Assumptions

Ethan Shen and Steven Herrera

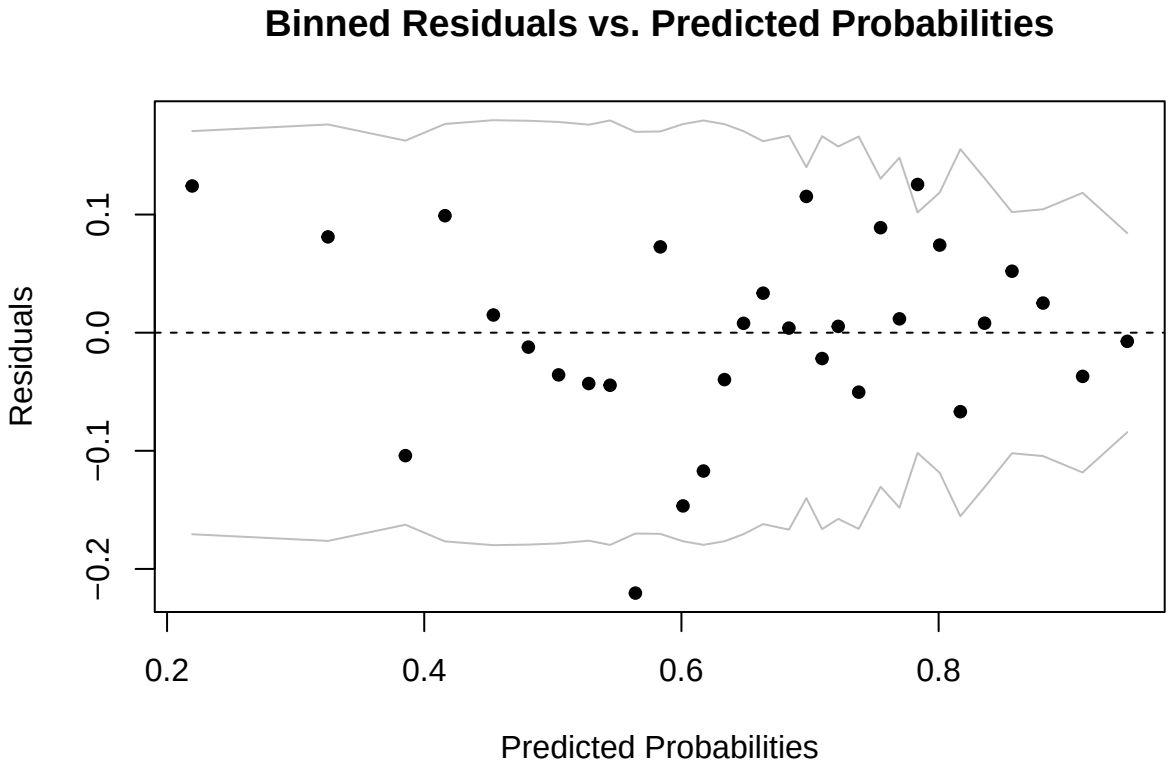
11/27/2018

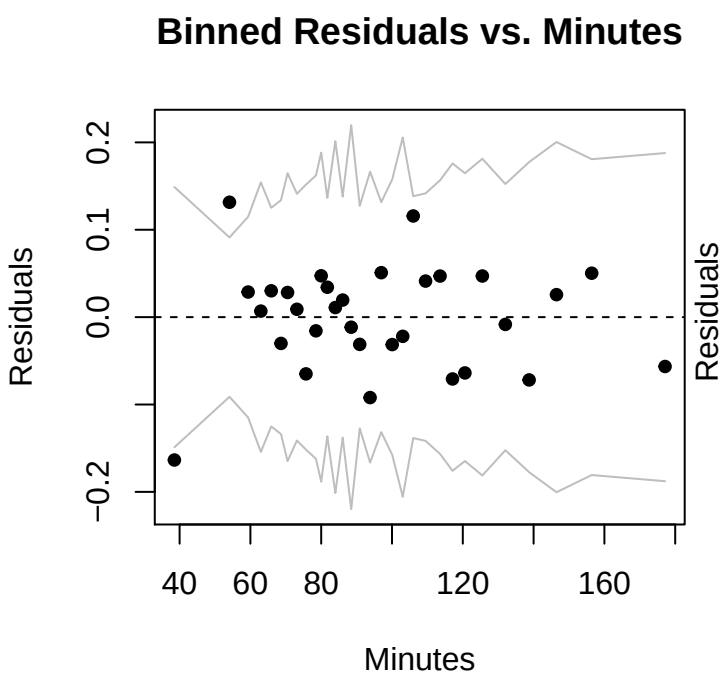
Linear Regression Assumptions

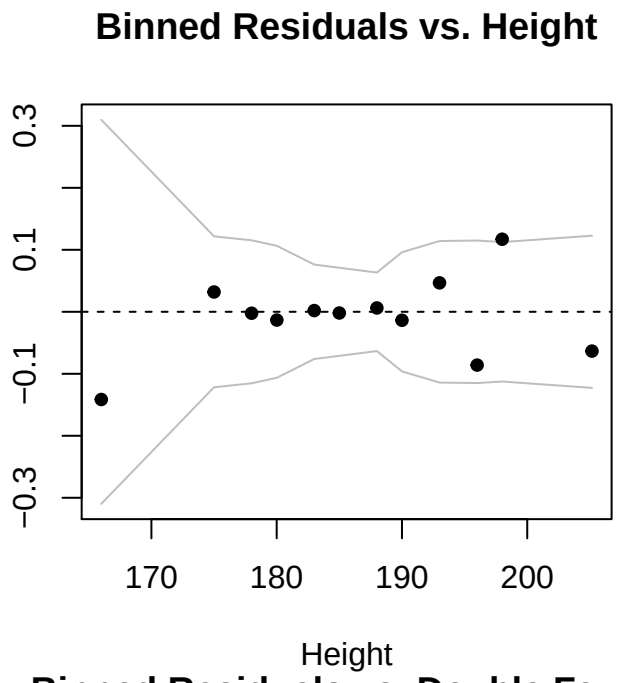
Model Assessment

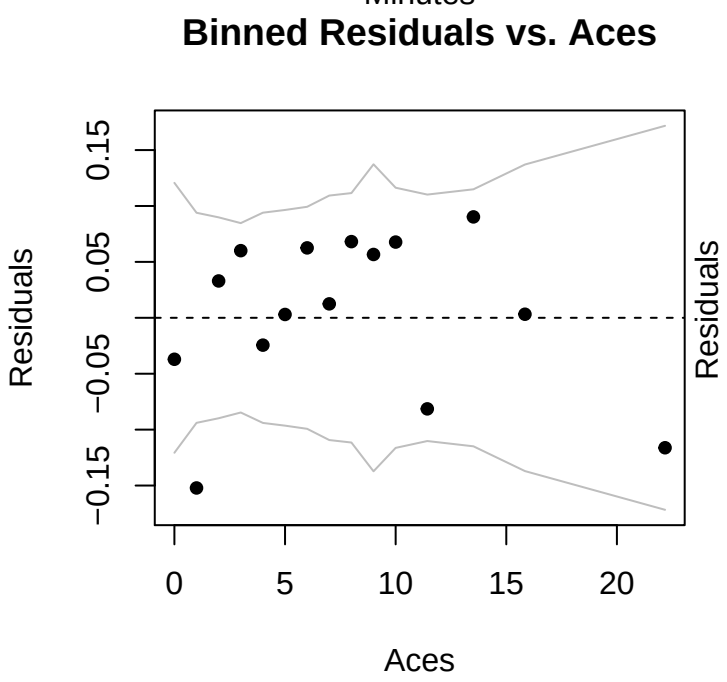
Binned Plots with Residuals vs Predicted

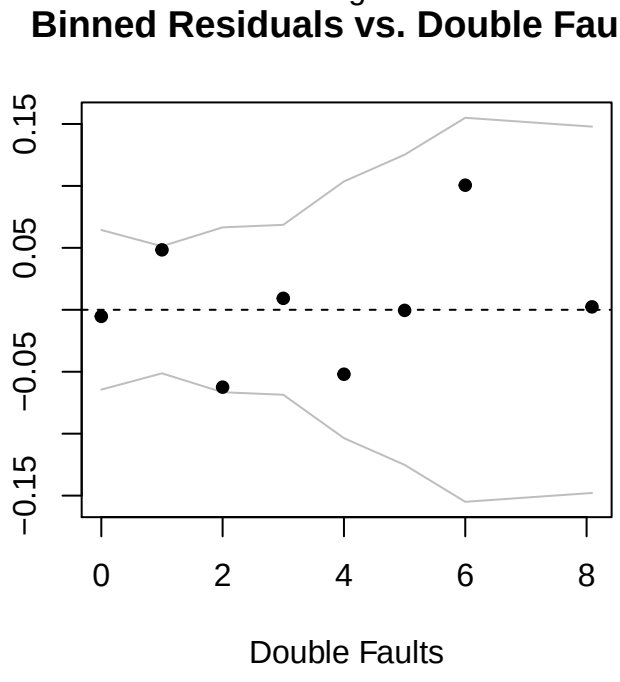
```
ten <- ten %>% mutate(Residuals = residuals.glm(final.base.model,type="response"),
                      Predicted = predict.glm(final.base.model,type="response"))

binnedplot(ten$Predicted, ten$Residuals,xlab="Predicted Probabilities",
           ylab="Residuals",main="Binned Residuals vs. Predicted Probabilities")
```

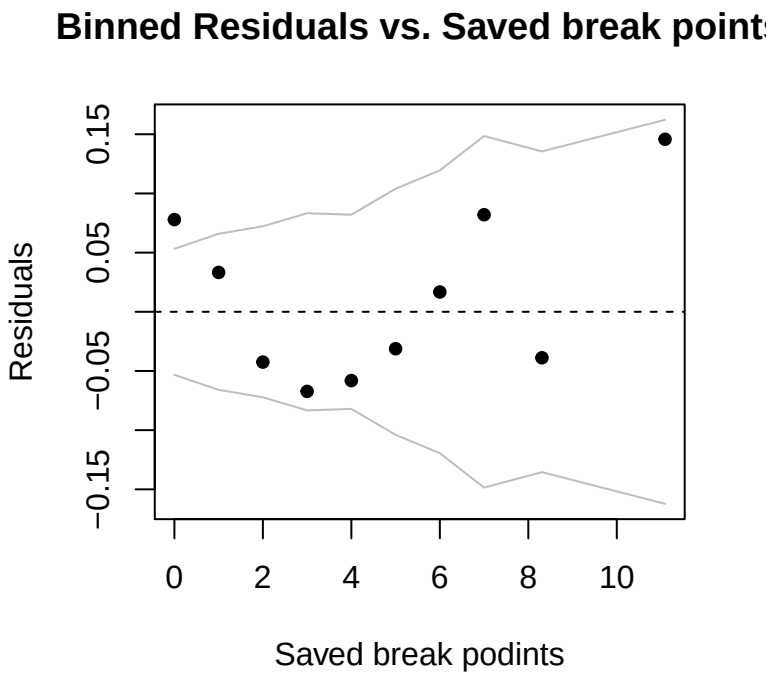



Binned Residuals vs. Minutes
Residuals
0.2
0.1
0.0
−0.2
40
60
80
120
160
Minutes


Binned Residuals vs. Height
Residuals
0.3
0.1
−0.1
−0.3
170
180
190
200
Height

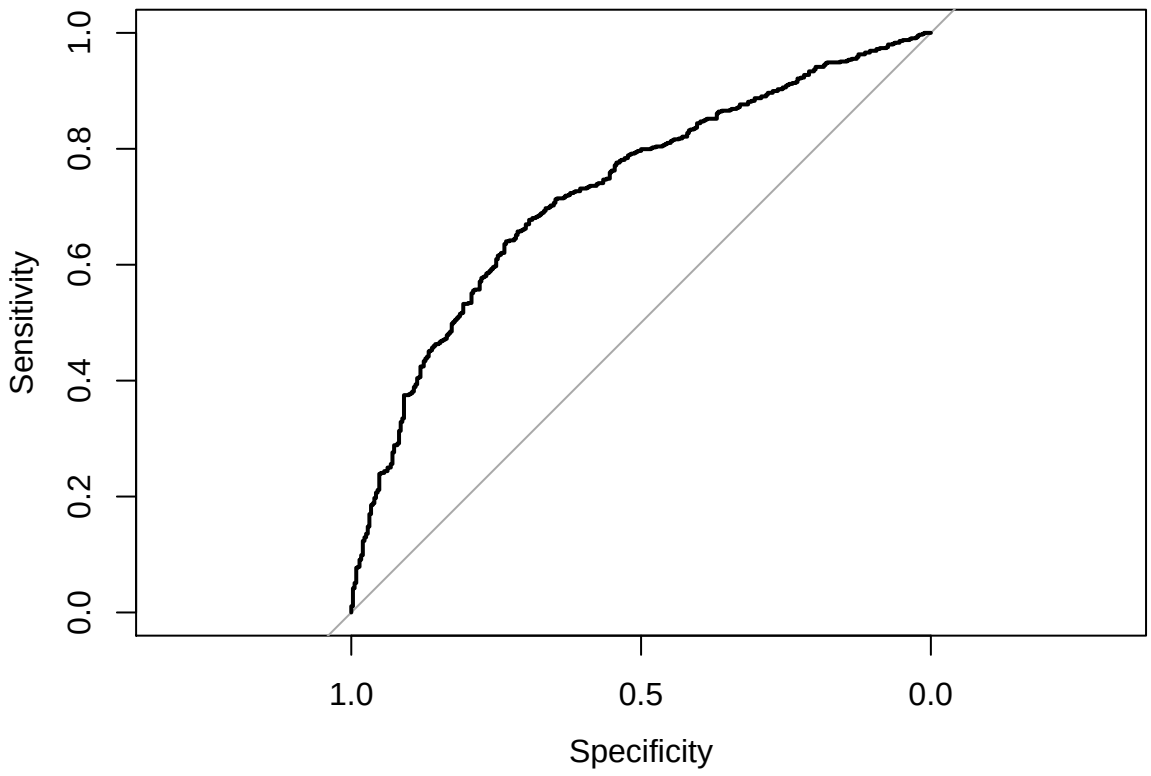

Binned Residuals vs. Aces
Residuals
0.15
0.05
−0.05
−0.15
0
5
10
15
20
Aces


Binned Residuals vs. Double Faults
Residuals
0.15
0.05
−0.05
−0.15
0
2
4
6
8
Double Faults



Looking at the binned residual plots, we see that all of the plots except for the binned residuals vs. saved break point have random scatter. The binned residuals vs. saved break shows a pattern. This is a violation of the assumptions.

```
ROC.ten <- roc(ten$status,ten$Predicted,plot=T)
```




```
ROC.ten$auc
```

```
## Area under the curve: 0.7268
```

```r
threshold = 0.30
table(ten$status, ten$Predicted > threshold)
```

```
##
##     FALSE TRUE
##   0    26  326
##   1    13  635
```

```r
(326 + 13)/(14+13+326+635)
```

```
## [1] 0.3431174
```