Assign 5

Natalie Schmer

## 1.

Refer to Problem 6.42 which deals with lung capacity of rats exposed to ozone. Note: For consistency, please calculate the differences as After – Before for all questions. (12 pts)

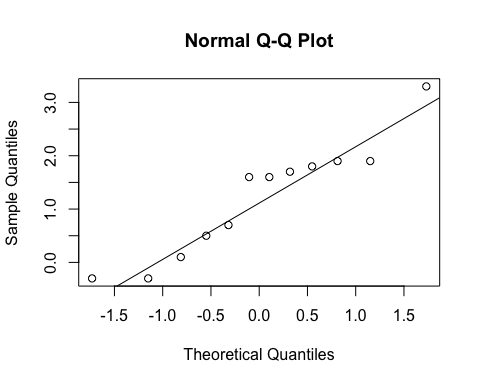
library(tidyverse)  
  
#load data and arrange data to Rats and Rats2  
Rats <- read.csv("/Users/natalieschmer/Desktop/GitHub/stats\_511/data/ASCII-comma/CH06/ex6-42.TXT", quote ="'")  
  
str(Rats)

## 'data.frame': 12 obs. of 3 variables:  
## $ Rat : int 1 2 3 4 5 6 7 8 9 10 ...  
## $ Before: num 8.7 7.9 8.3 8.4 9.2 9.1 8.2 8.1 8.9 8.2 ...  
## $ After : num 9.4 9.8 9.9 10.3 8.9 8.8 9.8 8.2 9.4 9.9 ...

### 1A.

Are the differences normally distributed? Support your answer by including a qqplot of differences in your assignment.

Rats <- Rats %>%  
 mutate(differences = After - Before)  
  
qqnorm(Rats$differences)  
qqline(Rats$differences)



**Based on the qqplot and the fact that the points generaly fall along the 1:1 line, the differences do appear normally distributed**

### 1B.

Is there sufficient evidence to support that ozone exposure increases lung capacity? Use the paired t-test with =0.05. State the hypotheses, test statistic, p-value and conclusion. (4 pts)

t.test(Rats$differences)

##   
## One Sample t-test  
##   
## data: Rats$differences  
## t = 3.885, df = 11, p-value = 0.002541  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## 0.5237735 1.8928932  
## sample estimates:  
## mean of x   
## 1.208333

**Hypotheses: Null hypothesis is that the mean difference in lung capacities is 0, and the alternative is that the mean difference in lung capacities is NOT 0. t = 3.885, p-value = 0.002541. Since p < 0.05, we reject the null hypothesis and conclude the true mean in difference of lung capacity is NOT 0.**

### 1C.

Estimate the size of the increase in lung capacity after exposure and construct a 95% t-based CI. Note: provide a standard “two-sided” CI here.

t.test(Rats$differences)

##   
## One Sample t-test  
##   
## data: Rats$differences  
## t = 3.885, df = 11, p-value = 0.002541  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## 0.5237735 1.8928932  
## sample estimates:  
## mean of x   
## 1.208333

**The size of the increase in lung capacity after exposure is about 1.21 and the 95% CI is (0.5237735, 1.8928932)**

### 1D.

Rerun the test from part B using the Wilcoxon Paired (Signed Rank) test. Give your p-value and conclusion. Use the wilcoxsign\_test() function from the coin package with distribution = “exact”. Remember to use a one-sided alternative. (4 pts)

# below code just opens the help for these two functions  
?wilcox.test  
  
library(coin)

## Loading required package: survival

?wilcoxsign\_test  
  
wilcoxsign\_test(After ~ Before, data = Rats, distribution = "exact", alternative = "greater")

##   
## Exact Wilcoxon-Pratt Signed-Rank Test  
##   
## data: y by x (pos, neg)   
## stratified by block  
## Z = 2.6692, p-value = 0.002441  
## alternative hypothesis: true mu is greater than 0

**In this case, p-value = 0.002441, so much les than 0.05, and we can still conclude that the true mean in difference of lung capacity is NOT 0, and is greater than 0.**

## 2

Designs for comparing means using College Board data on SAT scores. State whether the following is testing: 1) a single mean versus a claim in null hypothesis, 2) Means from two independent samples, or 3) Means from matched pairs. Also state the degrees of freedom associated with the t-test. (1 pt for hypotheses of test, 1 pt for df => 10 total))

### 2A.

**Hypothesis is testing a single mean versus a claim in null hypothesis, and df = 49**

### 2B.

**Hypothesis is testing means from two independant samples, and df = 50**

### 2C.

**Hypothesis is testing means from two independant samples, df= 45.**

### 2D.

**Hypothesis is testing means from matched pairs, df= 24.**

### 2E.

**Hypothesis is testing means from matched pairs, df = 19**

## 3.

Refer to problem 7.6 Vehicle Speeds Single Standard Deviation

### 3A.

Construct a 95% CI for .

s=11.35  
n=100

df <- 99  
s <- 11.35  
  
upper <- sqrt((df\*s^2)/(qchisq(.025, 99)))  
lower <- sqrt((df\*s^2)/(qchisq(.975, 99)))

### 3B.

Using =0.05, test : > 10. Give your test statistic, rejection rule and conclusion. (4pts)

x2 <- (99\*(s^2))/ 10^2  
  
pval\_3b <- 1-pchisq(x2, df= 99)

**The test statistic is 127.53, and p = 0.0282775. Since p < 0.05, we reject the null hypothesis and conclude that the true standard deviation is > 10.**

### 3C.

**The distributional assumptions include being a random sample, independent observations, normally distributed data.**

### 3D.

**Even though 10 is in our confidence interval, we still reject the null because our p value is less than 0.05**