

STA305 a2

TONGFEI ZHOU

2/12/2021

Part I

1.) & 2.)

```
set.seed(1004738448) # set the seed as my student number
S <- rexp(9, rate = 1/3) # generate 9 observations from the treatment S
order_value <- 1:9
data.frame(round(S, 3), order_value) # print the observed value and the order value
```

##	round.S..3.	order_value
## 1	4.277	1
## 2	4.287	2
## 3	6.989	3
## 4	1.436	4
## 5	0.567	5
## 6	4.856	6
## 7	11.635	7
## 8	1.685	8
## 9	4.470	9

3.)

```
T <- rexp(9, rate = 1) # generate 9 observations from the treatment T
order_value <- 1:9
data.frame(round(T, 3), order_value) # print the observed value and the order value
```

##	round.T..3.	order_value
## 1	2.311	1
## 2	0.591	2
## 3	0.047	3
## 4	0.209	4
## 5	1.245	5
## 6	0.038	6
## 7	2.267	7
## 8	0.459	8
## 9	0.520	9

4)

```
data.frame(round(S, 3), round(T, 3)) # List the pairs of observations as required.
```

```
##   round.S..3. round.T..3.
## 1      4.277      2.311
## 2      4.287      0.591
## 3      6.989      0.047
## 4      1.436      0.209
## 5      0.567      1.245
## 6      4.856      0.038
## 7     11.635      2.267
## 8      1.685      0.459
## 9      4.470      0.520
```

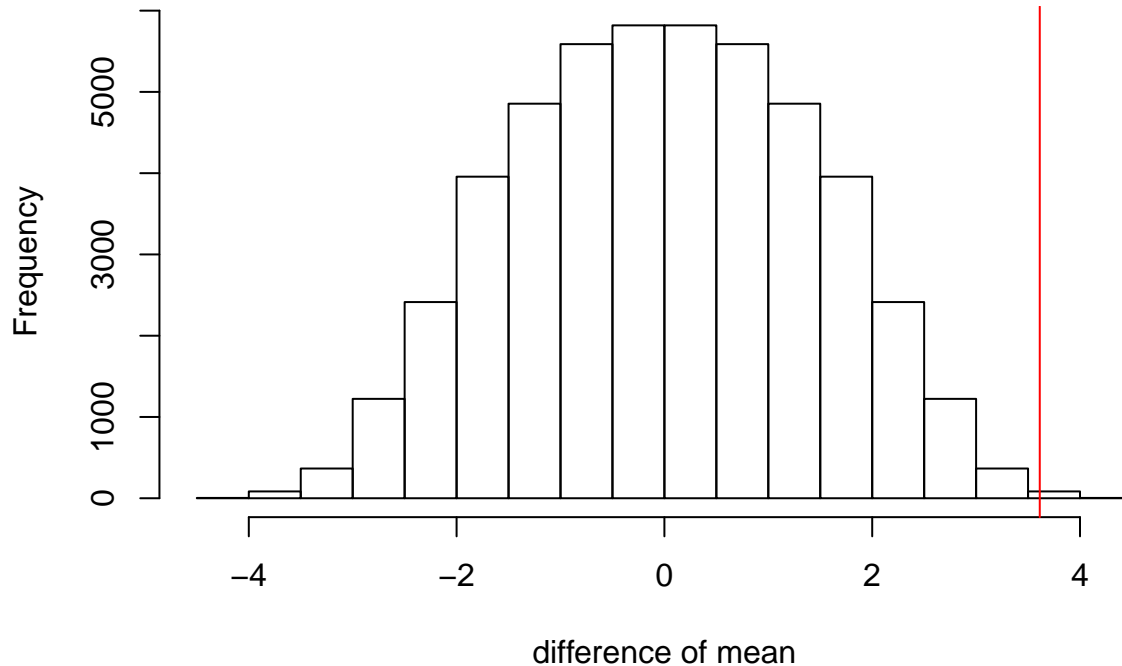
Above is the pairs of observations for the randomized paired design.

Part II

For completely randomized design, I set the null hypothesis $H_0 : \text{mean}(S) - \text{mean}(T) = 0$ and $H_a : \text{mean}(S) - \text{mean}(T) > 0$, and we have the following results:

```
bulbs_TZ <- c(S, T) # combine together to be ready for the randomization
N <- choose(18,9) # generate total number of the possible groups
observed_TZ <- mean(S) - mean(T) # observed difference mean
res <- numeric(N) # store the result
index <- combn(1:18, 9) # generate all the possible groups
# generate the results
for (i in 1:N){
  res[i] <- mean(bulbs_TZ[index[,i]]) - mean(bulbs_TZ[-index[,i]])
}
tbar_TZ <- mean(res)
pval_8448 <- sum((res-tbar_TZ) >= (observed_TZ - tbar_TZ))/N # one-sided p-value
hist(res, xlab = "difference of mean", main = "randomization distribution with P-value area 8448") # hi
abline(v = observed_TZ, col="red")
```

randomization distribution with P-value area 8448



```
print(pval_8448, digits = 2) # round the p-value to 2 significant digits
```

```
## [1] 0.0012
```

```
length(res) # print the number of values that the distribution contains
```

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

- i) As we can see from above, since we have two groups each with 9 numbers that are randomly assigned, we have total of $\binom{18}{9} = 48620$ number of values contained in the distribution. Therefore, the probability of a single observation, i.e., an observed treatment allocation is $\frac{1}{48620} \approx 2.06 \times 10^{-5}$.
- ii) I use the red line to mark the one-sided P-value as shown above.
- iii) Since the P-value is 0.0012, which means that we will only have the probability of 0.0012 to observe more or same extreme difference result as what we observed here, which is 3.612622 under the assumption of H_0 is true. However, 0.0012 is far smaller than the significance level of 5%, and so we reject the null hypothesis and so there is evidence of a difference in means between the two treatments.

For the randomized paired design, I set the null hypothesis $H_0 : \text{mean}(S) - \text{mean}(T) = 0$ and $H_a : \text{mean}(S) - \text{mean}(T) > 0$, and we have the following results:

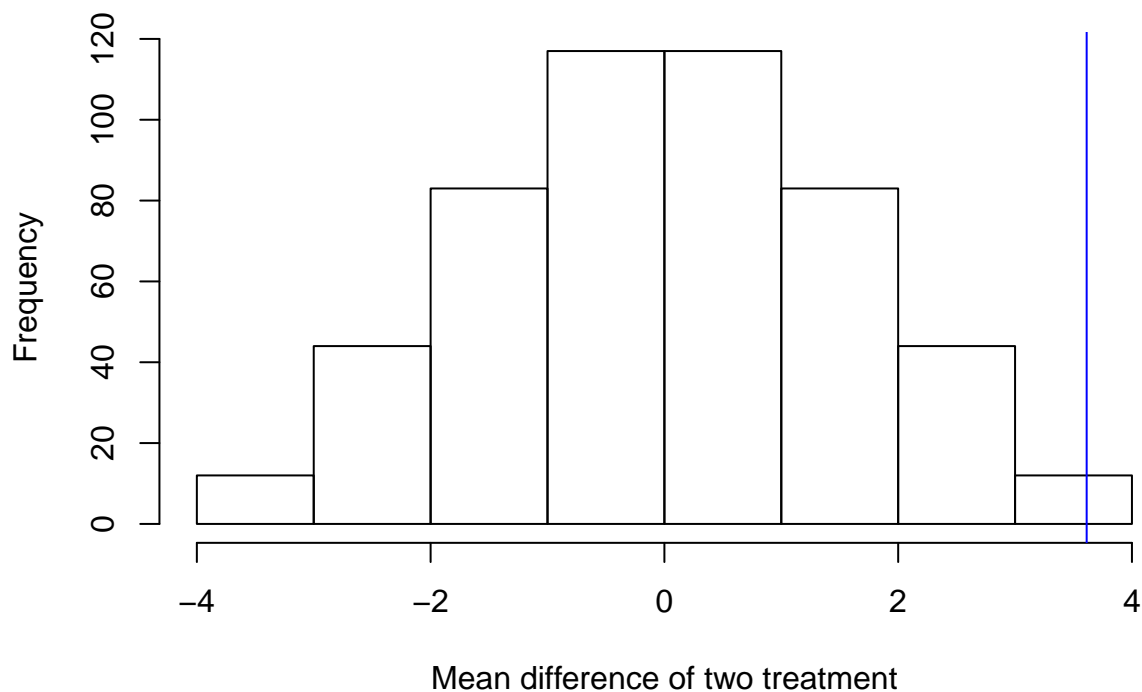
```

difference_8448 <- S - T
meandiff_TZ <- mean(difference_8448)
N_paired <- 2^9 # since we have 9 elements for each group
res_paired <- numeric(N_paired) # vector to store results
LR <- list(c(-1,1)) # difference is multiplied by -1 or 1
# generate all possible treatment assign
trtassign_8448 <- expand.grid(rep(LR, 9))

for (i in 1:N_paired){
  res_paired[i] <- mean(as.numeric(trtassign_8448[i,]) * difference_8448)
}
pval_paired <- sum(res_paired >= meandiff_TZ) / N_paired # p-value
hist(res_paired, xlab = "Mean difference of two treatment", main = "Randomizatio distribution with paired", col = "white", border = "black")
abline(v = meandiff_TZ, col = "blue") # p-value area indicating line

```

Randomizatio distribution with paired design



```

print(pval_paired, digits = 2) # round the p-value to 2 significant digits

```

```
## [1] 0.0039
```

```

length(res_paired) # print the number of values that the distribution contains

```

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

```
meandiff_TZ
```

```
## [1] 3.612622
```

- i) Since we are conducting the randomization paired test, we have $2^9 = 512$ number of ways in randomization since we need to test the two different treatments in pairs, and each treatment group contains 9 elements. Therefore, the probability of the observed treatment allocation is $\frac{1}{512} \approx 1.95 \times 10^{-3}$.
- ii) The histogram is shown above, with the blue vertical line to indicate the p-value area, which is the right side of the blue vertical line.
- iii) Since the p-value here is 0.0039, which means that we will only have the probability of 0.0039 to observe more or same extreme difference result as what we observed here, which is 3.612622 under the assumption of H_0 is true. However, 0.0039 is less than the significance level of 5%, and therefore we reject the null hypothesis and there is evidence of a difference, further, with mean of treatment S greater than that of treatment T.

Part III

For the completely randomized design, we have the t-test as follows:

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

##
## Welch Two Sample t-test
##
## data: S and T
## t = 3.1214, df = 9.1207, p-value = 0.006048
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  1.494216      Inf
## sample estimates:
## mean of x mean of y
## 4.466585 0.8540364
```

For the paired randomized design, we have the t-test as follows:

```
t.test(S, T, paired = TRUE, alternative = "greater")

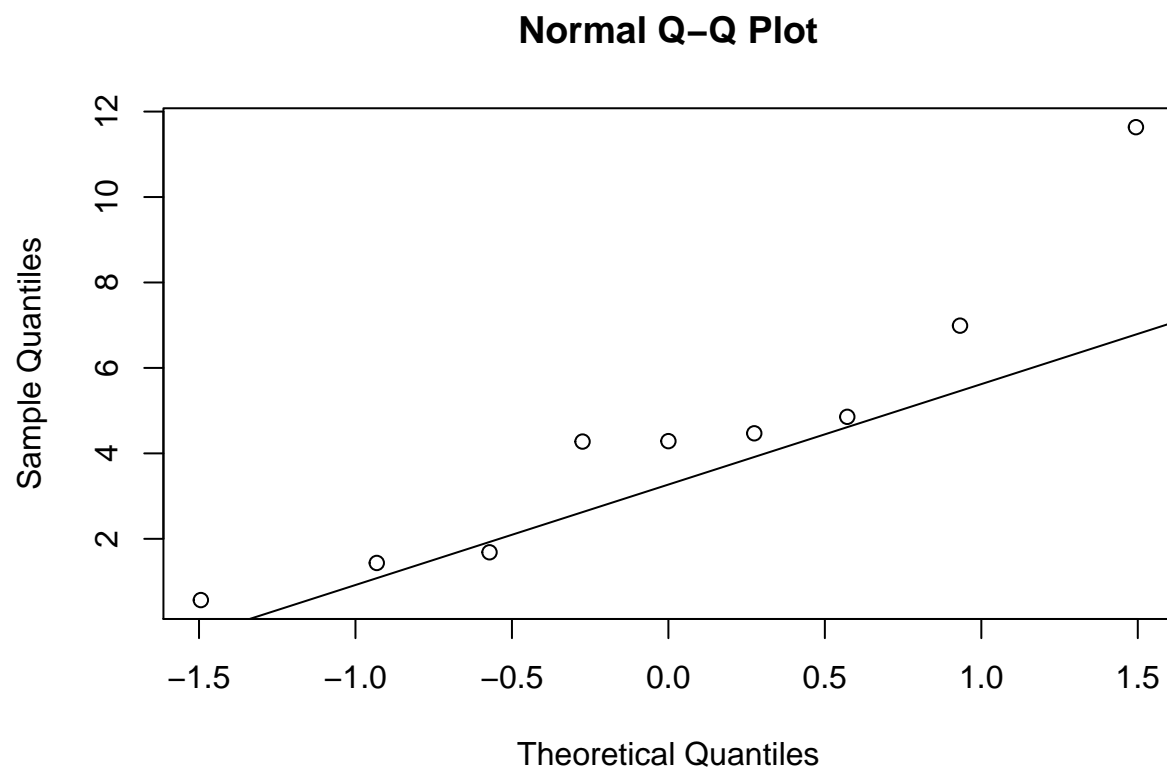
##
## Paired t-test
##
## data: S and T
## t = 3.4643, df = 8, p-value = 0.004257
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  1.673444      Inf
## sample estimates:
## mean of the differences
##              3.612622
```

i) As we can see from above, the p-value generated from both designs are significant, one is 0.006048 and the other is 0.004257, with one test statistic being 3.1214 and the other being 3.4643, and both p-values are less than 0.05. Therefore, we reject the null hypothesis that the mean of two treatments are the same, and accept the alternative hypothesis that true difference in means is greater than 0 as listed in the results printed above.

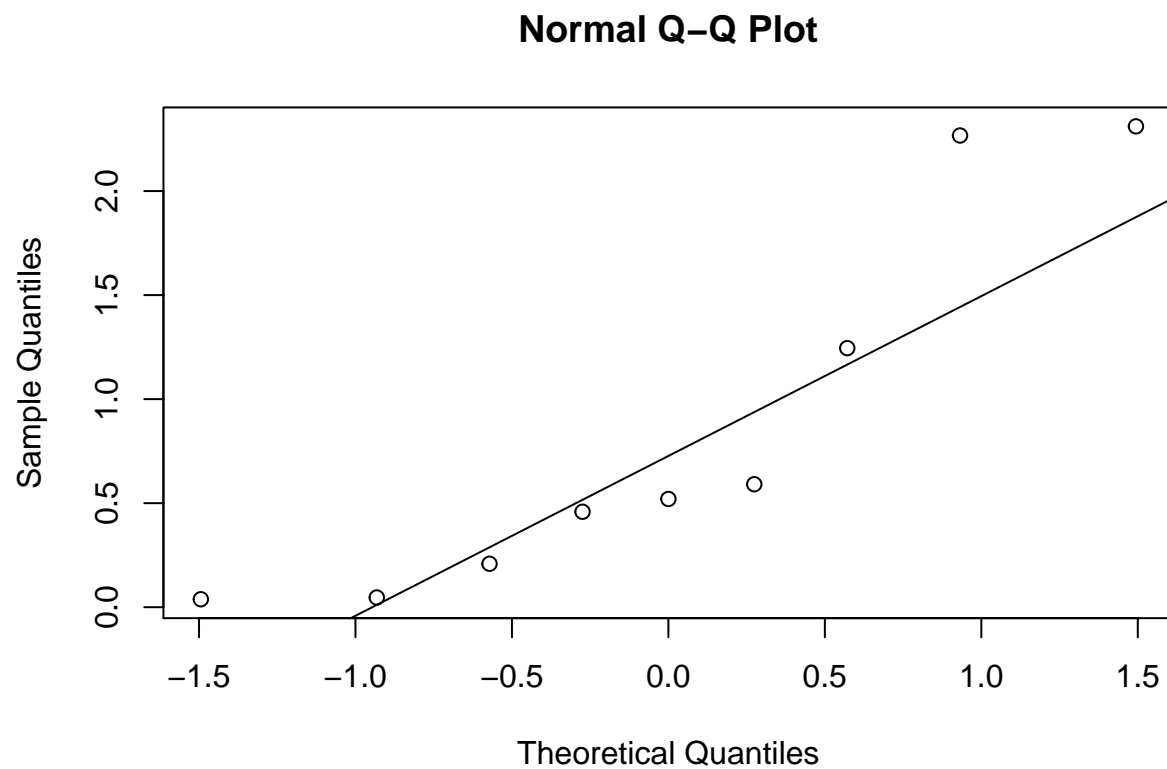
ii)

For the completely randomized design, we check whether two treatments' elements are normally distributed.

```
qqnorm(S) # check for treatment S
qqline(S)
```

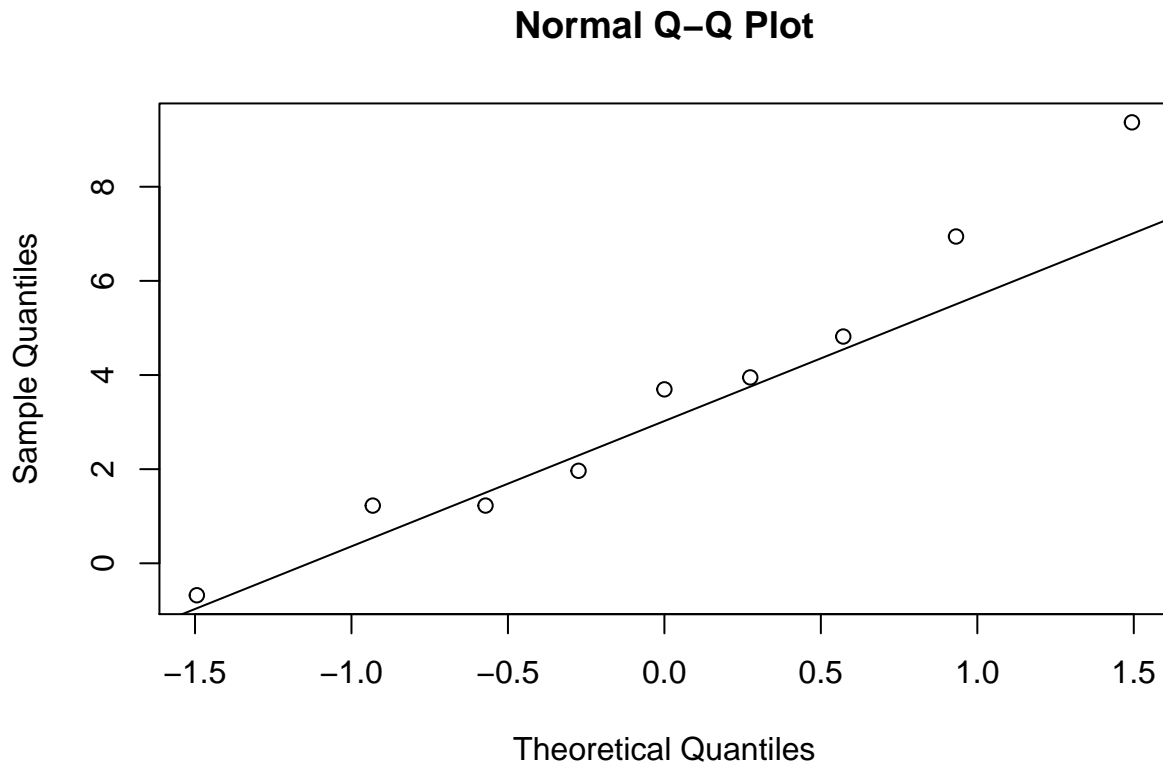


```
qqnorm(T) # check for treatment T
qqline(T)
```



As we can see, the points are deviating quite far from the line. Therefore, the assumption dose not hold.
For the paired randomized design, we check whether the difference is normally distributed.

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



As we can see here, the points deviate the line quite much, so the assumption does not hold for t-test under paired randomized design.

- iii) As we can see from above, the p-values in the both t-test are not the same as what we get from the randomization tests. In the completely randomized design, the p-value for t-test is 0.006048 and the p-value for completely randomized design is 0.0012 since the assumption for the t-test in the completely randomized design does not hold well. However, both p-values indicate that we have evidence that there is difference in the mean of two treatments.

In the paired randomized design, the p-value for t-test is 0.004257 and the p-value for the paired randomized design is 0.0039 since the variance is not equal and against normal assumption. However, both p-values indicate that we have evidence that there is difference, and further, mean of treatment S is greater than mean of treatment T.

However, since the assumption for the t-test does not hold very well here, the results produced by the t-test should not be considered.

Part IV

- i)

```
set.seed(1004738448) # set the seed as my student number
# calculate the simulated power of completely randomized design with t test
pvals_ttest_CRD <- replicate(1000, t.test(rexp(20, 1/3),
                                           rexp(20, 1),
```



```

                                var.equal = FALSE)$p.value)
power_ttest_CRD <- sum(pvals_ttest_CRD<=0.05)/1000

```

ii)

```

# calculate the simulated power of randomized paired design with t test
pvals_ttest_RPD <- replicate(1000, t.test(rexp(20, 1/3),
                                rexp(20, 1),
                                var.equal = FALSE, paired = TRUE)$p.value)
power_ttest_RPD <- sum(pvals_ttest_RPD<=0.05)/1000

```

iii)

```

# calculate the simulated power of completely randomized design with Wilcoxon test
pvals_wilcox_CRD <- replicate(1000, wilcox.test(rexp(20, 1/3),
                                rexp(20, 1),
                                var.equal = FALSE)$p.value)
power_wilcox_CRD <- sum(pvals_wilcox_CRD<=0.05)/1000

```

iv)

```

# calculate the simulated power of randomized paired design with Wilcoxon test
pvals_wilcox_RPD <- replicate(1000, wilcox.test(rexp(20, 1/3),
                                rexp(20, 1),
                                var.equal = FALSE, paired = TRUE)$p.value)
power_wilcox_RPD <- sum(pvals_wilcox_RPD<=0.05)/1000

```

```

# List the power results
power_ttest_CRD

```

```
## [1] 0.855
```

```
power_ttest_RPD
```

```
## [1] 0.877
```

```
power_wilcox_CRD
```

```
## [1] 0.795
```

```
power_wilcox_RPD
```

```
## [1] 0.841
```

As we can see from above, under completely randomized design, we have power for t-test being greater than the Wilcoxon test, but since the assumption for the t-test does not hold in this simulation since the samples are from exponential distribution, we should not use t-test at all. Therefore, I will recommend the Wilcoxon test for completely randomized design.

As we can see from above, under randomized paired design, we have power for t-test being larger than the Wilcoxon test, and since the assumption for the t-test does not hold as well in this simulation, I will recommend the Wilcoxon test for randomized paired design as well.