**UNI: jh3561**

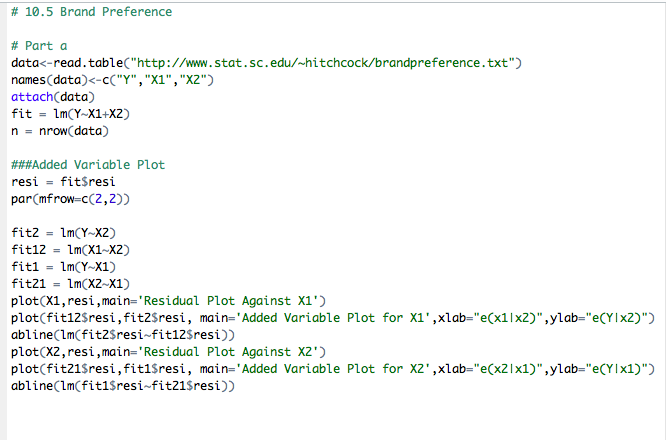
**Name : Jiahong Hu**

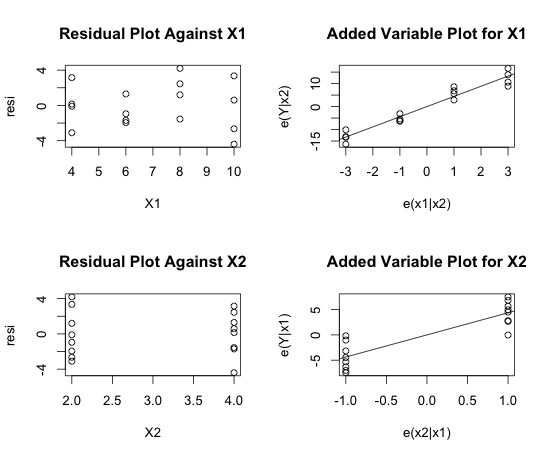
**Stat 4315**

**Spring 2015**

Problem 10.5

Part a.

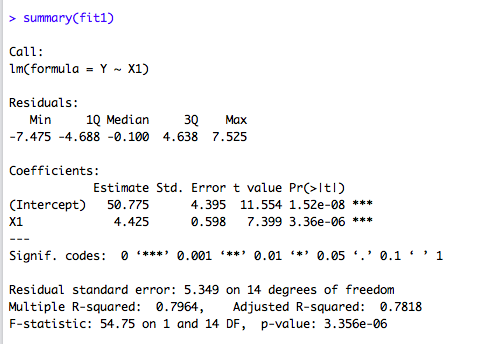


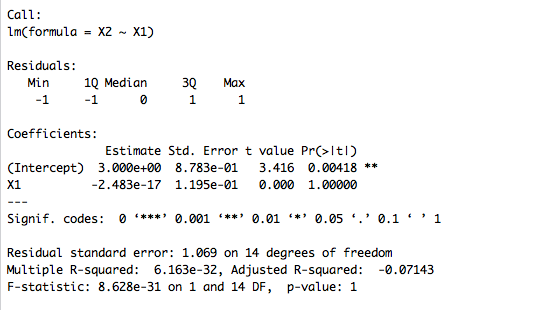


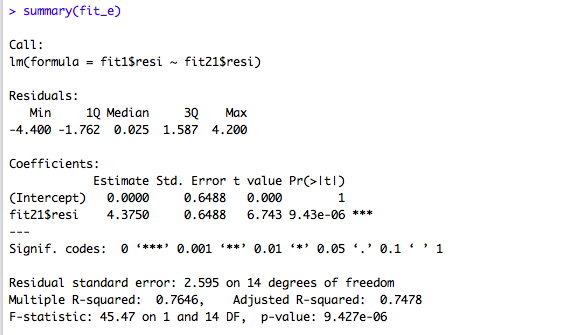
Part b

No. It is appropriate in fact. There is no trend shown in both residual plots for X1 and X2 and the residuals are spread around 0. In addition, both added-variable plots for X1 and X2 present an upward trend, which suggest that the linear term X1 or X2 may be helpful addition to the regression model already containing the other variable.

Part c





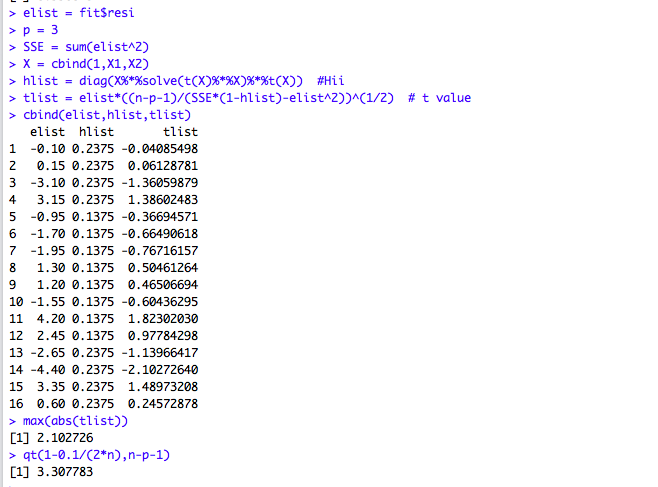




Problem 10.9

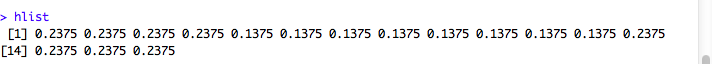
Part a



Bonferroni Outlier Test

I find . Hence, we should conclude there is no outlier. Otherwise, we conclude there are outliers.

Part b

****

If the ith case is outlying in terms of its X observations and therefore has a large value, it exercises substantial leverage in determining the fitted value .

The  in this example satisfies two properties:



hii indicates the distance between the X values for the ith case and the means of the X values for all n cases. It measures the effect to determine the fitted value of a given observation.

In this case, we have two groups of equally influential observations, one is 0.2375 and the other one is 0.1375. The numbers are small in the middle and high in two ends.

To be more specific,

h11=h22=h33=h44=h1313=h1414=h1515=h1616=0.2375,

h55=h66=h77=h88=h99=h1010=h1111=h1212=0.1375

It is because X1, X2, X3,X4, X13, X14, X15, X16 have the same distance from the mean value of X; X5, X6, X7, X8, X9, X10, X11, X12 have the same distance from the mean value of X.

Part c

. Hence, we consider there’re no outlying cases with respect to their X values, according to rule of thumb.

Part d

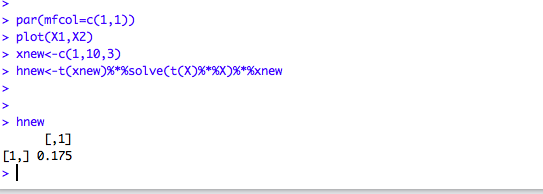
Method 1: The scatter plot indicates there is no extrapolation when x1= 10, x2=3 beyond the range of the data.

****

Method 2 :







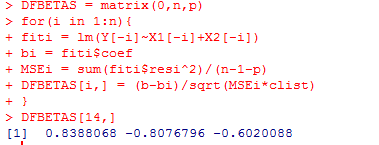
It indicates no extrapolation and hence, the conclusions of the two methods agree with each other.

Part e

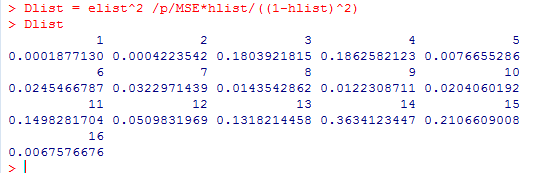
.

DFBETAS

**C:\Users\Senyu\Desktop\1.png**

****

Cook’s distance for case 14 is 0.3634.

****

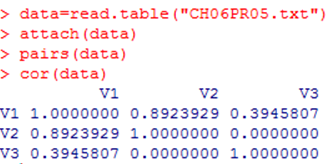
According to DFFITS, the 14th case does have a significant influence in the fitted function.



According to the value of DFFITS, Cook’s Distance and DFBETAS the 14th case does have a significant influence on its fitted value, but dos not have a significant influence on the whole regression.

Problem 10.15

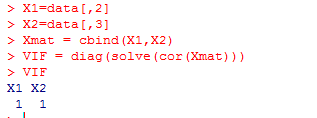
Part a.



****

There doesn’t exist multicolinearity among the predictor variables.

Part b.

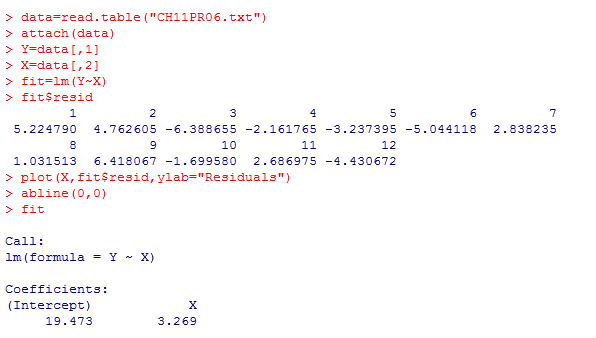
****

*VIF1*=*VIF2*=1 and therefore there’s no correlation between the predictors.

Problem 11.6

Part a

.

****

****

The residuals plot shows that the variance of the error terms increases with the increase of the X values. There is no outlier, but there may have non-constancy of the error variance.

Part b.

If |t\*BF|<=t-critical, then we conclude error variance is constant, otherwise, error variance not constant.

,

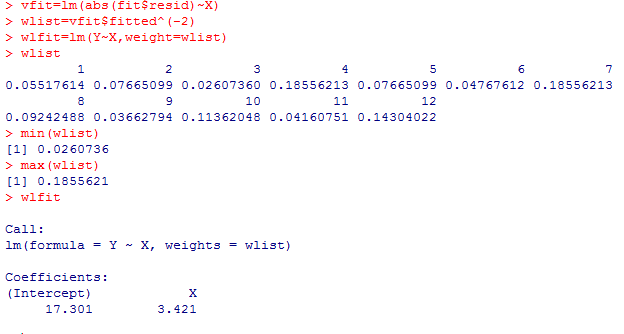
Thus, we can conclude that the error variance is constant.

Part c.

****

This plot shows that the standard deviation of the error terms increases with the increase of the X values.

Part d

****

Case 4 and 7 receives the largest weight 0.1855621 and case 3 receives the smallest weight 0.0260736.

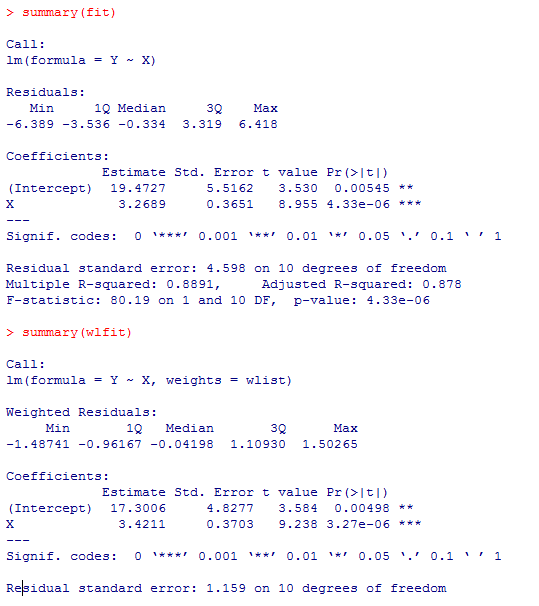
Go back to the ordered linear regression, we have

S^=-0.9049+0.3226\*X.

Part e

The weighted least squares regression function is .

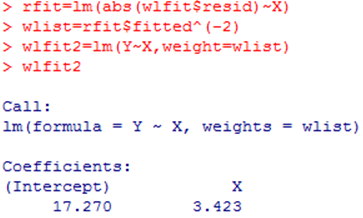
They are similar with some differences from those produced in the ordinary least squares method in (a).

Part f ****

1. OLS method, , 
2. weighted LS method, , 

The standard deviations of the coefficients did not change much after weighted.

Part g



and the result is similar to what we have in the estimated regression coefficients. There are no substantial changes, so we will accept the former regression form.