Malavika Elay

Stat 21 – Test 1 – Corrections (3/1)

MCQ

(1) A, C

(5) C, D

Short Answer

(6) 1) I am interested to explore a linear relationship between my payments made on my credit card account (predictor) & my annual percent interest rate (response). I will choose a simple linear regression model to display & analyze my data.

2) I will use the lm() function & ggplot() function on R to create a scatterplot of my data with "Credit Card payments" on the x-axis & "APR" on the y-axis. The lm() function will tell me the estimated slope (β_1) & intercept (β_0) coefficients for my modeled data such that $\hat{y} = \beta_0 + \beta_1 x$.

3) I will use the mutate() function on R to create a scatterplot with my data set's residuals on the y-axis & fitted values on the x-axis. A geometric or mathematical pattern in the residuals plot (like a funnel shape or parabola) could indicate nonlinearity and/or heteroscedasticity in the data. I will also generate a normal quantile plot using R with my sample being the data's residuals. A strong strength of correlation with the quantile line suggests that the data is linear & that it can likely be well-modeled by a SLR.

4) Given that my linear model proved to meet the SLR conditions (linearity, ~~zero mean of error, constant variance,~~ independence) in the previous step, I can safely use my calculated p-value from the model to make an inference about the x-y relationship. If the p-value of my regression is under 0.05 (my alpha level), I can reject the H_0.

(9) a.) We would estimate the SD of the number of calories burned to be **30.84** calories burned per this linear model.

c.) I would use a two-tailed one-sample t-test to figure out if there is a statistically significant difference between our runner's rate of burning calories + the age grap average. $H_0 \rightarrow \beta_1 = 100$, $H_A \rightarrow \beta_1 \neq 100$. If my test p-value is less than my alpha level of 0.05, or my 95% CI for my runner's estimated change in number of calories burned per additional mph doesn't include the number 100, I can reject my null hypothesis.