compare the simple linear regression to the piecewise linear model.

e) Which analysis among 2. a), b), and c) do you prefer for these data? Why? Present specific numeric evidence to support your choice.