STAT 315 Chapter 11 Review Questions

1. Two variables X and Y are observed for a sample. Answer the following questions by selecting one of the options to the right:

If X and Y are positively correlated, then the regression slope for predicting Y with X will be positive.	Always True	Always False	Sometimes True Sometimes False
If the assumptions of simple linear regression are met, then b1 follows the t-distribution.	Always True	Always False	Sometimes True Sometimes False
If X and Y have no linear relationship, then they also have no quadratic relationship	Always True	Always False	Sometimes True Sometimes False
If X and Y have a positive linear relationship, then increasing the value of X will cause an increase in the value of Y.	Always True	Always False	Sometimes True Sometimes False
If X and Y are strongly correlated, then the regression slope between X and Y will be large.	Always True	Always False	Sometimes True Sometimes False
MSE is always negative	Always True	Always False	Sometimes True Sometimes False
The regression slope from predicting Y with X is the same as the slope from predicting X with Y.	Always True	Always False	Sometimes True Sometimes False
b0 is the predicted response for a predictor value of 0.	Always True	Always False	Sometimes True Sometimes False
b1 is the predicted change in the response for a one unit increase in the predictor.	Always True	Always False	Sometimes True Sometimes False
If we reject the null hypothesis for the slope, then there is evidence to suggest that the population slope is zero	Always True	Always False	Sometimes True Sometimes False
The uncertainty in predicting the mean response is greater than the uncertainty in predicting a new observation.	Always True	Always False	Sometimes True Sometimes False

Scenario A: Data were collected on the fuel efficiency of cars (mpg), and the total displacement volume of their engines (disp). We'd like to predict efficiency using displacement. Below is some R code and output. View the output and answer the questions that follow.

```
> fit <- lm(mpg ~ disp, mtcars)
> summary(fit)

Call:
lm(formula = mpg ~ disp, data = mtcars)

Residuals:
    Min      1Q  Median      3Q     Max
-4.8922 -2.2022 -0.9631  1.6272  7.2305

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 29.599855   1.229720  24.070  < 2e-16 ***
disp        -0.041215   0.004712  -8.747 9.38e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.251 on 30 degrees of freedom
Multiple R-squared:  0.7183,    Adjusted R-squared:  0.709
F-statistic: 76.51 on 1 and 30 DF,  p-value: 9.38e-10
```

n-2 = 30

2. Write the population regression model, and label each component of the equation

$(mpg)_i = \beta_0 + \beta_1(disp)_i + \varepsilon_i$

3. Write the estimated regression equation and label each component of the equation.

$\widehat{(mpg)}_i = 29.60 - 0.04\,(disp)_i$

4. Is there a linear relationship between engine displacement and fuel efficiency? Explain how you know

yes: slop is significantly different from zero ($p = 9.38 \times 10^{-10}$)

5. What is the result of the hypothesis test $H_0: \beta_1 = 0$ using significance level 0.05?

$p = 9.38 \times 10^{-10} < 0.05$ Reject H_0.

6. Construct a 99% confidence interval for the population slope.

$\beta_1 = b_1 \pm t_{\frac{\alpha}{2}, n-2} \cdot SE$ / $t_{0.005, 30} = 2.75$ / $\beta_1 = -0.041 \pm (2.75)(0.005)$

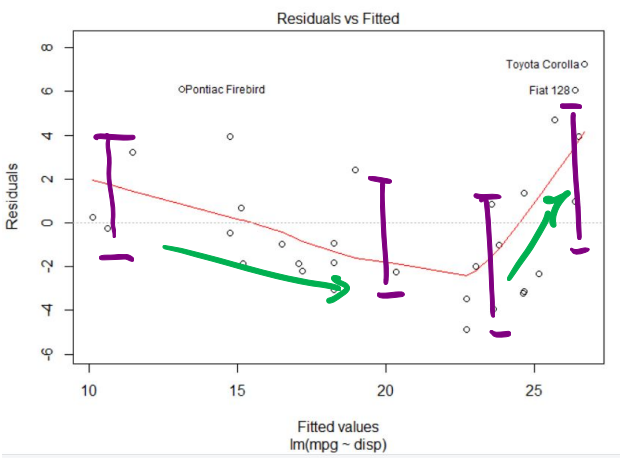
$-0.055 \leq \beta_1 \leq -0.027$

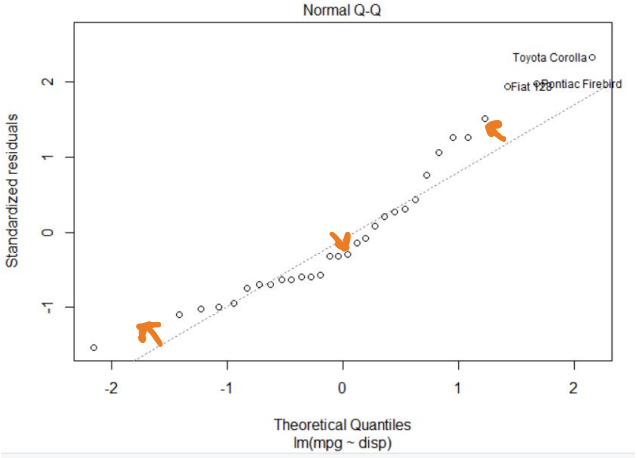
7. How much variability in fuel efficiency is explained by engine displacement?

$R^2 = 0.718 \Rightarrow 71.8\%$

8. Does this model demonstrate that engine displacement is a leading cause of fuel efficiency? Is it reasonable to conclude that engine displacement is a leading cause of fuel efficiency, despite what can be concluded from the model? This model does not establish a causal relationship, but it's reasonable to conclude that a higher displacement volume leads to a lower fuel efficiency.

Scenario A (again): Below are some model diagnostics for the linear model above. View them and answer the questions that follow:





9. Do the diagnostics suggest that the relationship is linear?

no. look at residuals vs. fitted.

10. Do the diagnostics suggest that the residuals are normally distributed?

no. look at Q-Q plot

11. Do the diagnostics suggest that the variance of the residuals are constant?

Possibly: though it's difficult to tell. see Res. vs. fitted again.