Engsci 211 Task 1: Sleep Deprivation

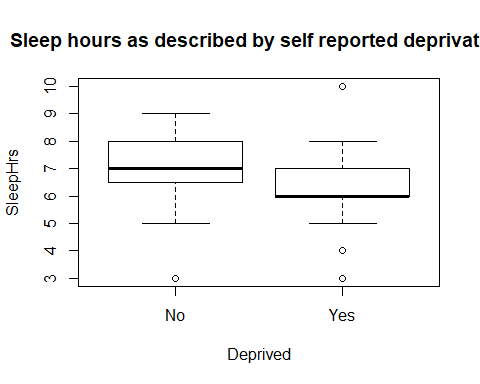
Alex Verkerk

library(s20x)  
sleepDep.df = read.table("sleepDeprivation.txt", header = TRUE)  
head(sleepDep.df)

## Deprived SleepHrs  
## 1 No 8  
## 2 No 6  
## 3 No 7  
## 4 Yes 6  
## 5 Yes 6  
## 6 Yes 7

### **Exploratory analysis**

boxplot(SleepHrs ~ Deprived, main = "Sleep hours as described by self reported deprivation", data = sleepDep.df)

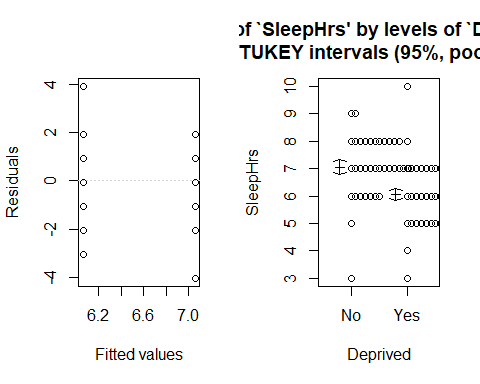


summaryStats(SleepHrs ~ Deprived, data = sleepDep.df)

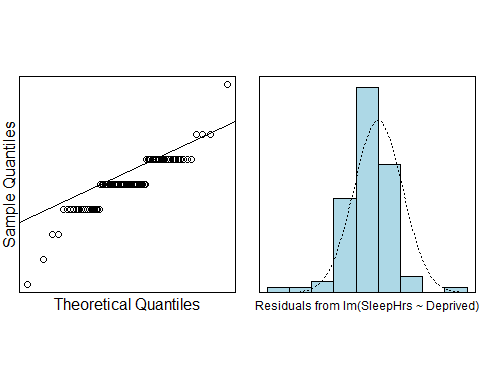
## Sample Size Mean Median Std Dev Midspread  
## No 35 7.057143 7 1.161714 1.5  
## Yes 50 6.060000 6 1.057722 1.0

* Centre: Mean sleep hours for deprived is lower than mean sleep hours for non deprived as seen by the summary stats.
* Spread: The spread is slightly wider (1.5) for non deprived as compared to deprived (1.0)
* Skew: Both Yes and No appear fairly symmetrical without too significant of a tail, so no significant skew. The boxplots for the two appear to be quite similar.

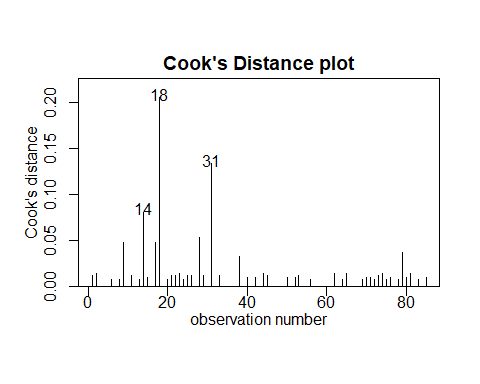
layout20x(1,2)  
eovcheck(SleepHrs ~ Deprived, data = sleepDep.df)  
onewayPlot(SleepHrs ~ Deprived, data = sleepDep.df)



normcheck(lm(SleepHrs ~ Deprived, data = sleepDep.df))



cooks20x(lm(SleepHrs ~ Deprived, data = sleepDep.df))



t.test(SleepHrs ~ Deprived, data = sleepDep.df, var.equal = TRUE)

##   
## Two Sample t-test  
##   
## data: SleepHrs by Deprived  
## t = 4.1075, df = 83, p-value = 9.329e-05  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 0.5143015 1.4799842  
## sample estimates:  
## mean in group No mean in group Yes   
## 7.057143 6.060000

### Methods and Assumption Checks

We are attempting to find the mean difference between two groups of measurements, and as such we need to take a two sample t-test. We assume the individual reports of sleep hours and deprivation are independant measurements.

We checked Equality of Variance, and found that the EOV check for Yes and No appears to show equality of variance, however due to the way measurements were collected (as integer levels) this means that there could be no equality of variance as clusters of values will collapse into a single point. The one-way plot we did alongside the EOV check shows that indeed there are clusters of values at certain levels, but our assumption of equality of variance is satisfied, as the plots appear to be spread similarly.

We then checked the normality of the sample, and while the Q-Q plot looked slightly abnormal due to the integer levels of measurements, the histogram showed that our assumption of normality is satisfied. On the cooks distance plot, no influential observations were observed.

This allows us to use a standard two sample t-test. This had a p-value of which is less than 0.05. This shows that the difference between the two groups is statistically significant.

The model fitted is where is the effect of whether the person self-reports feeling sleep deprived or not, and .

### Executive Summary

This data was collected to see if those who self-report sleep deprivation sleep on average for less time than those who self-report not feeling sleep deprived.

We have observed that those who self-report that they feel sleep deprived do sleep, on average, less than those who self-report not feeling sleep deprived.

We estimate that students who self-reported no sleep deprivation slept for an average of 0.51 to 1.48 more hours than those who self-reported being sleep deprived.