STA305 Assignment 2

Linxia Li (1005715488)

Part I

1. Set the seed of your randomization to be your student number

```
set.seed(1005715488) # Set the seed
```

2. Randomly generate 9 observations from the $Exp(\lambda = 1/3)$ distribution to correspond to treatment S. List the observed values, to 3 decimal places, and the order in which they appeared. (Hint: A random sample from an exponential distribution can be generated in R using the function rexp(n =, rate=))

```
exp_S <- rexp(9, rate = 1/3) # Randomly generate 9 observations
treatment_S <- round(exp_S, 3) # Round the observed values to 3 decimal places
treatment_S
```

```
## [1] 17.630  4.560  3.551  0.088 12.400  7.910  1.889  8.379  3.615
```

3. Randomly generate 9 observations from the $Exp(\lambda = 1)$ distribution to correspond to treatment T. List the observed values, to 3 decimal places, and the order in which they appeared.

```
exp_T <- rexp(9, rate = 1) # Randomly generate 9 observations
treatment_T <- round(exp_T, 3) # Round the observed values to 3 decimal places
treatment_T
```

```
## [1] 0.506 0.110 0.565 4.121 0.258 0.166 0.997 2.310 0.514
```

4. Use the order of the observations in 2) and 3) to from pairs of observations. Display the pairs of observations of treatment S and T for the randomized paired design.

Here is the plot for the pairs of observations of treatment S and T for the randomized paired design

```
plot(c(1,2,3,4,5,6,7,8,9), treatment_S, pch = 16, col = "blue", cex = 1.5, xlab = "observations",
ylab = "lifetime")
points(c(1,2,3,4,5,6,7,8,9), treatment_T, pch = 17, col = "purple", cex = 1.5)
mtext("Observed lifetime of LED bulbs under treament S and T")
legend("topright", legend = c("Treatment S", "Treatment T"), pch = c(16, 17), col =
c("blue", "purple"), bty = "n", horiz = TRUE)
```

Observed lifetime of LED bulbs under treament S and T
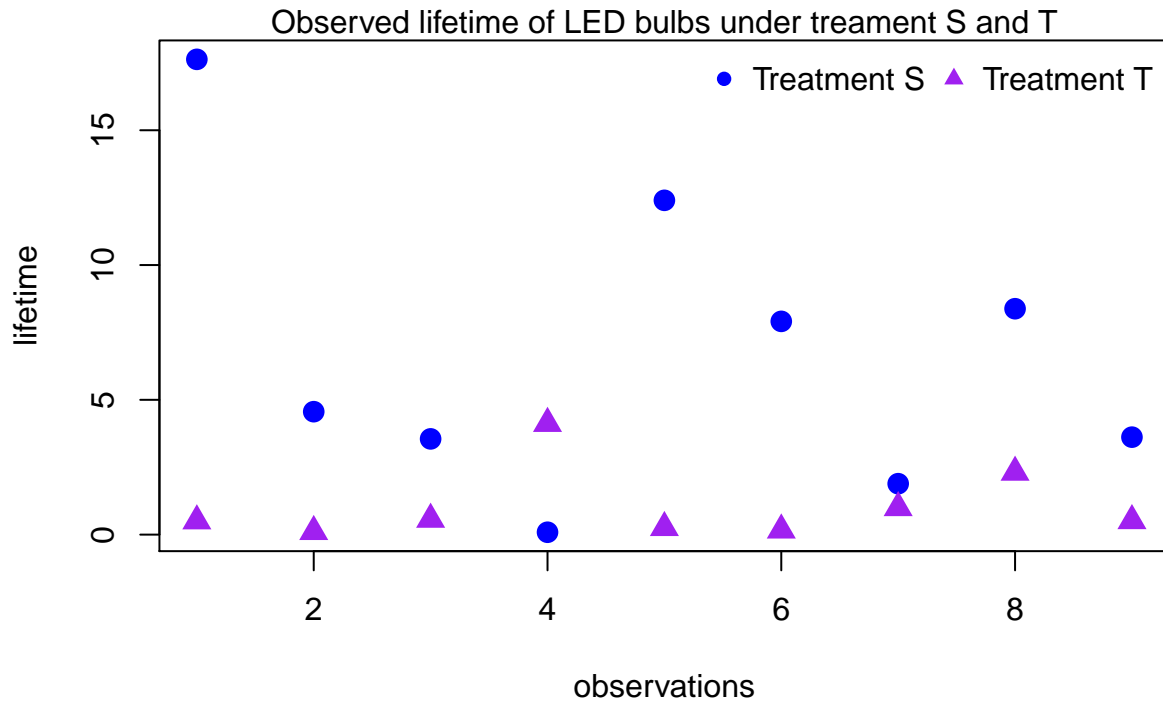
Here is the table for the pairs of observations of treatment S and T for the randomized paired design

```
diff <- treatment_S - treatment_T
meandiff <- mean(diff)
randomized_paired_comparison_table <- data.frame(treatment_S, treatment_T, diff)
randomized_paired_comparison_table
```

```
##   treatment_S treatment_T    diff
## 1      17.630       0.506  17.124
## 2       4.560       0.110   4.450
## 3       3.551       0.565   2.986
## 4       0.088       4.121  -4.033
## 5      12.400       0.258  12.142
## 6       7.910       0.166   7.744
## 7       1.889       0.997   0.892
## 8       8.379       2.310   6.069
## 9       3.615       0.514   3.101
```

Part II

Completely Randomized Design

i. Describe the randomization distribution for this comparison by stating the number of values that this distribution contains and the probability of the observed treatment allocation?

The completely randomized design contains the number of values of $\binom{18}{9}$ which is 48620 and the probability of the observed treatment allocation is $\frac{1}{\binom{18}{9}}$ which is $\frac{1}{48620}$.

ii. Create a histogram of this randomization distribution; include vertical line(s) to mark the area(s) corresponding to the p-value.

```r
fert <- c(treatment_S, treatment_T)
N <- choose(18, 9)
res <- numeric(N)
index <- combn(1:18, 9)
for (i in 1:N)
{res[i] <- mean(fert[index[,i]]) - mean(fert[-index[,i]])}

# Plotting the histogram
hist(res, xlab = expression(bar(y)[S] - bar(y)[T]), main = "Randomized Distribution of Difference
in Means by two different treatments")
observed <- mean(treatment_S) - mean(treatment_T)
abline(v = observed, col = "blue")
abline(v = -observed, col = "blue")
```
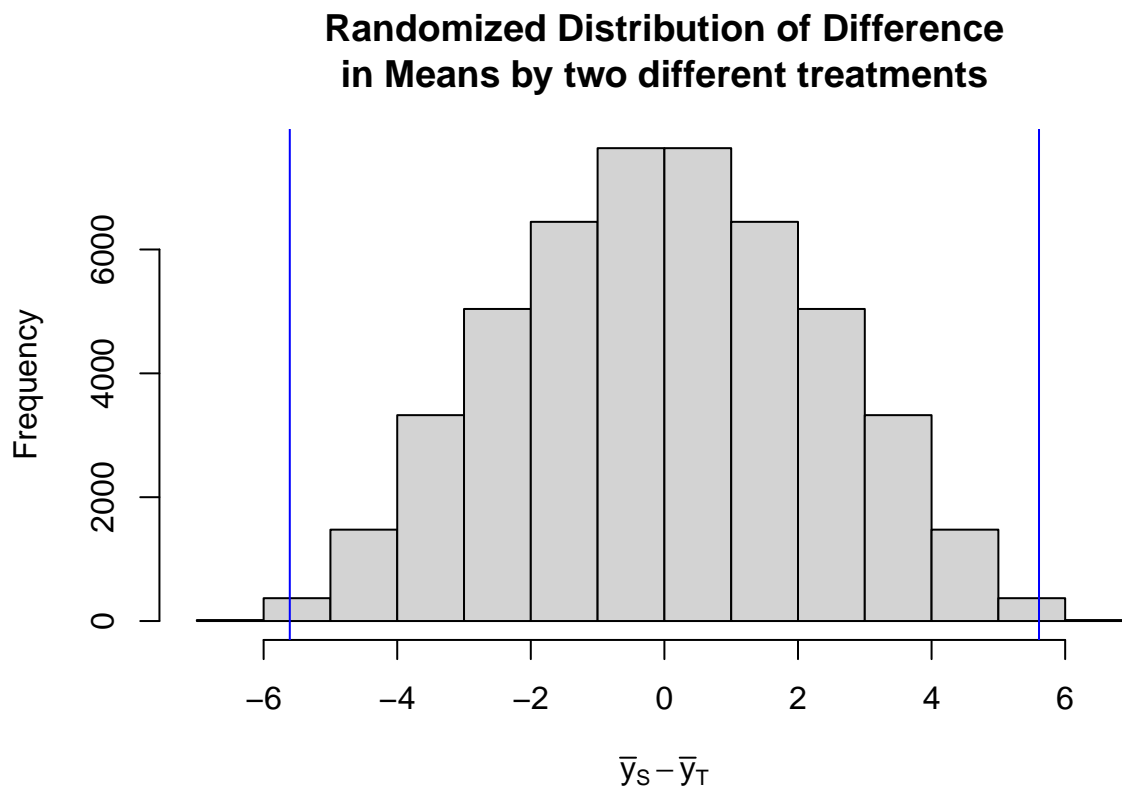

Randomized Distribution of Difference in Means by two different treatments

The right blue vertical line represents the test statistic of this randomization test which is the value of difference between the mean of lifetime of LED bulbs under treatment S and the mean of lifetime of LED bulbs under treatment T and the left blue vertical line represents the opposite value of the test statistic of this randomization test.

The p-value is as follows:

```r
tbar <- mean(res)
pval <- sum(abs(res - tbar) >= abs(observed - tbar))/N
round(pval, 4)
```

[1] 0.0032

iii. Use the randomization test to determine if there is evidence of a difference in means between the two treatments. Explain your answer, including the P-value of your test.

Null Hypothesis H_0: There is no difference in means of LED bulbs lifetime between treatment S and treatment T, i.e., $\mu_S = \mu_T$.

Alternative Hypothesis H_1: There is a difference in means of LED bulbs lifetime between treatment S and treatment T, i.e. $\mu_S \neq \mu_T$.

Under the null hypothesis, the p-value calculated above is 0.0032 which is smaller than the significance level of 5%. Therefore, we can reject the null hypothesis and conclude that there is evidence of a difference in means of LED bulbs lifetime between the two treatments which are treatment S and T.

Randomized Paired Design

i. Describe the randomization distribution for this comparison by stating the number of values that this distribution contains and the probability of the observed treatment allocation?

The randomized paired design contains the number of values of 2^9 which is 512 and the probability of the observed treatment allocation is $\frac{1}{2^9}$ which is $\frac{1}{512}$.

ii. Create a histogram of this randomization distribution; include vertical line(s) to mark the area(s) corresponding to the p-value.

```
diff <- treatment_S - treatment_T
meandiff <- mean(diff)

N <- 2^(9) # number of treatment assignments
res <- numeric(N) # vector to store results
LR <- list(c(-1,1)) # difference is multipled by -1 or 1
# Generate all possible treatment assignments as a 1024x9 array
trtassign <- expand.grid(rep(LR, 9))
for (i in 1:N)
{res[i] <- mean(as.numeric(trtassign[i,])*diff)}

# Plotting the histogram
hist(res, xlab = "Mean Difference", main = "Randomized Paired Distribution of Mean
Difference by two different treatments")
abline(v = meandiff, col = "blue")
abline(v = -meandiff, col = "blue")
```
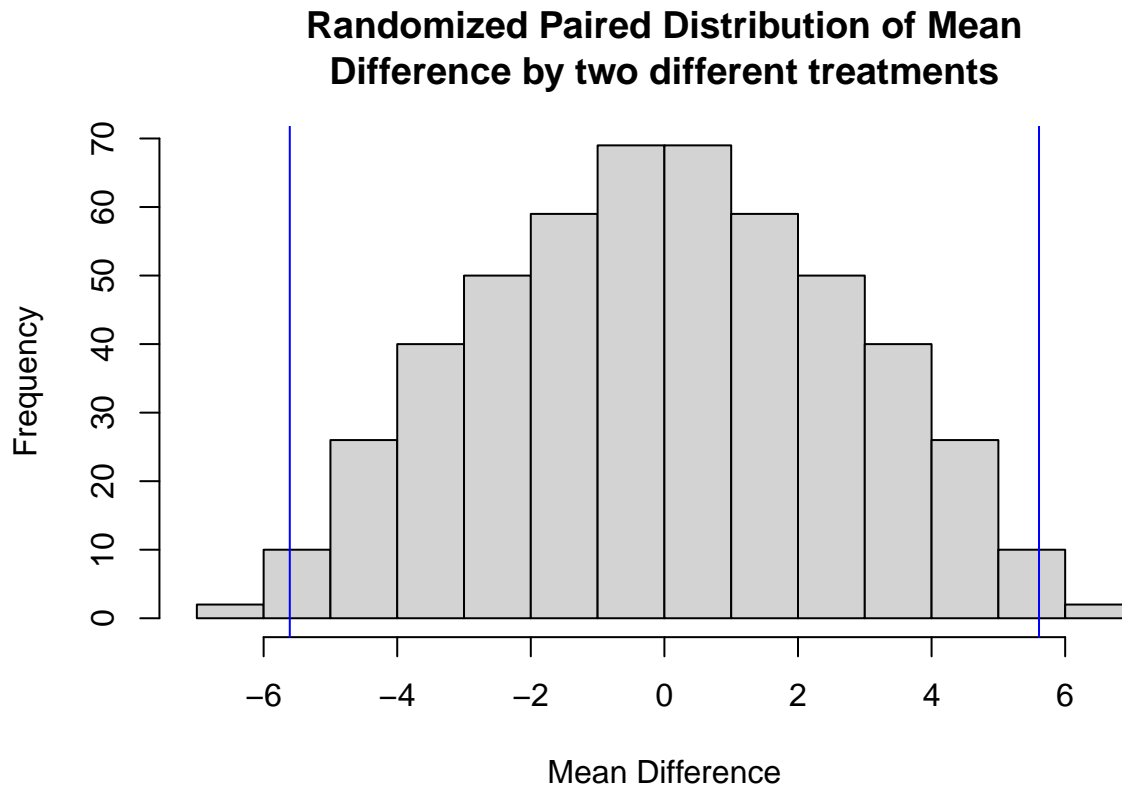
Randomized Paired Distribution of Mean Difference by two different treatments

The right blue vertical line represents the test statistic of this randomization test which is the value of mean difference of lifetime of LED bulbs under treatment S and treatment T and the left blue vertical line represents the opposite value of the test statistic of this randomization test.

The p-value is as follows:

```
tbar <- mean(res)
pval <- sum(abs(res - tbar) >= abs(meandiff - tbar))/N
round(pval, 3)
```

```
## [1] 0.027
```

iii. Use the randomization test to determine if there is evidence of a difference in means between the two treatments. Explain your answer, including the P-value of your test.

Null Hypothesis H_0: There is no mean difference of LED bulbs lifetime between treatment S and treatment T, i.e., $\bar{d} = 0$.

Alternative Hypothesis H_1: There is a mean difference of LED bulbs lifetime between treatment S and treatment T, i.e. $\bar{d} \neq 0$.

Under the null hypothesis, the p-value calculated above is 0.027 which is smaller than the significance level of 5%. Therefore, we can reject the null hypothesis and conclude that there is evidence of a mean difference of LED bulbs lifetime between the two treatments which are treatment S and T.

Part III

For both designs, based on the data simulated in part I, conduct a t-test to compare the means of the two treatments. Note: Assume that the population distributions and parameters are unknown.

Two sample t-test for the completely randomized design

i. Explain your answer, including the P-value of your test.

```
t.test(treatment_S, treatment_T, var.equal = FALSE, alternative = "two.sided")
```

```
##
##  Welch Two Sample t-test
##
## data:  treatment_S and treatment_T
## t = 2.9443, df = 8.912, p-value = 0.01654
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.292857 9.923809
## sample estimates:
## mean of x mean of y
##  6.669111  1.060778
```

The above is a two sample t-test for the completely randomized design. The null hypothesis is that there is no difference in means of LED bulbs lifetime between treatment S and treatment T, i.e., $\mu_S = \mu_T$ and the alternative hypothesis is that there is a difference in means of LED bulbs lifetime between treatment S and treatment T, i.e., $\mu_S \neq \mu_T$. From the above two sample t-test, we can find out that the two-sided p-value is approximately 0.017 which is less than the significance level 5%. Therefore, we can reject the null hypothesis and conclude that there is a difference in means of LED bulbs lifetime between treatment S and treatment T.
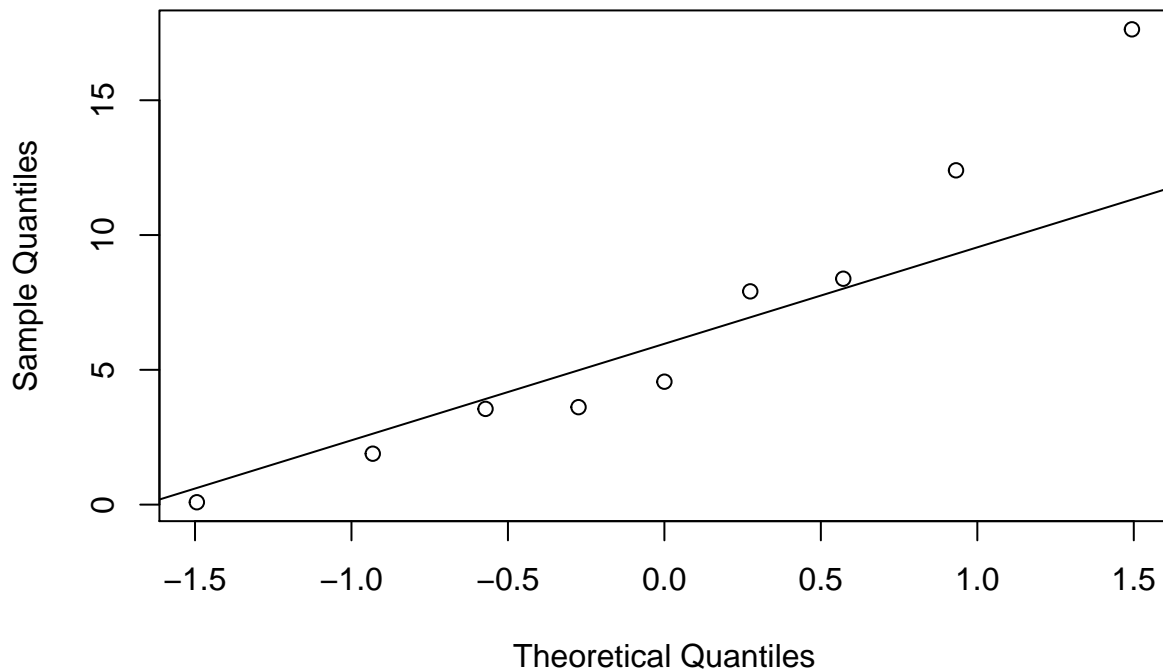
ii. Are the assumptions behind the t-test satisfied?

The assumption of normality can be checked by using normal quantile plots (QQ-plot).

QQ-plot of treatment S

```
qqnorm(treatment_S, main = " Normal QQ-plot of treatment S") # Normal QQ-plot of treatment S
qqline(treatment_S) # Fit a straight line
```
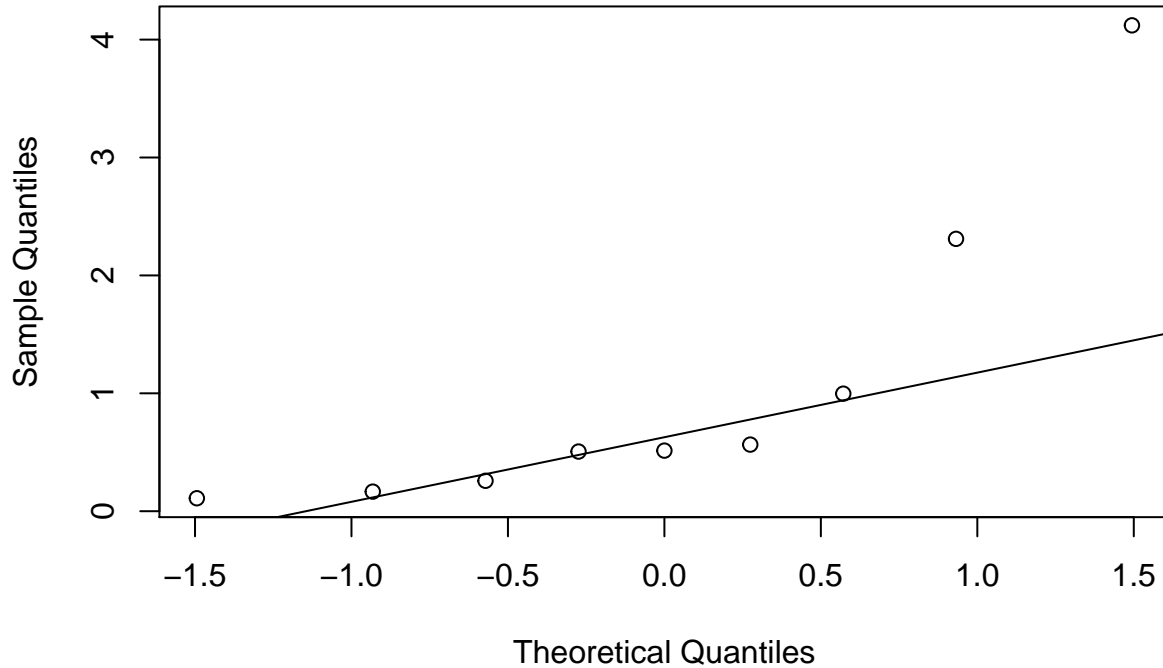
Normal QQ–plot of treatment S

The QQ-plot shows if the data is normally distributed. We have known that if the points are lied well on the straight line, then the observations are normally distributed. From the above normal QQ-plot for treatment S, we can find out that there are some points which are not on the straight line and they are quite far from the straight line so that the normality assumption is not satisfied. Additionally, the observations within the sample for treatment S are independent so that the independence assumption is satisfied. Therefore, the normality assumption behind the t-test is not satisfied and the independence assumption behind the t-test is satisfied.

QQ-plot of treatment T

```
qqnorm(treatment_T, main = "Normal QQ-plot of treatment T") # Normal QQ-plot of treatment T
qqline(treatment_T) # Fit a straight line
```

Normal QQ–plot of treatment T

The QQ-plot shows if the data is normally distributed. We have known that if the points are lied well on the straight line, then the observations are normally distributed. From the above normal QQ-plot for treatment T, we can find out that there are some points which are not on the straight line and they are quite far from the straight line so that the normality assumption is not satisfied. Additionally, the observations within the sample for treatment T are independent so that the independence assumption is satisfied. Therefore, the normality assumption behind the t-test is not satisfied and the independence assumption behind the t-test is satisfied.

iii. Do the results of the t-test agree with the results of the randomization test? Explain

There are some similarities in the two sample t-test and the randomization test. Firstly, the null hypothesis and the alternative hypothesis are the same. The null hypothesis in both tests is that there is no difference in means of LED bulbs lifetime between treatment S and treatment T, i.e., $\mu_S = \mu_T$ and the alternative hypothesis is that there is a difference in means of LED bulbs lifetime between treatment S and treatment T, i.e. $\mu_S \neq \mu_T$. Secondly, in the completely randomized design, the p-value of the randomization test is 0.0032 which is less than the significance level 5% so that we can reject the null hypothesis and conclude that there is evidence of a difference in means of LED bulbs lifetime between the two treatments. In the two sample t-test, the p-value of the t-test is approximately 0.0095 which is also less than the significance level 5% so that we can also reject the null hypothesis and conclude that there is evidence of a difference in means of LED bulbs lifetime between the two treatments. Therefore, the results of the t-test agree with the results of the randomization test.

Paired t-test

i. Explain your answer, including the P-value of your test.

```
t.test(treatment_S, treatment_T, paired = TRUE, alternative = "two.sided")
```

```
##
##  Paired t-test
##
## data:  treatment_S and treatment_T
## t = 2.7038, df = 8, p-value = 0.02691
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##    0.8251357 10.3915309
## sample estimates:
## mean of the differences
##                5.608333
```

The above is a paired t-test for the randomized paired design. The null hypothesis is that there is no difference in means of LED bulbs lifetime between treatment S and treatment T, i.e., $\bar{d} = 0$ and the alternative hypothesis is that there is a difference in means of LED bulbs lifetime between treatment S and treatment T, i.e., $\bar{d} \neq 0$. From the above paired t-test, we can find out that the two-sided p-value is approximately 0.027 which is less than the significance level 5%. Therefore, we can reject the null hypothesis and conclude that there is evidence of a mean difference of LED bulbs lifetime between treatment S and treatment T.
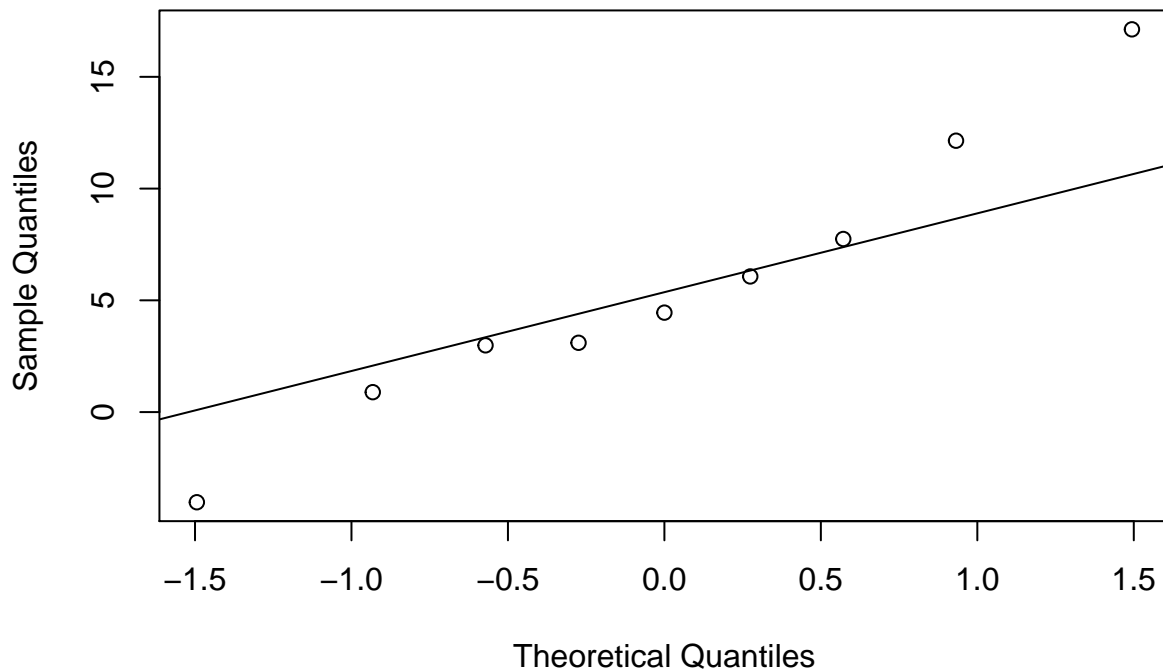
ii. Are the assumptions behind the t-test satisfied?

The assumption of normality can be checked by using normal quantile plots (QQ-plot).

QQ-plot of the difference between treatment S and T

```
qqnorm(diff, main = "Normal QQ-plot of the difference between treatment S and T")
qqline(diff)
```

Normal QQ–plot of the difference between treatment S and T

The QQ-plot shows if the data is normally distributed. We have known that if the points are lied well on the straight line, then the observations are normally distributed. From the above normal QQ-plot for the difference between treatment S and T, we can find out that there are some points which are not on the straight line and they are quite far from the straight line so that the normality assumption is not satisfied. Additionally, the two samples which are two groups for treatment S and T are independent so that the independence assumption is satisfied. Therefore, the normality assumption behind the t-test is not satisfied and the independence assumption behind the t-test is satisfied.

iii. Do the results of the t-test agree with the results of the randomization test? Explain

There are some similarities in the paired t-test and the randomized paired test. Firstly, the null hypothesis and the alternative hypothesis are the same. The null hypothesis in both tests is that there is no mean difference of LED bulbs lifetime between treatment S and treatment T, i.e., $\bar{d} = 0$ and the alternative hypothesis is that there is a mean difference of LED bulbs lifetime between treatment S and treatment T, i.e. $\bar{d} \neq 0$. Secondly, in the randomized paired design, the p-value of the randomization test is 0.027 which is less than the significance level 5% so that we can reject the null hypothesis and conclude that there is evidence of a mean difference of LED bulbs lifetime between the two treatments. In the paired t-test, the p-value of the t-test is approximately 0.027 which is also less than the significance level 5% so that we can also reject the null hypothesis and conclude that there is evidence of a mean difference of LED bulbs lifetime between the two treatments. Therefore, the results of the t-test agree with the results of the randomization test.

Part IV

```
set.seed(1005715488) # Set the seed
```

i. Completely randomized design and t-test

```
pvals <- replicate(1000, t.test(rexp(20, rate = 1/3), rexp(20, rate = 1), var.equal = FALSE,
alternative = "two.sided")$p.value)
sum(pvals <= 0.05)/1000 # power of two sample t-test
```

```
## [1] 0.85
```

The power of the two sample t-test for the completely randomized design is 0.85.

ii. Randomized paired design and t-test

```
pvals <- replicate(1000, t.test(rexp(20, rate = 1/3), rexp(20, rate = 1), paired = TRUE,
alternative = "two.sided")$p.value)
sum(pvals <= 0.05)/1000 # power of paired t-test
```

```
## [1] 0.873
```

The power of the paired t-test for the randomized paired design is 0.873.

iii. Completely randomized design and Wilcoxon test

```
pvals <- replicate(1000, wilcox.test(rexp(20, rate = 1/3), rexp(20, rate = 1), var.equal = FALSE,
alternative = "two.sided")$p.value)
sum(pvals <= 0.05)/1000
```

```
## [1] 0.829
```

The power of the Wilcoxon test for the completely randomized design is 0.829.

iiii. Randomized paired design and Wilcoxon test

```
pvals <- replicate(1000, wilcox.test(rexp(20, rate = 1/3), rexp(20, rate = 1), paired = TRUE,
alternative = "two.sided")$p.value)
sum(pvals <= 0.05)/1000
```

```
## [1] 0.829
```

The power of the Wilcoxon test for the randomized paired design is 0.829.

We have known that the power is the proportion of times that the test correctly rejects the null hypothesis in repeated sampling and it indicates the probability of avoiding a type II error. Additionally, a high statistical power means that the test results are likely valid. As the power increases, the probability of making a Type II error decreases and a low statistical power means that the test results are questionable. From the above tests for two different designs, we can find out that the power of the two sample t-test for the completely randomized design is 0.85 and the power of the Wilcoxon test for the completely randomized design is 0.829 so that the power of the two sample t-test for the completely randomized design is larger than that of the Wilcoxon test for the completely randomized design. In addition, the power of the paired t-test for the randomized paired design is 0.873 and the power of the Wilcoxon test for the randomized paired design is 0.829 so that the power of the paired t-test for the randomized paired design is larger than that of the Wilcoxon test for the randomized paired design. Therfore, I intend to recommend t-test since the power of the t-test is larger than that of the Wilcoxon test. However, I would like to recommend the Wilcoxon test for each of the two experimental designs since it is a non-parametric test. We have known that t-test assumes the normality while the data does not follow a normal distribution so that we should not choose t-test. Therefore, I would like to recommend the Wilcoxon test.