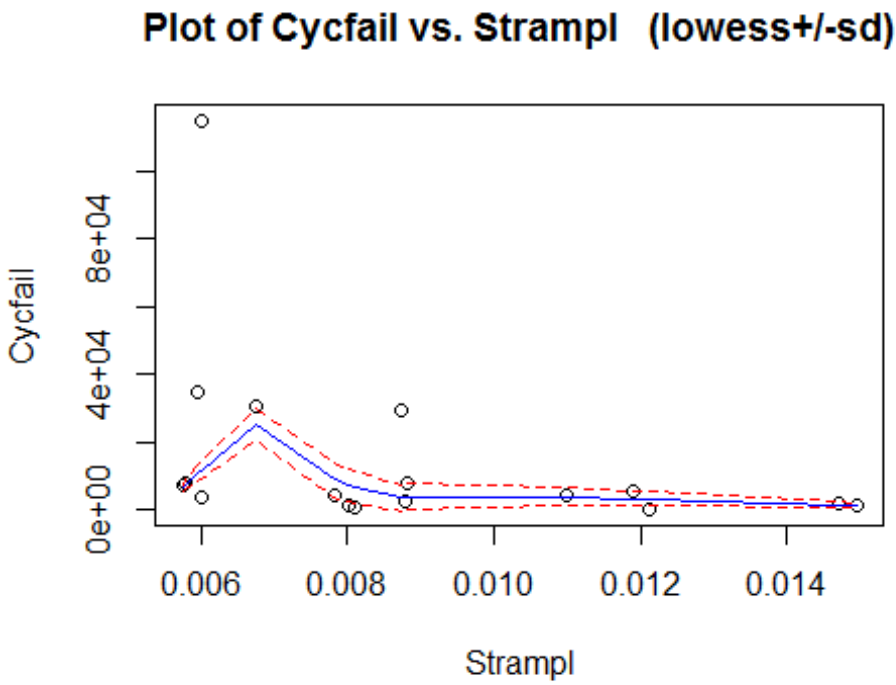
Task One

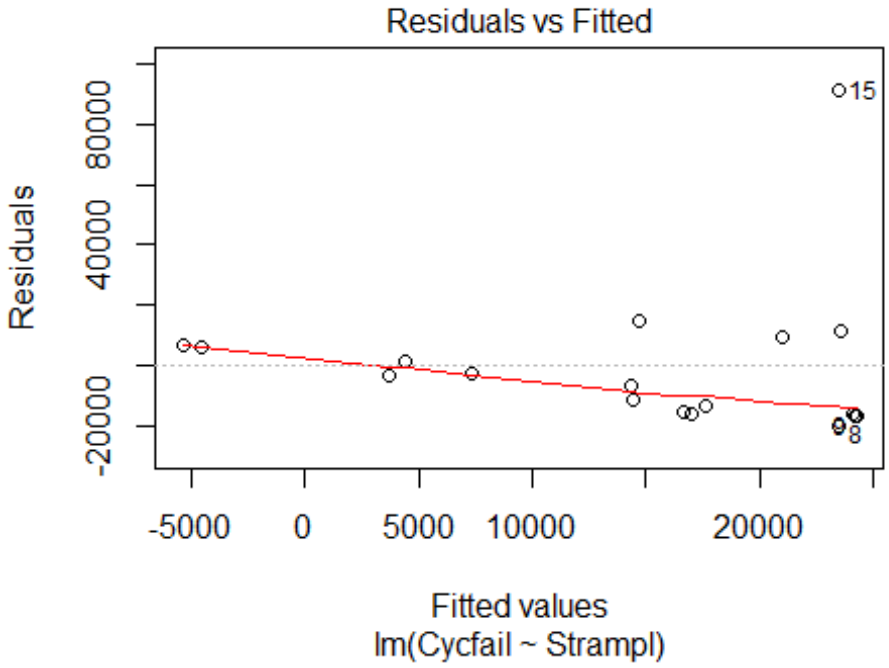
Sang Hyun Kho

9 September 2016

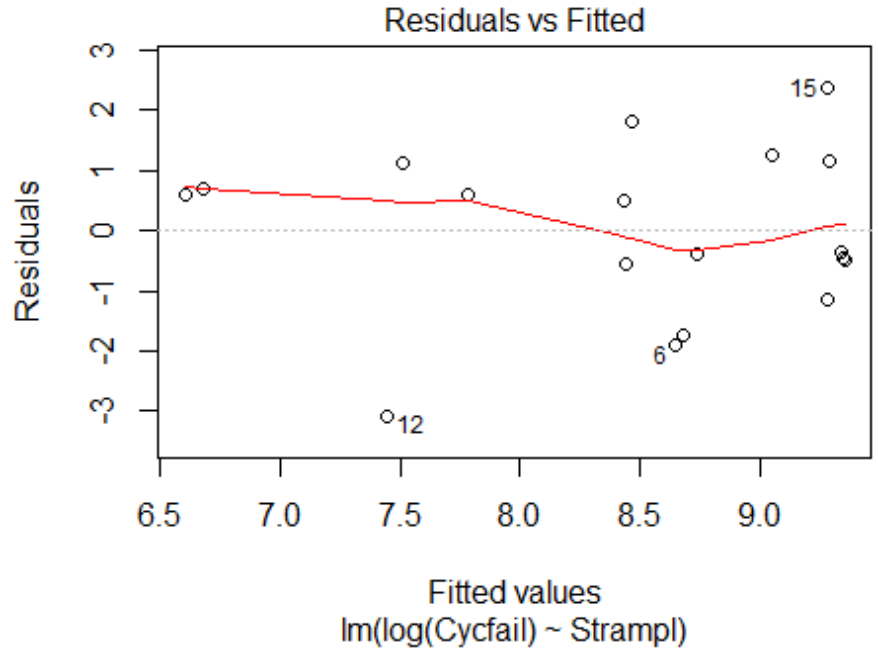
```
library(s20x)
setwd("C:/Users/Buzz/Desktop/Uni/MM3/AssignmentTwo")
gasturbines.df = read.table(file.choose(), header = TRUE) # read text file
trendscatter(Cycfail ~ Strampl, data = gasturbines.df)
```



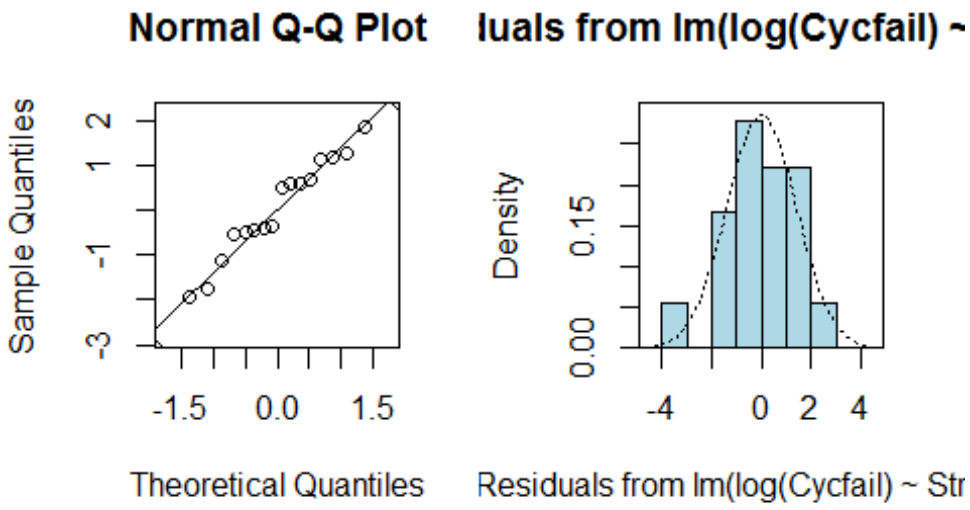

```
gasturbines.lm1 = lm(Cycfail ~Strampl, data = gasturbines.df)
# fitting a linear model
plot(gasturbines.lm1, which = 1)
```



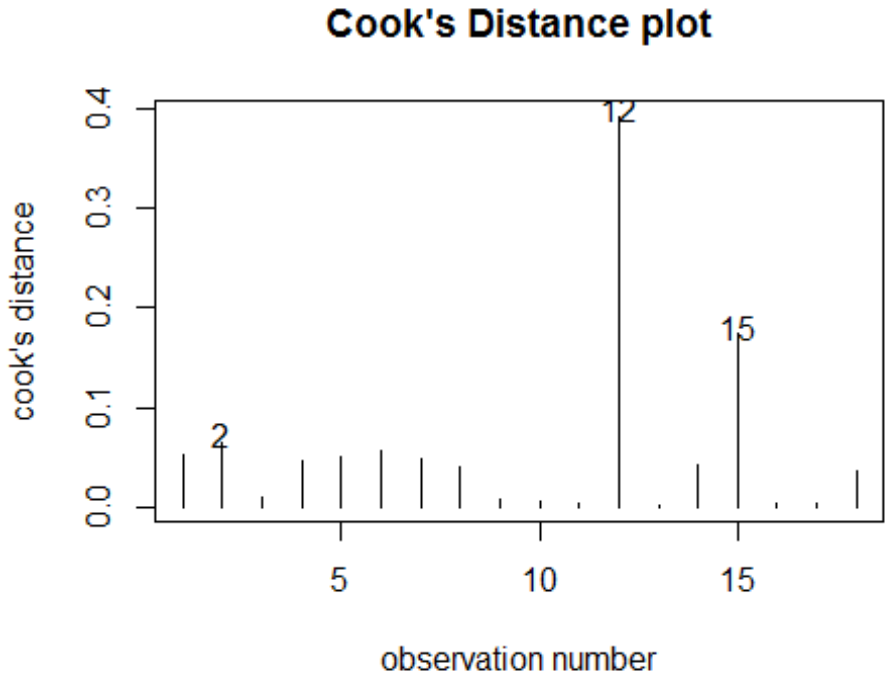

```r
gasturbines.lm2 = lm(log(Cycfail) ~ Strampl, data = gasturbines.df)
# logging the Cycfail and re-fit the linear model
plot(gasturbines.lm2, which = 1)
```



```
normcheck(gasturbines.lm2)
```




```
cooks20x(gasturbines.lm2)
```



```r
summary(gasturbines.lm2)
```

```
##
## Call:
## lm(formula = log(Cycfail) ~ Strampl, data = gasturbines.df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.07869 -0.53533  0.06447  1.02150  2.37065
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   11.076      1.063  10.418 1.55e-08 ***
## Strampl     -299.357    115.522  -2.591   0.0197 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.441 on 16 degrees of freedom
## Multiple R-squared:  0.2956, Adjusted R-squared:  0.2516
## F-statistic: 6.715 on 1 and 16 DF,  p-value: 0.01968
```

```r
exp(0.001*confint(gasturbines.lm2))
```

```
##                2.5 %    97.5 %
## (Intercept) 1.008861 1.0134195
## Strampl     0.580275 0.9469961
```

Method and Assumption Checks

The scatter plot of cycles to failure vs strain amplitude showed clear non-linearity. Residuals from a simple linear repression model suggested clear failure of equality of variance. Observation 12 showed up as close to 0.4 which is our rule of thumb value, but we retained the observation as it is still less than 0.4 and it is expensive to collect data. Theory suggested a multiplicative relationship, so a log transformation was attempted. We assume the observations are independent. A simple linear model fitted to log cycles to failure satisfied the equality of variance and Normality assumptions. There was no unduly influential observations.

The final model is:

$\log(Cycfail_i) = \beta_0 + (\beta_1 \text{ x } Strampl_i) + \varepsilon_i$, where $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$.

Executive Summary

We are interested in seeing if the new material, a single crystal nickel base super alloy can solve the problem of failures in aircraft gas turbine engine due to high cycle fatigue. The cycles to failure decreases in value with a decaying relationship relative to strain amplitude i.e. the rate of cycles to failure decreases with an increase in strain amplitude.

We estimate that 0.001 strain amplitude increase is associated with a decrease in median cycles to failure by between 0.58 and 0.95.

Our model explains almost 30% of the variation in the cycles to failure which is not really a reasonable model for prediction.