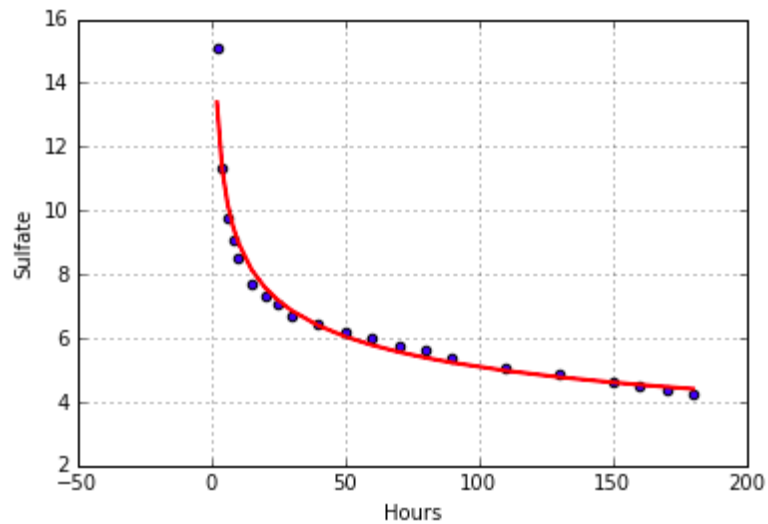# CS 498 HW 4 Report

gcui3,colinli2,sdai7
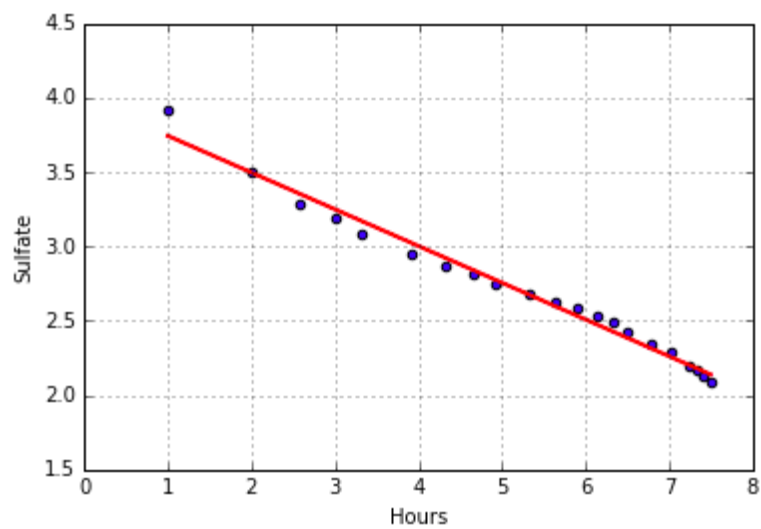
7-9)

The details of the implementation can also be viewed at:

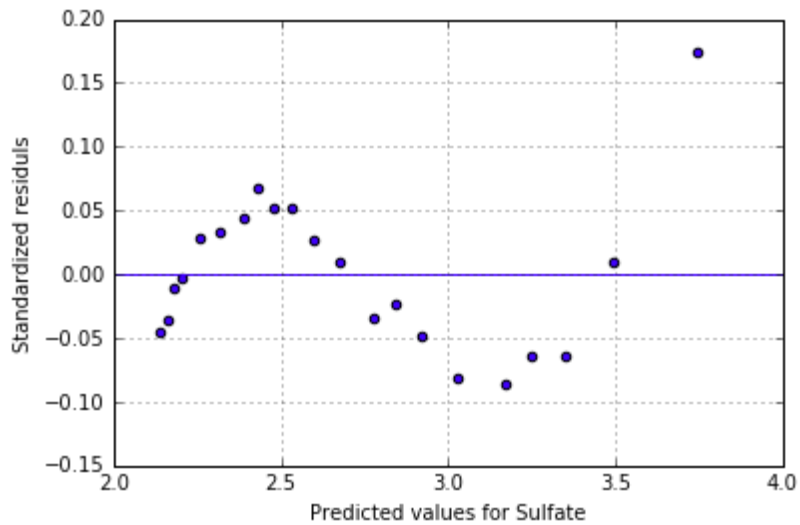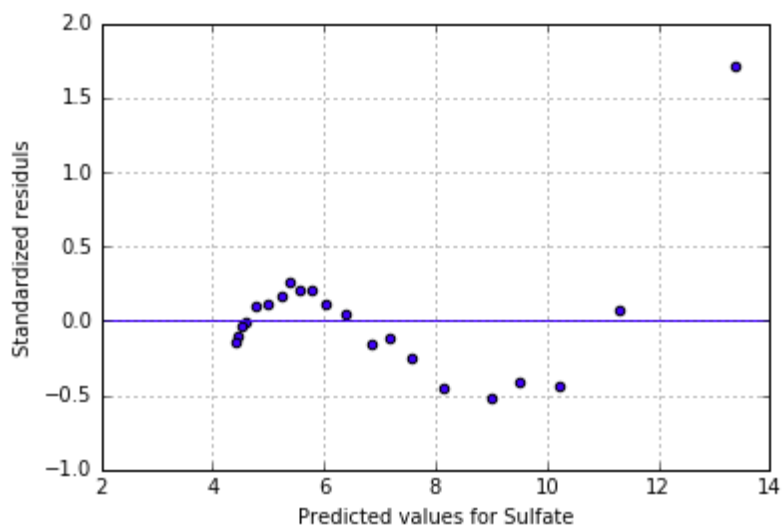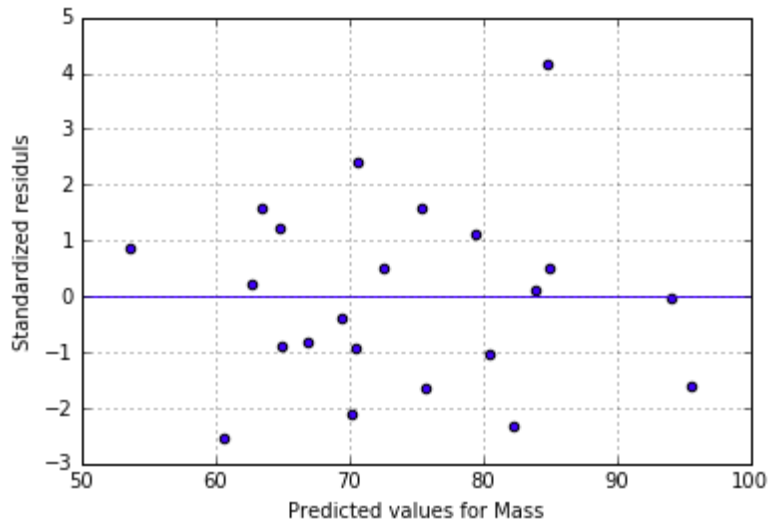**7-9_7-10.ipynb**
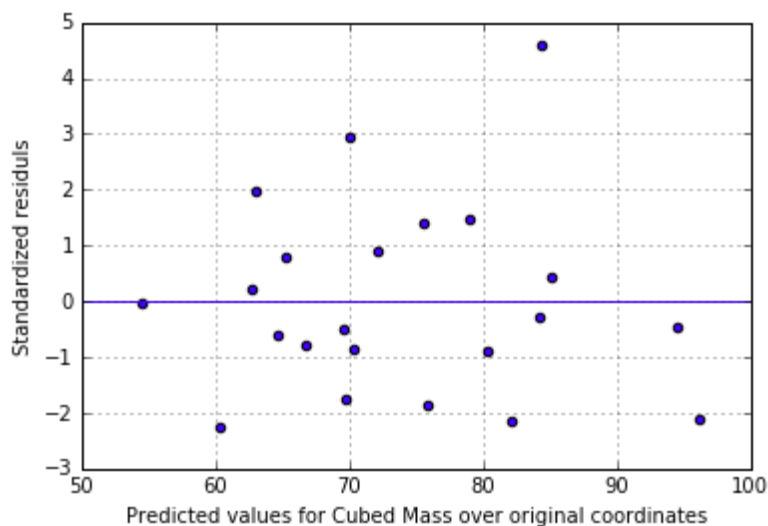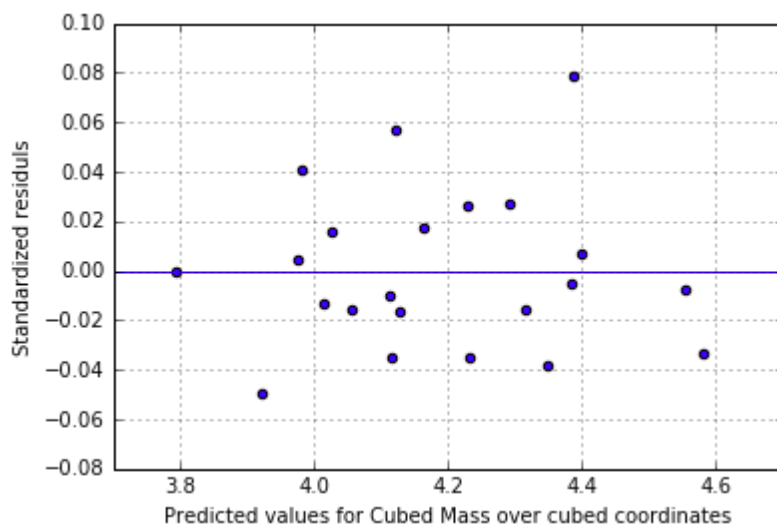
a.

b.

c.log-log:

original:

d. There is a sin pattern in both regression, the regression could be improved. The log-log version has smaller sum of squared residuals.
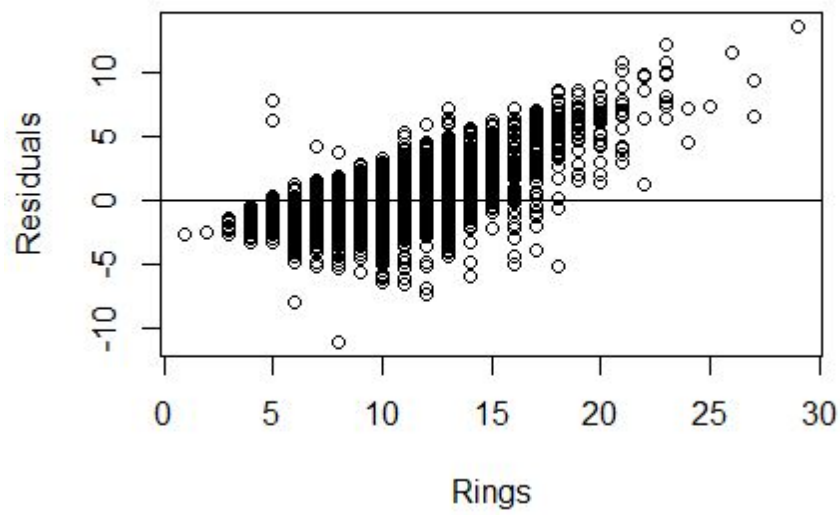
7-10)

a.

b.

c.
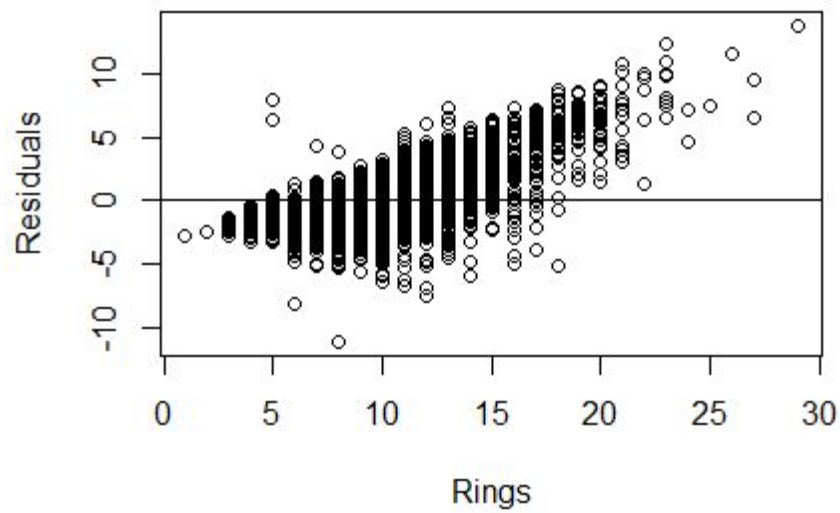The points are equally spread over on each side of y=0 in both regression, in general there aren't clear pattern, but the Cubed Mass has lower R square value, which is 0.976. The original data has a R square 0.977. So I would believe the cubed mass regression is better.
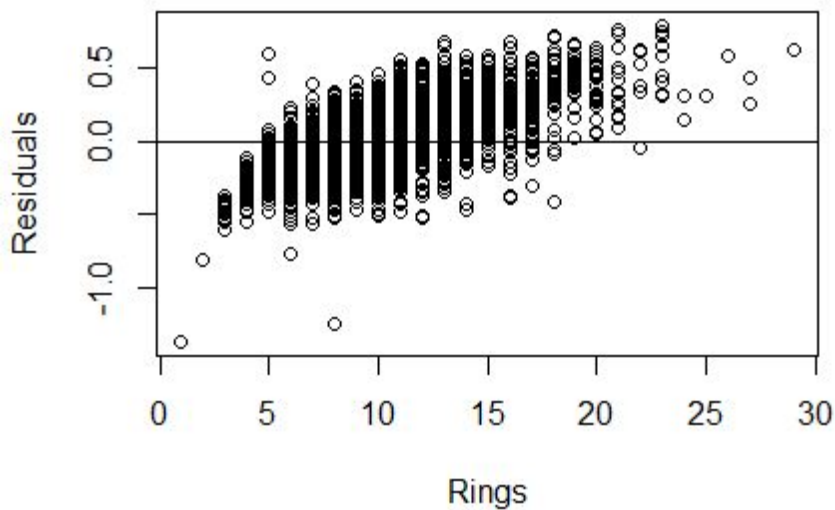
7-11)

A) Residuals Predicting Rings w/o Gender
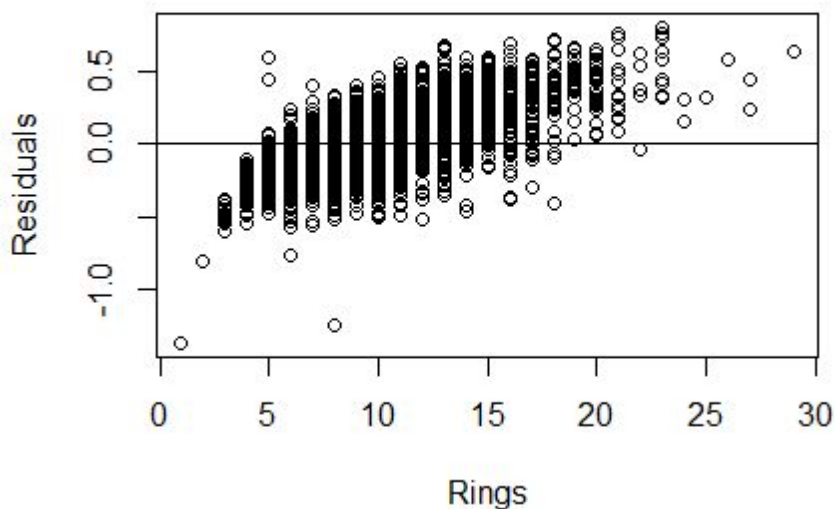

B) Residuals Predicting Rings w/ Gender

## C) Residuals Predicting log(Rings) w/o Gende

D) Residuals Predicting log(Rings) w/ Gender

E)
If I had to replace the ring counting procedure with one of these linear models, I would use the model created in part C or D. All four of the linear models overpredict younger Abalone snails and underpredict older ones, but the residual values of the two models where we used the log of their ages are overall much smaller, being from ~(-1.5 to) ~.7 as opposed to ~ (-10) to ~12. While some of this difference can be attributed to the fact that the values are overall smaller due to the log() transformation, $e^{1.5}$ is only about 4.5, which is still less than the original highest residual. Overall, the data in the models using log(Rings) is much more closely clustered to the fitted line.

The Genders plotted against Age with M=1, F = -1, I = 0 look like:
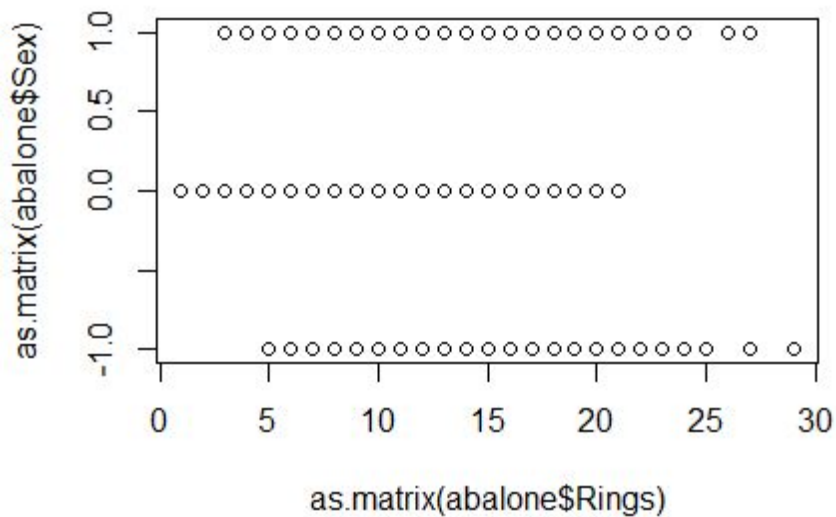
The distribution of "I" gendered Abalone were overall slightly younger than the other two, and the "F" gendered Abalone distribution was slightly older than the other two. However, even with the genders re-assigned to {I = -1, M = 0, F = 1} to follow this order, the impact on the accuracy of the linear models was negligible, likely due to the fact that the three distributions still overlap massively despite their slightly different centers.
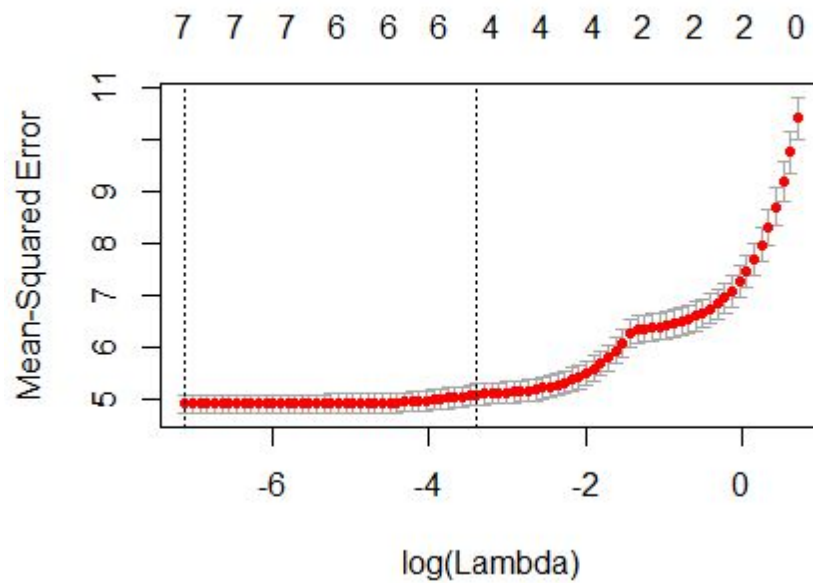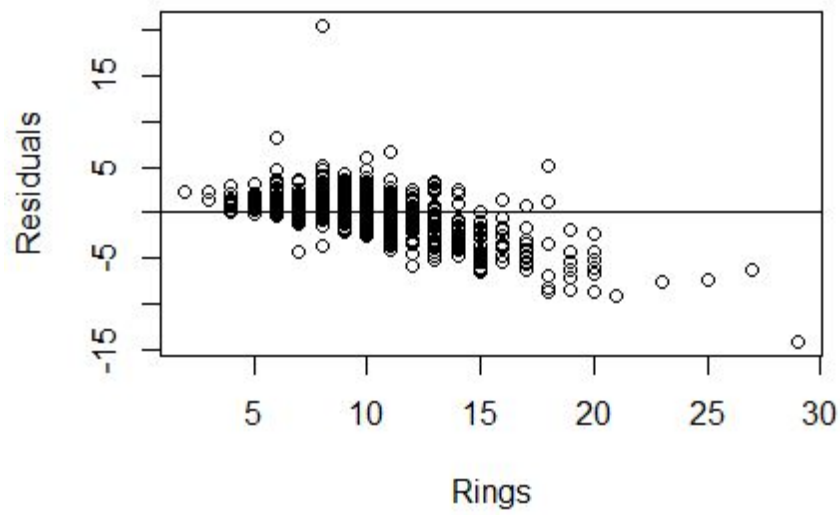
F)

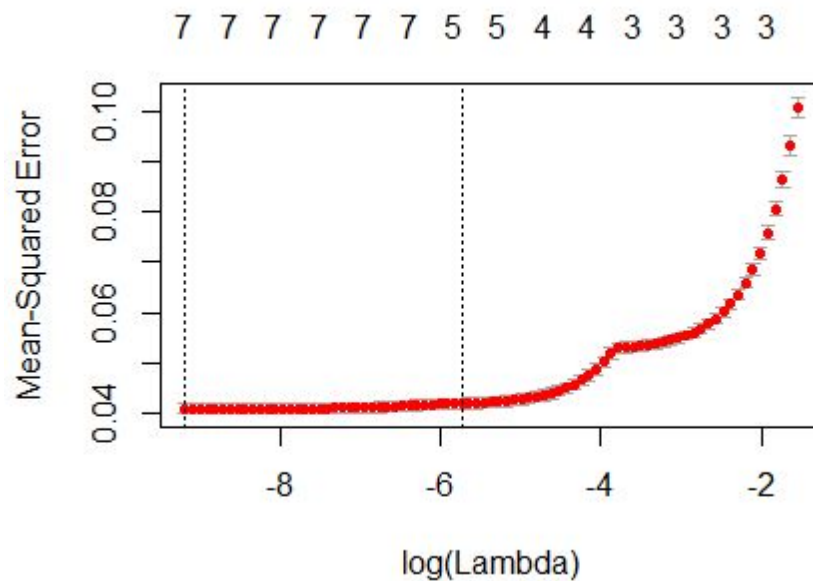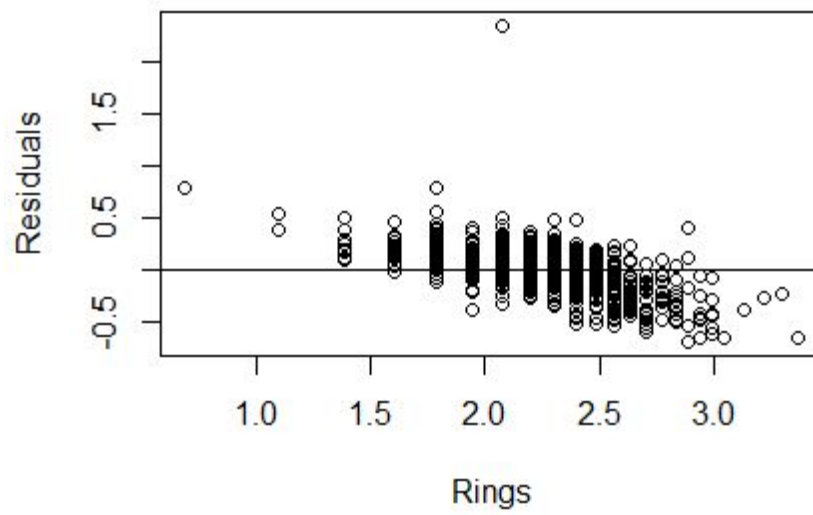## F-a) Regularizer predicting Rings at minimum Lambda w/o Sex
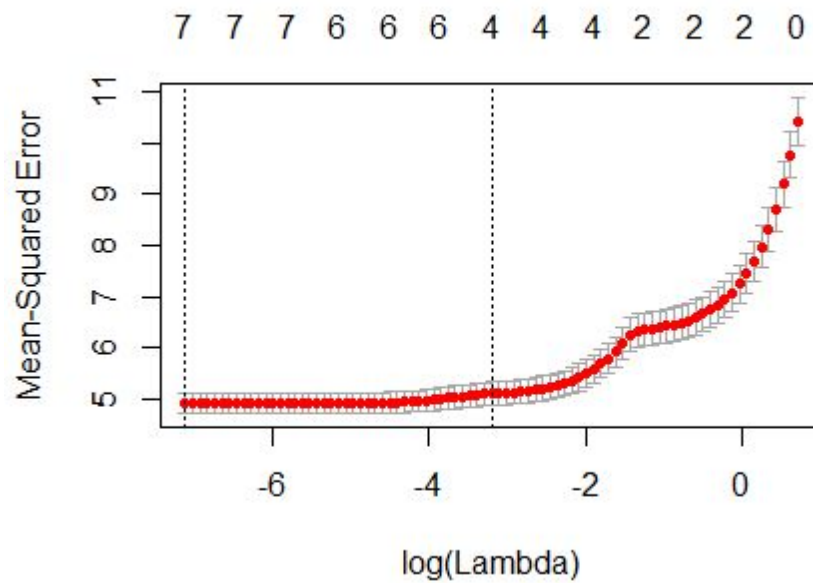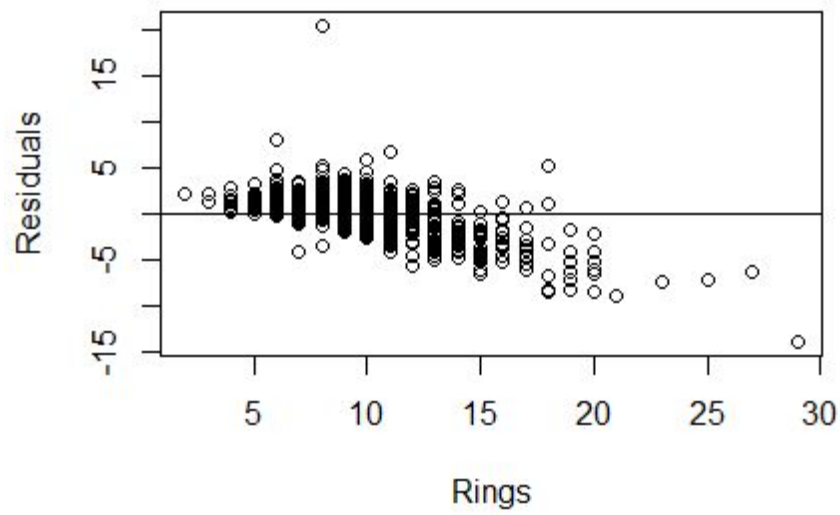
MSE against log(Lambda) for f-a)

F-b) Regularizer predicting Rings at minimum Lambda w/ Sex
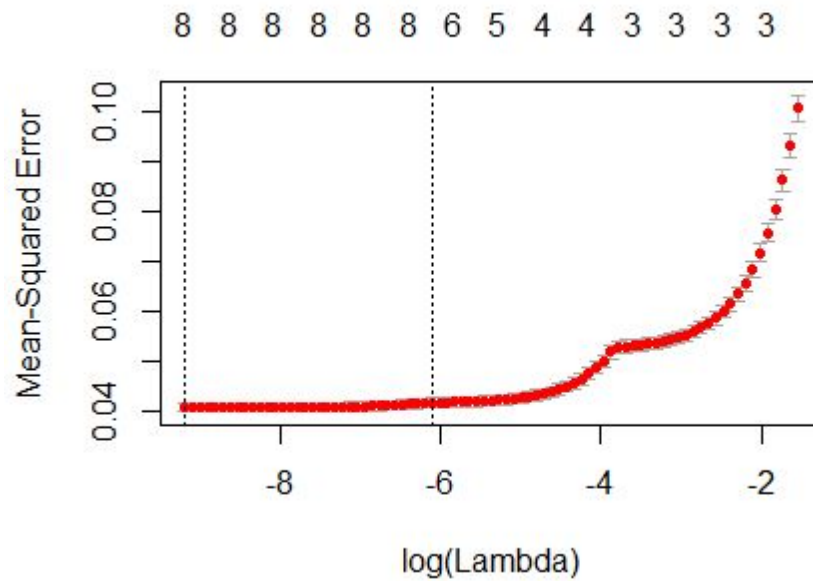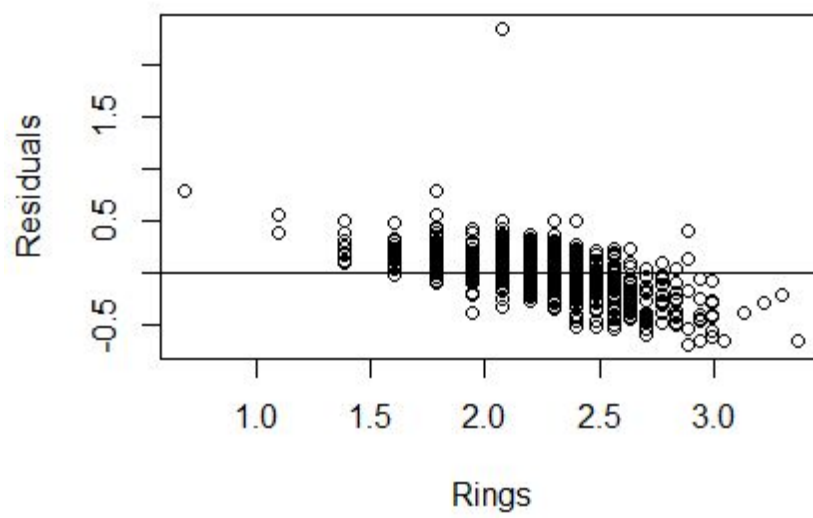

MSE against log(Lambda) for f-b)

F-c) Regularizer predicting log(Rings) at minimum Lambda w/o Sex

MSE against log(Lambda) for f-c)

The regularizer's prediction residuals are much more closely clustered around the model, with the few outliers being much further away from the rest of the points.

F-d) Regularizer predicting log(Rings) at minimum Lambda w/ Sex

MSE against log(Lambda) for f-d)