STAT4100 HW2

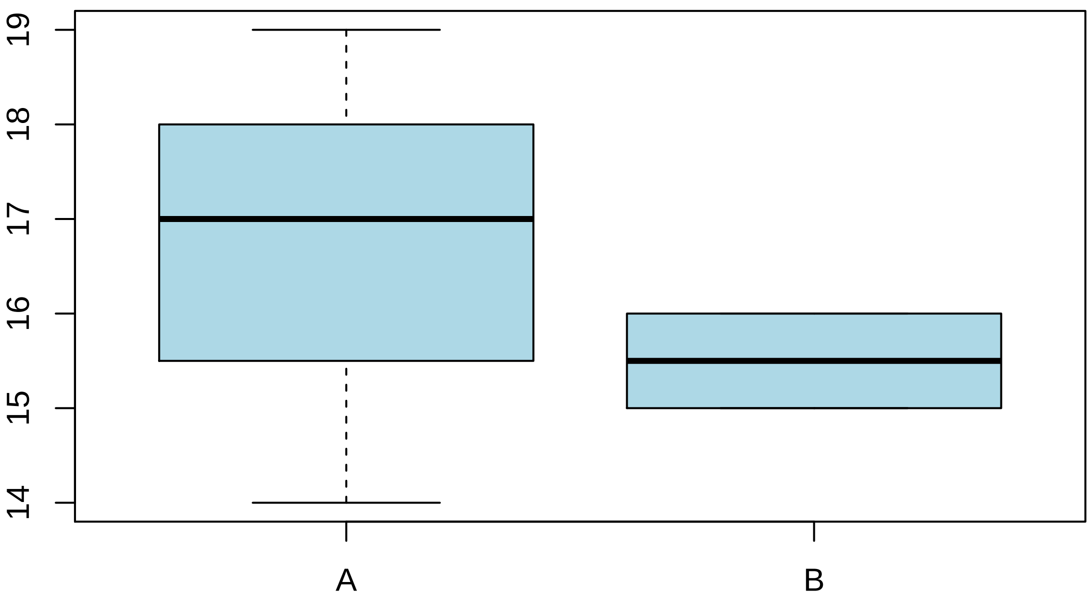
Cody Frisby

January 14, 2016

#3

- First I will bring the data into R, plot it, and run a quick t test.

```
dat <- data.frame(treatment = c("A", "B", "A", "A", "B"),
                  response = c(14, 16, 19, 17, 15))
boxplot(dat$response ~ dat$treatment, col = "lightblue")
```



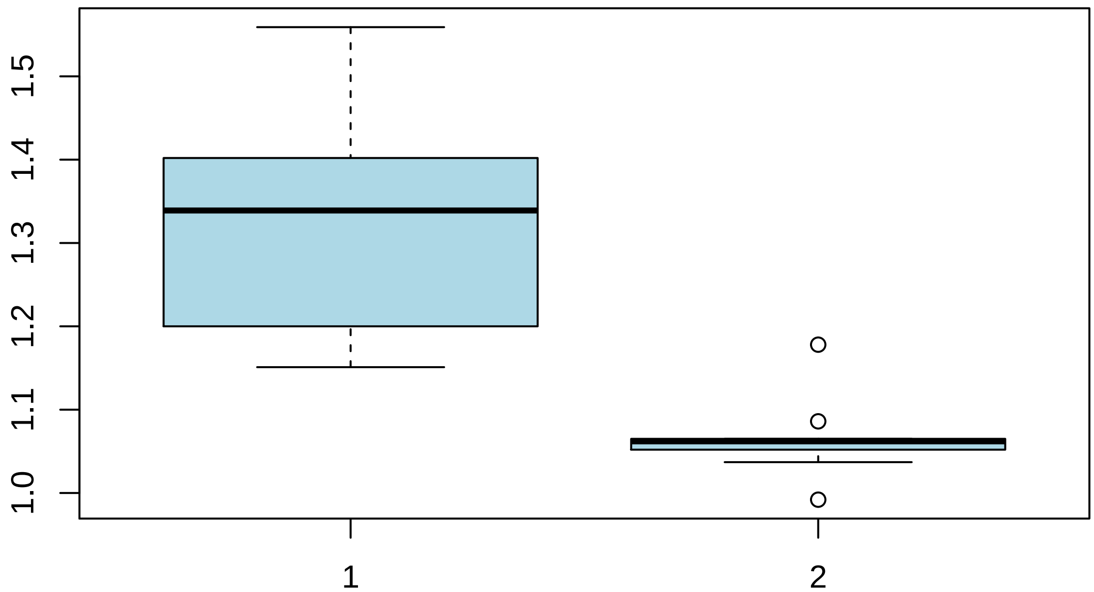

```
t <- t.test(response ~ treatment, data = dat, var.equal = TRUE, alternative = "less")
```

Do not reject null hypothesis. P value = 0.7075314.

#4 Two procedures for predicting shear strength will be compared.

```
Karl <- c(1.186, 1.151, 1.322, 1.339, 1.2, 1.402, 1.365, 1.537, 1.559)
Lehigh <- c(1.061, 0.992, 1.063, 1.062, 1.065, 1.178, 1.037, 1.086, 1.052)
s <- t.test(Karl, Lehigh, paired = TRUE)
boxplot(Karl, Lehigh, col = "lightblue")
```

a) Is there evidence to support the claim that there is a differnce between the two methods?




```
s
```

```
##
##  Paired t-test
##
## data:  Karl and Lehigh
## t = 6.0819, df = 8, p-value = 0.0002953
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1700423 0.3777355
## sample estimates:
## mean of the differences
##               0.2738889
```

Yes, there is evidence to suggest a difference in mean performance between the two methods at $\alpha = 0.05$.

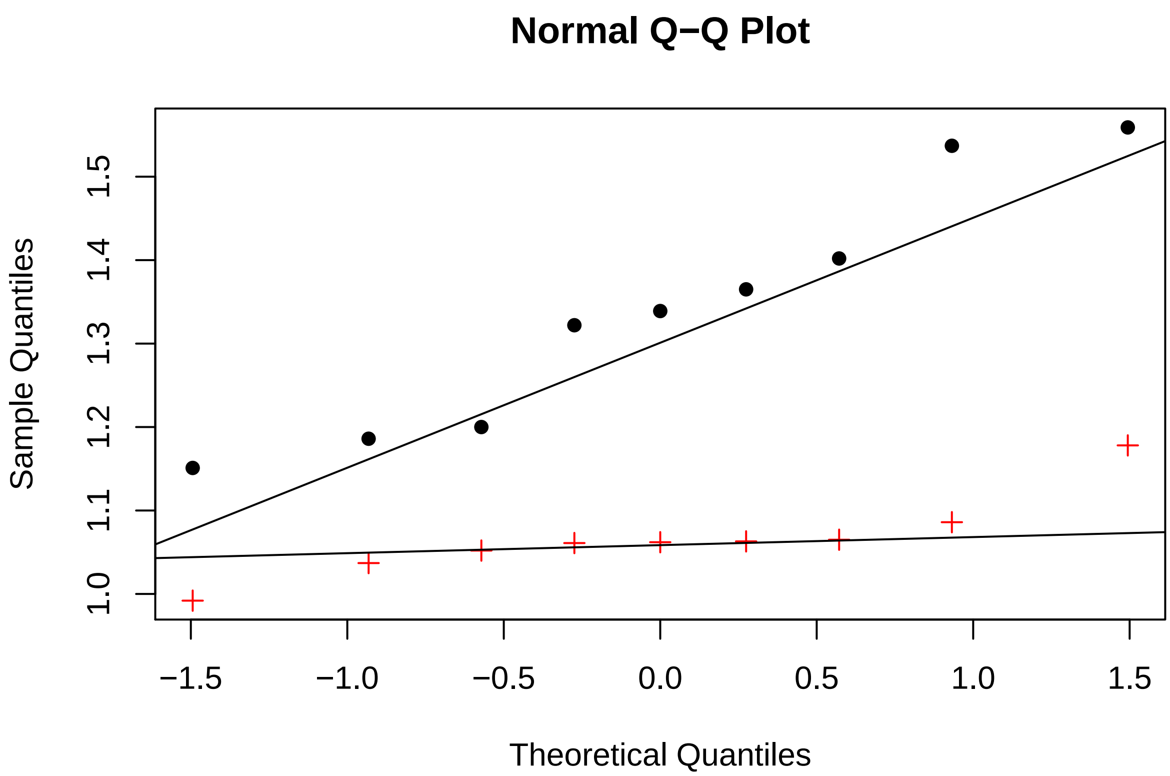
b) The p value for this test is 2.9529546×10^{-4}. This is much smaller than 0.05.

c) Confidence Interval: $[0.1700423, 0.3777355]$. Zero is not contained in this interval.

d) Normality Assumptions:

- QQ Plot for the Karl and Lehigh data.

```
q1 <- qqnorm(Karl, plot.it = FALSE)
q2 <- qqnorm(Lehigh, plot.it = FALSE)
plot(range(q1$x, q2$x), range(q1$y, q2$y), type = "n", xlab = "Theoretical Quantiles", ylab = "Sample Qu
points(q1, pch = 16)
points(q2, col = "red", pch = 3)
qqline(Karl); qqline(Lehigh)
```



The variances of the two samples is not equal. The var.test function in R can show this. The difference in the two lines also demonstrates this.

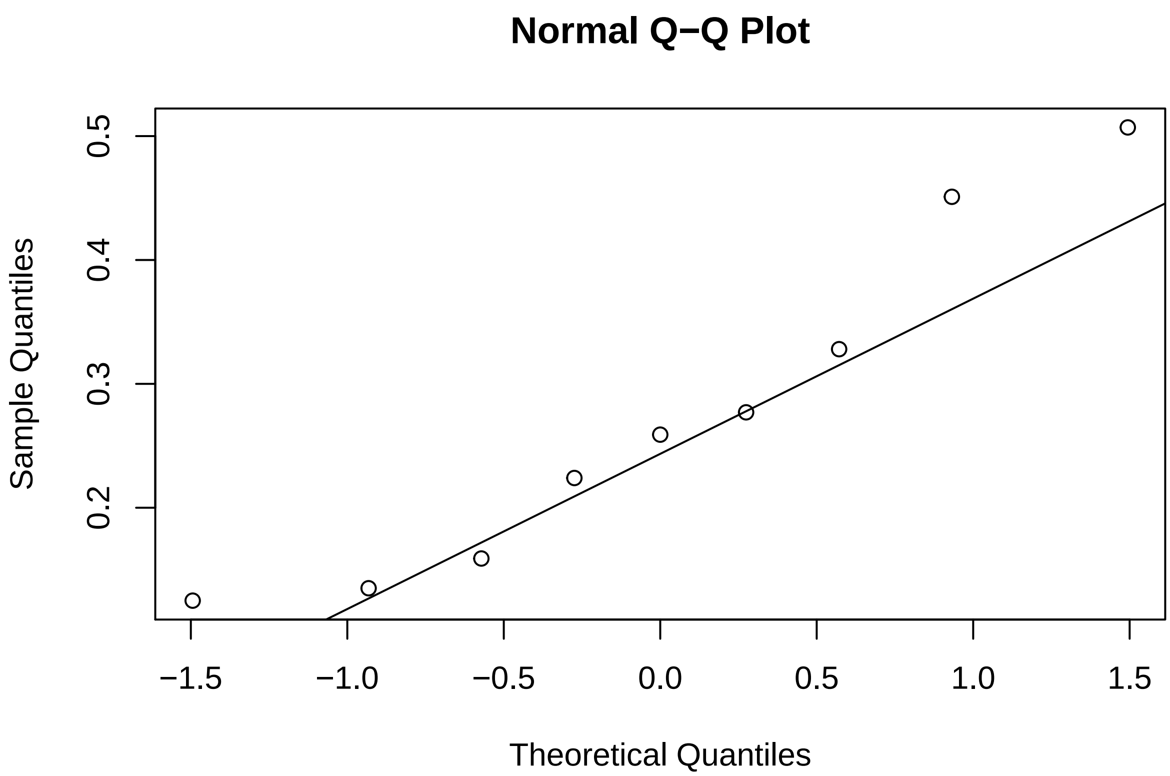
```
var.test(Karl, Lehigh)
```

```
##
##  F test to compare two variances
##
## data:  Karl and Lehigh
## F = 8.7454, num df = 8, denom df = 8, p-value = 0.006008
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##   1.972674 38.770520
## sample estimates:
## ratio of variances
##           8.745375
```

e) Normality for difference in ratios:

- Here we plot the difference between the two variables and run a shapiro.wilk test.

```
qqnorm(Karl - Lehigh)
qqline(Karl - Lehigh)
```




```r
shapiro.test(Karl - Lehigh)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Karl - Lehigh
## W = 0.91678, p-value = 0.3663
```

Looks like we can reasonably assume normality for the difference between the two methods.

f) Discuss the role of the normality assumption in the paired t test.

- The normality assumption is only moderately important with a paired t test just like other t tests. With a paired t test the assumption of normality applies to the differences. The individual sample measurements do not have to be normally distributed.

#5

```r
y1 <- 12.5
y2 <- 10.2
s1 <- 101.17
s2 <- 94.73
n1 <- 8
n2 <- 9
# compute the F statistic for 7,8 degrees of freedom at alpha=0.05
qf(0.975, 7, 8)
```

```
## [1] 4.528562
```

```r
s1/s2
```

```
## [1] 1.067983
```

```r
# is the ratio of the two variances greater?
s1/s2 > qf(0.975, 7, 8)
```

```
## [1] FALSE
```

a) No. 1.0679827 is not greater than 4.5285621. We cannot conclude that the variances are different.

```r
sp <- sqrt(((n1-1)*s1 + (n2-1)*s2) / (n1 + n2 - 2))
t0 <- (y1 - y2) / (sp * sqrt((1/n1) + (1/n2)))
qt(0.95, 15)
```

```
## [1] 1.75305
```

```r
t0
```

```
## [1] 0.4787886
```

```r
t0 > qt(0.95, 15)
```

```
## [1] FALSE
```

b) Do not reject. There is not evidence to indicate the new filtering device has affected the mean.