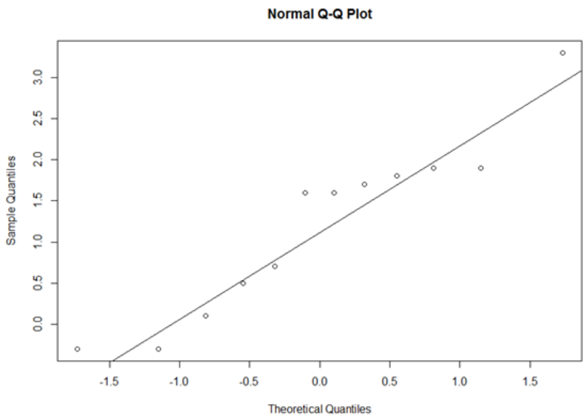
ST511 HW#5

Shaikh Shawon Arefin Shimon 830201247

1. a) Although there are some deviations present in the distribution of data, most data points fall near the normal line in QQplot. Hence, **it can be assumed that the data is normally distributed.**



b) Hypothesis:

$H_0 : \mu_D \leq 0$; where $\mu_D = \mu_{After} - \mu_{Before}$

$H_A : \mu_D > 0$

Test statistic, t = 3.885, df = 11, p-value = 0.001271

As p-value < α=0.05, null hypothesis can be rejected with 95% confidence. So, we can say with 95% confidence that ozone exposure increases lung capacity.

c) 95% confidence interval for the increase in lung capacity = (0.5237735, 1.8928932)

d) For $H_0 : \mu_D \leq 0$; where $\mu_D = \mu_{After} - \mu_{Before}$

$H_A : \mu_D > 0$

p-value = 0.002441 < 0.05

As p-value < α=0.05, null hypothesis can be rejected with 95% confidence. So, we can say with 95% confidence that ozone exposure increases lung capacity.

2) s = 11.35, df = 99 , n = 100

A) 95% CI = (9.965,13.185)

B) $H_0 : \sigma <= 10$

$H_A : \sigma > 10$; Rejection Region: $X^2 > X^2_{\alpha,n-1}$

TS: $X^2 = \frac{((n-1)s^2)}{\sigma^2}$ = 127.5343 ; $X^2_{\alpha,n-1}$ = 123.2252 ; $X^2 > X^2_{\alpha,n-1}$

So, H_0 can be rejected. We can conclude with 95% confidence that true standard deviation of the speed of vehicle is not less than 10 miles per hour.

C) For the CI and the Test to be valid, normal distributional assumption is required.

D)

skshimon@cs.colostate.edu

R Code:

```
> #Ans 1
> rats<-read.csv(file.choose())
> rats
  X.Rat. X.Before. X.After.
1      1       8.7      9.4
2      2       7.9      9.8
3      3       8.3      9.9
4      4       8.4     10.3
5      5       9.2      8.9
6      6       9.1      8.8
7      7       8.2      9.8
8      8       8.1      8.2
9      9       8.9      9.4
10    10       8.2      9.9
11    11       8.9     12.2
12    12       7.5      9.3
> Diff_After_Before <-rats$X.After.-rats$X.Before.
> Diff_After_Before
 [1]  0.7  1.9  1.6  1.9 -0.3 -0.3  1.6  0.1  0.5  1.7  3.3  1.8
> mean(Diff_After_Before)
[1] 1.208333
> sd(Diff_After_Before)
[1] 1.07742
>
>
> #A
> #Are the difference normally distributed?
> hist(Diff_After_Before)
> qqnorm(Diff_After_Before)
> qqline(Diff_After_Before)
> #NO, most data points are deviated from the straight line in the QQPlot
>
> #B
> t.test(Diff_After_Before,mu=0,alternative = "greater")

        One Sample t-test

data:  Diff_After_Before
t = 3.885, df = 11, p-value = 0.001271
alternative hypothesis: true mean is greater than 0
95 percent confidence interval:
 0.6497695       Inf
sample estimates:
mean of x
```

```
 1.208333

> #No, p value less than alpha
>
> #C Two sided CI
> t.test(Diff_After_Before,mu=0,alternative = "two.sided")

        One Sample t-test

data:  Diff_After_Before
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

>
>
> #D. Wilcoxon Paired test
> library(coin)
> wilcoxsign_test(X.After. ~ X.Before. , data = rats, distribution="exact", alternative = "gre
ater")

        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

> #P value less than alpha

>
```

```
> sigma0 <- 10
> df = 99
> s = 11.35
> Chi_Square <- df*s^2/sigma0^2
> qchisq(0.975,df=99,lower.tail = FALSE)
[1] 73.36108
> qchisq(0.025,df=99,lower.tail = FALSE)
[1] 128.422
```

```
>
> Lower <- sqrt((df*s^2)/(qchisq(0.025,df=99,lower.tail = FALSE)))
> Upper <- sqrt((df*s^2)/(qchisq(0.975,df=99, lower.tail = FALSE)))
> Lower
[1] 9.965378
> Upper
[1] 13.18501

>
```