R Assignment 3

Jacey Davies

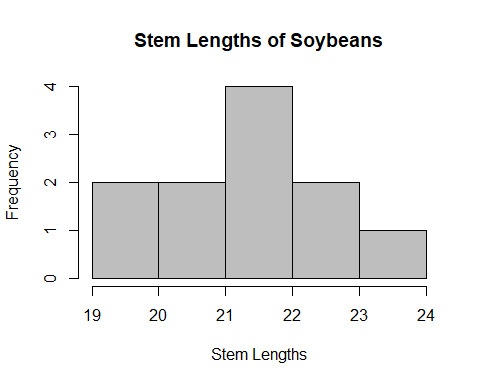
10/24/2020

The stem length of soybeans from an experiment are:

20.2, 22.9, 23.3, 20.0, 19.4, 22.0, 22.1, 22.0, 21.9, 21.5, 20.9

1. Create a histogram to visualize the data

stem = c(20.2, 22.9, 23.3, 20.0, 19.4, 22.0, 22.1, 22.0, 21.9, 21.5, 20.9)  
  
hist(stem, col="gray", main="Stem Lengths of Soybeans", xlab="Stem Lengths")



1. Test “t.test” whether the population mean is different from 23

t.test(stem, alternative = "two.sided", mu = 23, paired = FALSE, var.equal = FALSE, conf.level = 0.95)

##   
## One Sample t-test  
##   
## data: stem  
## t = -4.1467, df = 10, p-value = 0.00199  
## alternative hypothesis: true mean is not equal to 23  
## 95 percent confidence interval:  
## 20.65208 22.29338  
## sample estimates:  
## mean of x   
## 21.47273

The p-value is 0.00199, which assuming a 95% confidence level, means that there is significant evidence that the mean is different from 23.

1. Obtain a 2 sided 98% confidence interval on the true mean using “t.test”

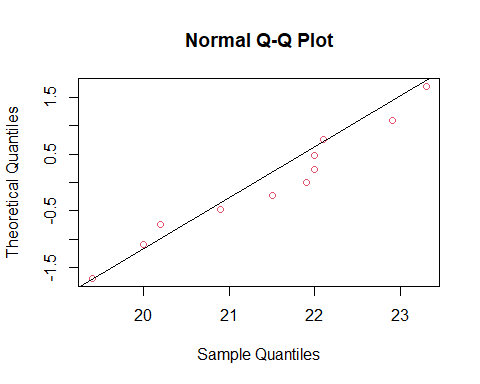
t.test(stem, alternative = "two.sided", mu = mean(stem), paired = FALSE, var.equal = FALSE, conf.level = 0.98)

##   
## One Sample t-test  
##   
## data: stem  
## t = 0, df = 10, p-value = 1  
## alternative hypothesis: true mean is not equal to 21.47273  
## 98 percent confidence interval:  
## 20.45480 22.49065  
## sample estimates:  
## mean of x   
## 21.47273

The 98% confidence interval is: (20.4548, 22.4907)

1. The researcher, by using “t.test” on a sample size of 11 was assuming that the data was normally distributed. Is that a valid claim? Create a QQ plot and interpret

qqnorm(stem,datax=TRUE,col=2)  
qqline(stem,datax=TRUE)



The points on the QQ plot do not deviate too far from the normal line except for towards the right, where the points are under the line. This implies a heavier right tail. Considering the small sample size, the data is relatively normal. Thus the researcher’s claim is valid enough for the purposes of the test.