MATH3424 HW4

Name: Leung Ko Tsun

SID: 20516287

1a

Graphical user interface, text, application

Description automatically generated

Chart, scatter chart

Description automatically generated

The residual plot shows a strong trend of M shape, so the errors should be autocorrelated.

1b

Text, letter

Description automatically generated

Since d = 0.19454 is close to 0 and p-value is 3.503e-14 which is smaller than 0.05, so it shows as an evidence of autocorrelation.

1c

Graphical user interface, text, application, email

Description automatically generated

Since p-value = 6.537e-5 which is smaller than 0.05, we can claim that there exist autocorrelation

1d

Graphical user interface, text, application

Description automatically generated

After doing Cochrane and Orcutt procedure for 1 iteration, the statistic d increases from 0.7338 to 0.80175, but the p-value is 6.909e-5, which is still smaller than 0.05, so that autocorrelation still exists.

2a

A picture containing text

Description automatically generated

All models have VIFj less than 10 which means collinearity has been removed in all models.

3.

a).

Text, letter

Description automatically generated

Chart, line chart

Description automatically generated

All variables becomes more stable at k = 0.65

Text, letter

Description automatically generated

As from the above R result, k1 = 0.2356, k2 = 0.5242, k3 = 0.6304, k4 = 0.6505, so it converges to k = 0.65 after 5 iterations

OLS Result:

Text, letter

Description automatically generated

Table

Description automatically generated

4.

Table

Description automatically generated

The p-value of every predictor is larger than 0.05. So we should not include all variables

4b.

Text

Description automatically generated

From the above R result, we will select modelF. If we see adjusted R squared or AIC, modelF has the largest adjusted R square, smallest AIC, close distance in Cp = p, and the second smallest BIC.

Chart, scatter chart

Description automatically generated

4c).

pairs(q4data[,c(1,2,3,9,11)])

A picture containing calendar

Description automatically generated

From the pairwise scatterplots, it indicated that there is a strong linear relationship between Y and X1, X2, X8, X10. Therefore, it suggests that there may be linear relationships between Y and the 11 predictors.

4d

Step1 (X1 is selected):

Text

Description automatically generated

The model is significant.

Step 2: X5 is selected

Text

Description automatically generated

From the result above, p-value of X5 is 0.477 > 0.05, and t-value is also not significant, so we should not include X5. Final model is Y = X\_1 + epsilon