STAT 608 HW 3

Jack Cunningham (jgavc@tamu.edu)

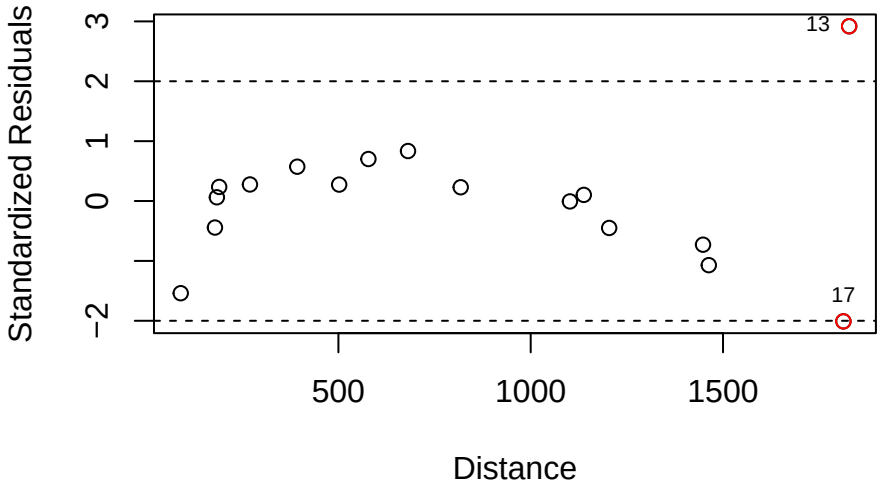
9/23/24

1.

a)

The business analyst claims that this model is highly effective for understanding the effects of Distance on Fare and predicting future values of Fare, there a few issues with this conclusion. First when we look at the standard residual plot we see that there appears to be a discernible quadratic pattern. With $n = 17$ it is very important that residuals be distributed i.i.d $N(0, \sigma^2)$ in order to make inference on coefficient estimates. Additionally with **any** sized sample residuals must be distributed i.i.d $N(0, \sigma^2)$ in order to create prediction intervals about the response variable.

With our current model there are also two outlier points I have labeled in red below. These need to be further analyzed.


Standardized Residuals
3
2
1
0
−2
13
17
500
1000
1500
Distance