

Problem Set 5

QTM 200: Applied Regression Analysis

Due: March 4, 2020

Instructions

- Please show your work! You may lose points by simply writing in the answer. If the problem requires you to execute commands in **R**, please include the code you used to get your answers. Please also include the **.R** file that contains your code. If you are not sure if work needs to be shown for a particular problem, please ask.
- Your homework should be submitted electronically on the course GitHub page in **.pdf** form.
- This problem set is due at the beginning of class on Wednesday, March 4, 2020. No late assignments will be accepted.
- Total available points for this homework is 100.

Using the **teengamb** dataset, fit a model with **gamble** as the response and the other variables as predictors.

```
1 gamble <- (data=teengamb)
2 # run regression on gamble with specified predictors
3 model1 <- lm(gamble ~ sex + status + income + verbal, gamble)
```

Answer the following questions:

- (a) Check the constant variance assumption for the errors by plotting the residuals versus the fitted values.

(b) Check the normality assumption with a Q-Q plot of the studentized residuals.

(c) Check for large leverage points by plotting the h values.

(d) Check for outliers by running an `outlierTest`.

(e) Check for influential points by creating a "Bubble plot" with the hat-values and studentized residuals.