Chapters 11 and 12 Homework – STA5176 – Fall 2019

Exercise 1

Data from problem 11.40 (page 612)

1. Create a scatter plot of the data. type of relationship, if any, appears to be between the dosage of the drug and the response?
2. Find Pearson's correlation coefficient.
3. Comment on the correlation: what type of linear relationship are we observing (negative, none, positive)? Is it a strong correlation?
4. Find Spearman's correlation coefficient.
5. Comment on the correlation: what type of linear relationship are we observing (negative, none, positive)? Is it a strong correlation?
6. How different is Spearman's correlation from Pearson's correlation in this problem (comment on the difference in values)? Does the correlation change drastically?
7. How is Spearman's correlation coefficient calculation different from Pearson's correlation coefficient?
8. Based on your previous answers, do you expect the slope of the regression line to be positive or negative? Why?
9. Construct the regression line. (**Note:** you must report the model as $\hat{y} = b_0 + b_1 x$ for full credit.)
10. Construct the 95% confidence interval for β_1.
11. Formally test to determine if dose level is a significant predictor of the protective strength of the drug at the $\alpha = 0.05$ level.
12. Construct the lack of fit ANOVA table.
13. Formally test to determine if there is lack of fit of the linear regression model at the $\alpha = 0.05$ level

Exercise 2

Data from problem 12.12 (page 688)

1. Create a matrix of scatter plots for revenue, population, and distance.
2. From your scatter plots in part (a), is there a data point that you think may be an "issue" for analysis?
3. Construct and report a multiple regression model relating revenue (y) to distance (x_1) and population (x_2). (**Note:** you must report the model as $\hat{y} = a + b_1 x_1 + b_2 x_2$ for full credit.
4. Construct the 95% confidence interval for β_1.
5. Construct the 95% confidence interval for β_2.
6. Formally test to determine if either x_1 or x_2 significantly predicts revenue at the $\alpha = 0.05$ level. (**Note**: This is requesting *two* hypothesis tests.)
7. Re-run and report your regression model *without* airport 20. (**Note:** you must report the model as $\hat{y} = a + b_1 x_1 + b_2 x_2$.
8. How is the model in part (g) different than the model you reported in part (c)?
9. Excluding airport 20, formally test to determine if either x_1 or x_2 significantly predicts revenue at the $\alpha = 0.05$ level. (**Note**: This is requesting *two* hypothesis tests.)
10. What differences do you observe between the tests in parts (f) and (i)?
11. What do you think I wanted you to take away from this homework problem?