MAT 355

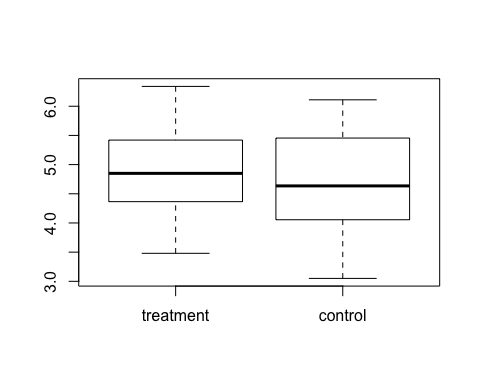
2.1a)

###read data#### local or web####  
dd=read.csv("/Users/celdrick/Downloads/plantweight.csv")

##Exploratory analysis of data###  
summary(dd)

## treatment control   
## Min. :3.480 Min. :3.050   
## 1st Qu.:4.388 1st Qu.:4.077   
## Median :4.850 Median :4.635   
## Mean :4.860 Mean :4.726   
## 3rd Qu.:5.390 3rd Qu.:5.393   
## Max. :6.340 Max. :6.110

boxplot(dd)



stem(dd$treatment)

##   
## The decimal point is at the |  
##   
## 3 | 568  
## 4 | 234477899  
## 5 | 024589  
## 6 | 03

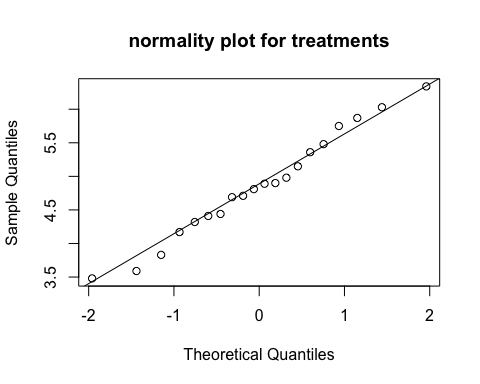
stem(dd$control)

##   
## The decimal point is at the |  
##   
## 3 | 1679  
## 4 | 0125567  
## 5 | 1223666  
## 6 | 11

The summary shows that Treatment has a higher minima and maxima than Control. Treatment has a mean of 4.860 and Control has a mean of 4.726. Their means are almost the same.

From the dot plots, Stem-and-leaf and probability plots it is clear that the observations of two groups are from normal distributions. Moreover mean is almost equal to median for both groups, it indicates that the data come from symmetric distribution.

qqnorm(dd$treatment, main = "normality plot for treatments")  
qqline(dd$treatment)

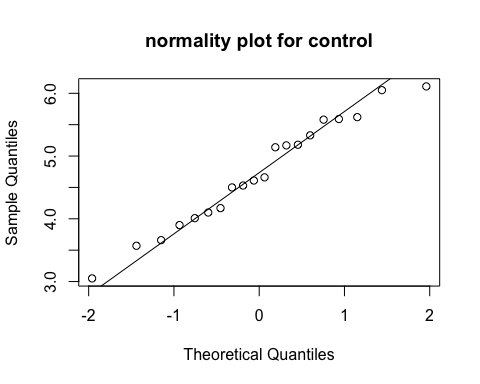


shapiro.test(dd$treatment)

##   
## Shapiro-Wilk normality test  
##   
## data: dd$treatment  
## W = 0.98061, p-value = 0.9417

Q-Q plot of the Treatment group shows that the data is perfectly normal wih a little variation. The Shapiro-wilk normality test which gives a p-value of 0.9417 supports this

qqnorm(dd$control, main = "normality plot for control")  
qqline(dd$control)



shapiro.test(dd$control)

##   
## Shapiro-Wilk normality test  
##   
## data: dd$control  
## W = 0.96909, p-value = 0.7356

Q-Q plot of the Control group shows that the data is quite normal wih a little variation. The shapiro-wilk normality test which gives a p-value of 0.7356. supports this.

2.1b)

###t-test###  
t.test(dd[,1],dd[,2])

##   
## Welch Two Sample t-test  
##   
## data: dd[, 1] and dd[, 2]  
## t = 0.50985, df = 37.711, p-value = 0.6131  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -0.3967069 0.6637069  
## sample estimates:  
## mean of x mean of y   
## 4.8600 4.7265

Assuming equal variances in the two groups, the unpaired t-test of mu\_1=mu\_2 against mu\_!≠mu\_2 gives t = 0.51, d.f. = 38, p-value = 0.613. We are 95% confident that the difference in mean yield between the two groups is between -0.3967069 and 0.6637069 and since the confidence interval includes 0, we can also conclude that there is no difference between the two groups..Hence we fail to reject

#sum(dd$control)  
#sum(dd$treatment)  
y\_bar=(sum(dd$control)+ sum(dd$treatment))/40  
y= (dd$control-y\_bar)^2  
x= (dd$treatment-y\_bar)^2  
S\_0=sum(x) + sum(y)  
S\_0

## [1] 26.23168

a= (dd$control-sum(dd$control)/20)^2  
b= (dd$treatment-sum(dd$treatment)/20)^2  
S\_1=sum(a) + sum(b)  
S\_1

## [1] 26.05345

F\_stat = (S\_0 - S\_1)/(S\_1/38)  
F\_stat

## [1] 0.2599446

and F statistics which is less than F(1,38). This shows that there is little evidence to reject .