STA305 Assignment 2

Qian Wang

2/14/2021

Part I

Suppose that I am an engineer who wants to compare the average lifetime of LED bulbs under two different temperature at $S°C$ and $T°C$. Suppose that the bulb's lifetime follows an exponential distribution with rate parameter λ, $\lambda>0$. The density function of this distribution is $f(x) = \lambda exp(-\lambda x), x \geq 0$. The mean and standard deviation of this exponential distribution is $1/\lambda$.

Hypothesis tests:

H_0(null hypothesis): Temperature does not affect the lifetime of LED bulbs. i.e. $\bar{y}_S = \bar{y}_T$

H_1(alternative hypothesis): Temperature does affect the lifetime of LED bulbs. i.e. $\bar{y}_S \neq \bar{y}_T$

I hypothesize that the expected lifetime of LED bulbs is 3 years under $S°C$ and 1 year under $T°C$. Then randomly generate 18 observations under two experimental designs which are completely randomized design and randomized paired design to compare the average lifetimes between two different temperatures $S°C$ and $T°C$.

```
#Set the seed of my randomization to be my student number
set.seed(1003946983)

#Randomly generate 9 observations from the Exp(rate=1/3) distribution to correspond to treatment S.
S=rexp(n=9, rate = 1/3)
#Listed the observed values with 3 decimal places.
S=round(S,3)

#Randomly generate 9 observations from the Exp(rate=1) distribution to correspond to treatment T.
T=rexp(n=9, rate = 1)
#Listed the observed values with 3 decimal places.
T=round(T,3)

#Use the order of the observations above to form pairs of observations.
bulbs <- c(S,T)
#Display the pairs of observations of treatment S and T for the randomized paired design
difference=S-T
diffdata=data.frame(S,T,difference)
diffdata
```

```
##        S     T difference
## 1  3.812 1.026      2.786
## 2  3.065 0.208      2.857
## 3  1.065 0.053      1.012
## 4  0.696 2.655     -1.959
## 5  3.137 1.625      1.512
## 6  4.034 0.609      3.425
```

```
## 7 13.316 0.309    13.007
## 8  2.111 0.448     1.663
## 9  4.652 0.006     4.646
```

Part II

For both designs which are completely randomized design and randomized paired design, conduct a randomization test to compare the means of the two treatments S and T.

For the completely randomized design:

i

```
N <- choose(18,9)
N
```

```
## [1] 48620
```

The number of values that this distribution contains is 48620.

```
possibility = format(1/N, scientific = FALSE)
possibility
```

```
## [1] "0.00002056767"
```

The possibility of the observed treatment allocation is 0.00002056767

ii

```
#The hypothetical value of the difference if the null hypothesis is true.
observed=mean(S)-mean(T)
observed
```

```
## [1] 3.216556
```

```
#Create vector to store results
res <- numeric(N)
#Create all 48620 arrangement to treatment S and the remainder will be the treatment T
index<-combn(1:18,9)
#Update the variable res
for(i in 1:N)
{
  res[i]=mean(bulbs[index[,i]])-mean(bulbs[-index[,i]])
}
tbar=mean(res)
tbar
```

```
## [1] -1.826776e-20
```

```
#The p-value under the null hypothesis of obtaining a more extreme result than the observed result
pval1=sum(abs(res-tbar)>=abs(observed-tbar))/N
pval1=signif(pval1,2)
#Create the histogram of this randomization distribution
```

```
histogram <- hist(res,breaks = 50,plot=F)
colors = rep("blue", length(histogram$breaks))
colors[histogram$breaks <= observed] = "white"
colors[histogram$breaks <= -observed] = "blue"
hist(res,breaks=50,col=colors,
     xlab = "Mean in Difference", ylab = "Frequency",
     main="Two Sample Randomization on Distribution of Mean in DIfference")
abline(v=observed,col="blue")
abline(v=-observed,col="blue")
```

Two Sample Randomization on Distribution of Mean in DIfference

iii
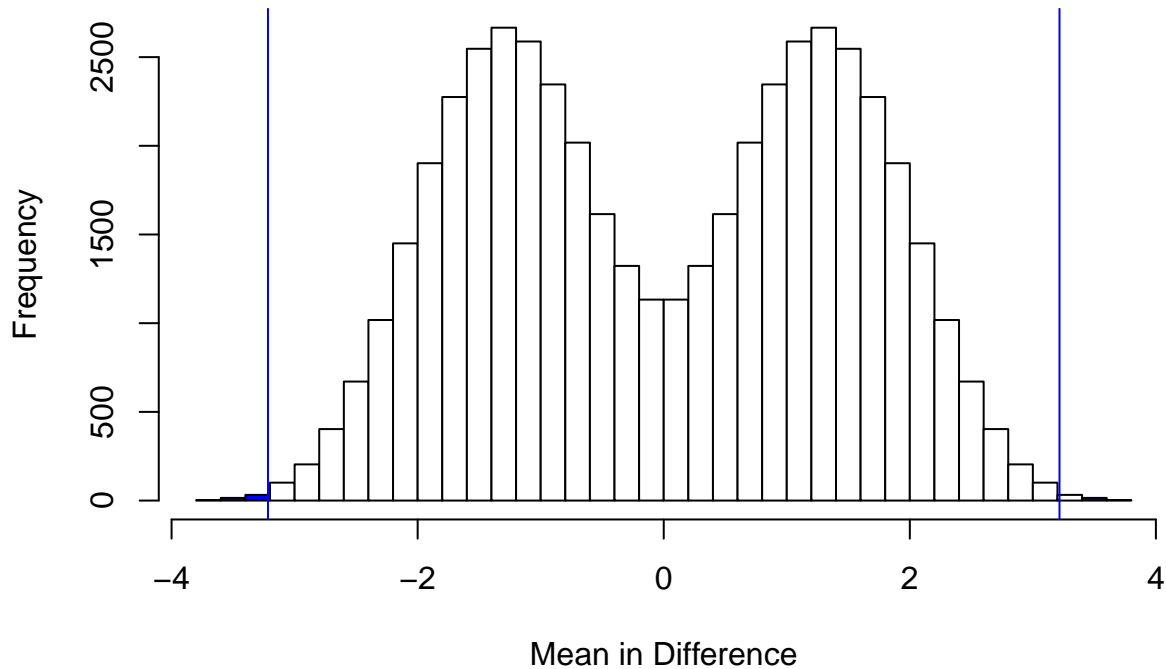
The blue area on the above histogram is the p-value. Using the complete randomization, the p-value is 0.0019 which is smaller than 5%, it indicates that we have evidence against the null hypothesis, the evidence of a difference in means between the two treatments is significant.

For the randomized paired design:

i

```
N=2^9
N
```

```
## [1] 512
```

The number of values that this distribution contains is 512

```r
possibility = format(1/N, scientific = FALSE)
possibility
```

```
## [1] "0.001953125"
```

The possibility of the observed treatment allocation is 0.001953125

ii

```r
#Looking at the differences
diff=S-T
#The hypothetical value of the average difference in the experiment that could have been observed if th
meandiff <- mean(diff)
meandiff
```

```
## [1] 3.216556
```

```r
#Create vector to store results
res=numeric(N)
#Get standard deviation of these differences as one sample.
blubs.data <- data.frame(S,T,diff)
#Difference is multiplied by -1 or 1
LR=list(c(-1,1))
#Generate all possible treatment assign
trtassign=expand.grid((rep(LR,9)))
#Update the variable res
for (i in 1:N)
{
  res[i]=mean(as.numeric(trtassign[i,])*diff)
}
tbar=mean(res)
tbar
```

```
## [1] 0
```

```r
#The p-value under the null hypothesis of obtaining a more extreme result than the observed result
pval2=sum(abs(res-tbar)>=abs(meandiff-tbar))/N
pval2=signif(pval2,2)
#Create the histogram of this randomization distribution
histogram <- hist(res,breaks = 50,plot=F)
colors = rep("blue", length(histogram$breaks))
colors[histogram$breaks >= -meandiff] = "white"
colors[histogram$breaks >= meandiff] = "blue"
hist(res,breaks=50,col=colors,xlab = "Mean Difference", ylab = "Frequency", main="Randomization on Dist
abline(v=meandiff,col="blue")
abline(v=-meandiff,col="blue")
```

Randomization on Distribution LED Bulbs

Mean Difference

iii
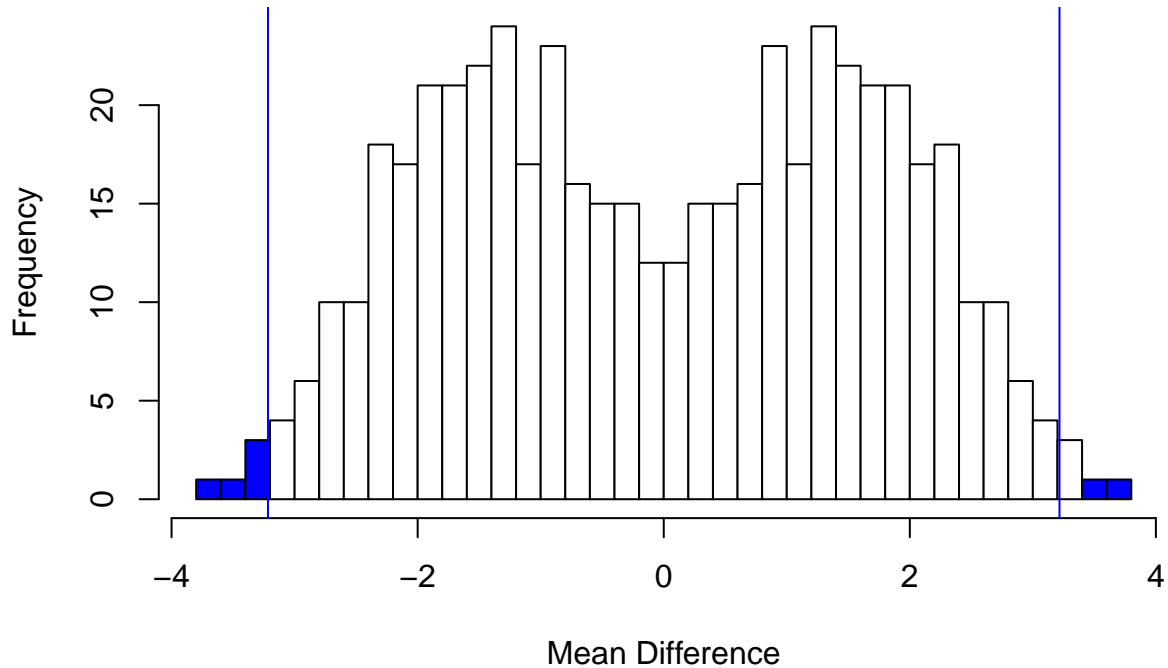
The blue area on the above histogram is the p-value. Using the paired randomization, the p-value is 0.02 which is smaller than 5%, it indicates that we don't have evidence against the null hypothesis, the evidence of a difference in means between the two treatments is significant.

Part III

For both designs and the data I simulated in part I, I conduct a t-test for each of the design to compare the means of the two treatments. Assume the population distribution and parameters are unknown.

For two-sample t test

For the two-sample t-test compare to the two-sample randomization test, we assume the two samples from S and T are independent random samples from a normal distribution with means μ_S and μ_T and the same variance.

i

```
statistics = t.test(S,T,var.equal=FALSE,alternative=)
statistics

##
##  Welch Two Sample t-test
##
```
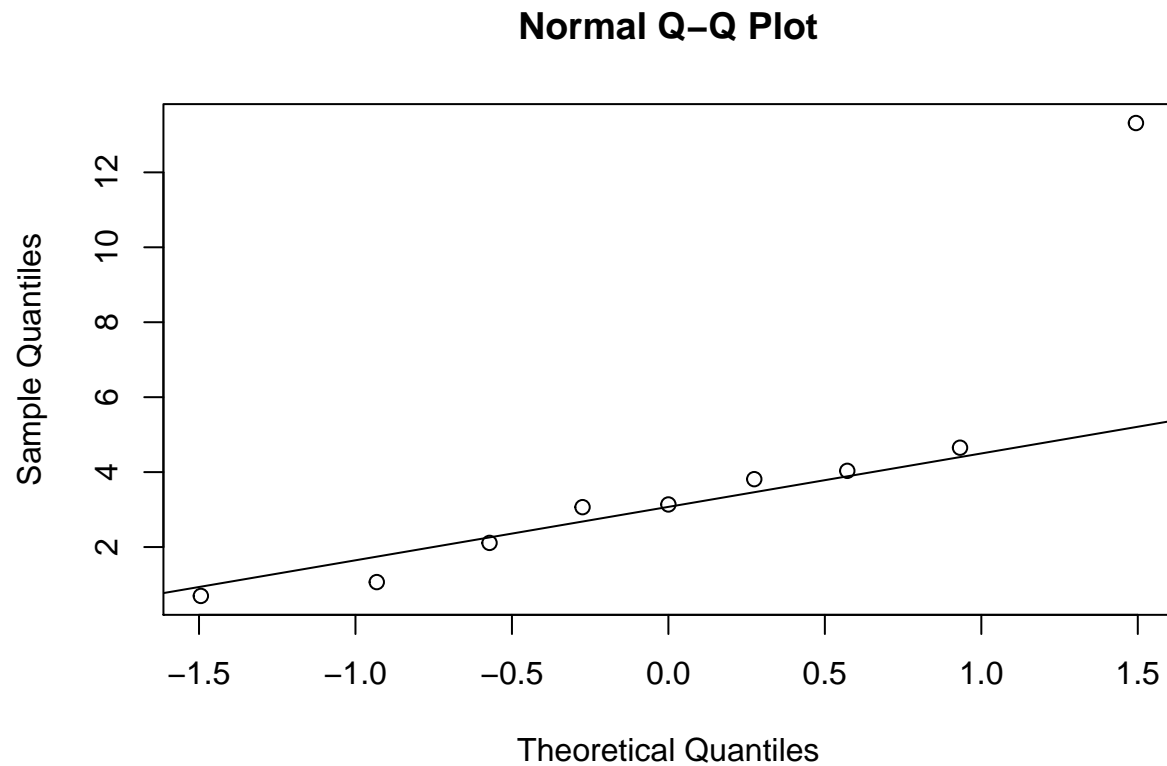
```
## data:  S and T
## t = 2.5118, df = 8.8699, p-value = 0.03357
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   0.3131944 6.1199167
## sample estimates:
## mean of x mean of y
##   3.987556  0.771000
```

```
pval3<-statistics$p.value
pval3=signif(pval3,2)
pval3
```
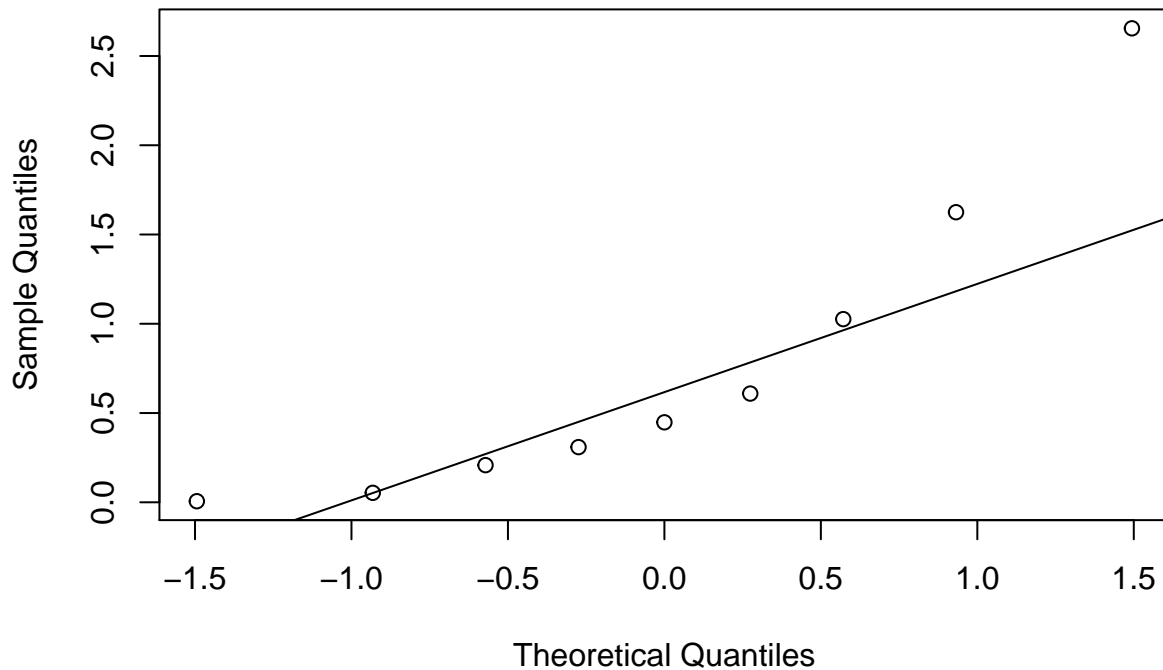
```
## [1] 0.034
```

ii

```
qqnorm(S)
qqline(S)
```

Normal Q–Q Plot


```
qqnorm(T)
qqline(T)
```

Normal Q−Q Plot

We see both samples S and T, the normality satisfies, since most points are on or close to the line. We don't see a dramatic difference from the line.

iii

The results of the two-sample t-test agree with the results of the complete randomization test. We see that the p-value of the complete randomization test and two-sample t-tests are 0.0019 and 0.034 respectively. They are both less than 0.05, but there are some differences in the assumption we make in the two-sample t-test. We assume the samples satisfy normality and independence.

For paired t test

i

For the paired t-test compare to the paired randomization test, we assume that the average differences are a random sample follows a normal distribution.

```
#The first sample comes from S, and the other sample comes from T. We are having differences from these
statistics <- t.test(blubs.data$S,blubs.data$T,paired=TRUE)
statistics
```

```
##
##  Paired t-test
##
## data:  blubs.data$S and blubs.data$T
## t = 2.3471, df = 8, p-value = 0.04689
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
```

```
##  0.05636095 6.37675016
## sample estimates:
## mean of the differences
##                 3.216556
```

```
#The p-value of the paired t-test
pval4 <- statistics$p.value
pval4=signif(pval4,2)
```

ii

```
qqnorm(S-T)
qqline(S-T)
```

Normal Q–Q Plot
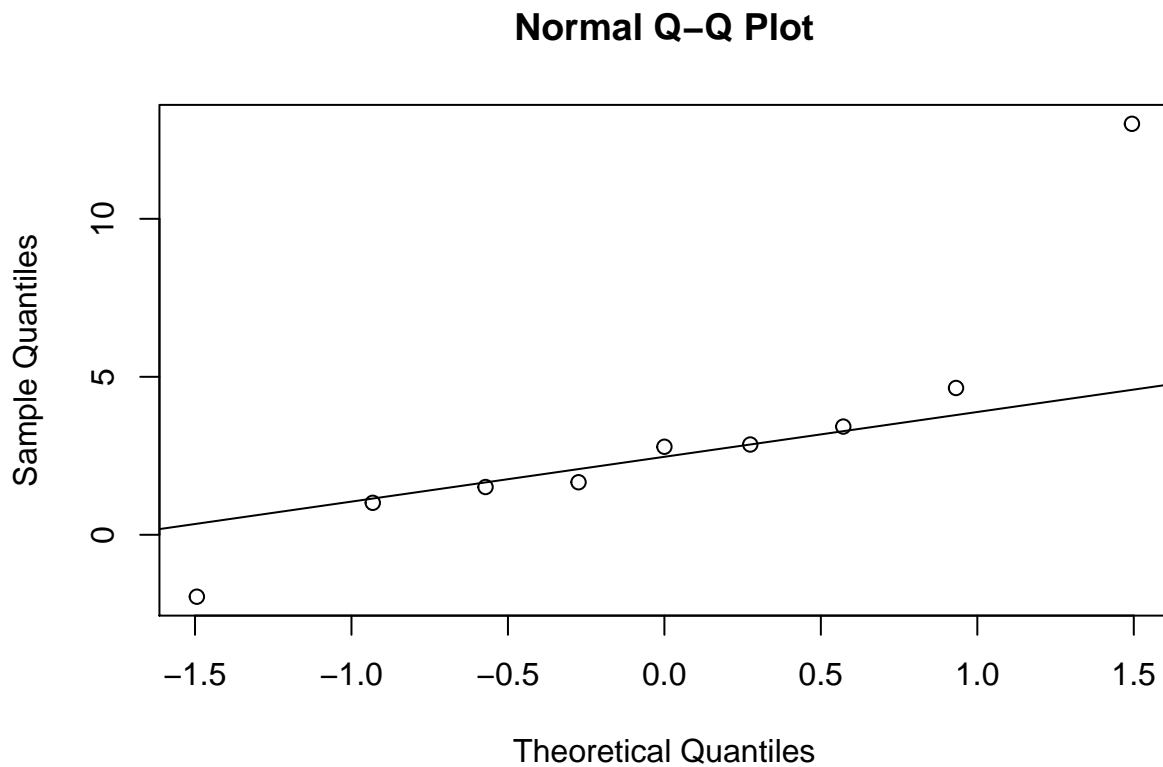
We see that majority of points close to the line, there is no great deviation from the line. Thus the assumption of normality is satisfied.

iii

The results of the t-test agree with the results of the paired randomization test since the p-value of the paired t-test is 0.047 which is similar to the randomized randomization test which has the p-value of 0.02. But there are differences between the two. More assumptions have to incorporate before we run the paired t-test, we have to assume the normality and observations are paired.

Part IV

I realize that I should use a non-parametric method called the Mann-Whitney Test since data does not follow a normal distribution. Then I can compare the average lifetimes under two different temperatures of the LED bulbs for both randomization designs. There are 20 bulbs are tested under each temperature of the $S°C$ and $T°C$. I am going to find and compare the power of the test to detect a difference in means at 5% significance level for the following four different cases.

Case i. Completely randomized design and t-test

```r
#Using my student number to set the seed of my randomization
set.seed(1003946983)
#Make 1000 replications
pvals <- replicate(1000,t.test(rexp(n=20, rate = 1/3),
                                rexp(n=20, rate = 1),
                               var.equal = FALSE
                               )$p.value)
#Calculate power
power1=sum(pvals<=0.05)/1000
```

The power of the two-sample t-test is approximately 0.872 for the completely randomized design.

Case ii. Randomized paired design and t-test

```r
set.seed(1003946983)
#Make 1000 replications
pvals <- replicate(1000,t.test(rexp(n=20, rate = 1/3),
                                rexp(n=20, rate = 1),
                               paired = TRUE)$p.value)
#Calculate power
power2=sum(pvals<=0.05)/1000
```

The power of the two-sample t-test is approximately 0.86 for the randomized paired design.

Case iii. Completely randomized design and Wilcoxon test

```r
set.seed(1003946983)
#Make 1000 replications
pvals <- replicate(1000,wilcox.test(rexp(n=20, rate = 1/3),
                                rexp(n=20, rate = 1),
                               var.equal = FALSE)$p.value)
#Calculate power
power3=sum(pvals<=0.05)/1000
```

The power of the Wilcoxon test is approximately 0.809 for the completely randomized design.

Case iv. Randomized paired design and Wilcoxon test

```r
set.seed(1003946983)
#Make 1000 replications
pvals <- replicate(1000,wilcox.test(rexp(n=20, rate = 1/3),
                                rexp(n=20, rate = 1),
                               paired = TRUE)$p.value)
#Calculate power
power4=sum(pvals<=0.05)/1000
```

The power of the Wilcoxon test is approximately 0.834 for the randomized paired design.

The t-test should be recommended for both of the two experimental designs since the t-test is robust and more powerful. We could see from the above calculation, for the completely randomized design, the power of the t-test and Wilcoxon test are 0.872 and 0.809 respectively, and for the randomized paired design, the power of the t-test and Wilcoxon test are 0.86 and 0.834 respectively. It is obvious that the power of each experimental design, t-test has higher power than the Wilcoxon test. Although the data does not satisfy the normality assumption of the t-test, the t-test is robust against non-normality, in other words, the t-test is robust to resistant to extreme values. While, the Wilcoxon test does not assume the normality of the data, but it has lower power. Thus, if I use the Wilcoxon test instead of the t-test, the lower power means I will less likely to do the right thing when the alternative is true. I will less likely to see the difference in lifetimes of the LED bulbs under two different temperature at S°C and T°C. Therefore, the t-test is recommended.