Analysis of Average Lifetime of LED Bulbs by Different Tests

Xingnuo Zhang

2/9/2021

Part I

```r
library(knitr)
set.seed(1006145306)       # generate a sequence of random observations below
treS <- rexp(9, rate=1/3)  # Exponential Distribution, lambda equals 1/3
round(treS, 3)             # Round the observation values to 3 decimal places
```

```
## [1] 0.086 3.206 4.322 3.032 2.477 6.636 0.784 1.578 0.811
```

```r
treT <- rexp(9, rate=1)    # Exponential Distribution, lambda equals 1
round(treT, 3)
```

```
## [1] 0.734 0.510 2.587 0.080 0.508 0.572 0.782 0.571 0.414
```

```r
difference <- treS-treT
meandiff <- mean(difference)
# Display the pairs of observations
plot_data_1 <- data.frame(treS, treT, difference)
# Keep treS, treT, difference to 3 decimal places in the presented table
plot_data <- round(plot_data_1, digits = 3)
kable(plot_data)
```

treS	treT	difference
0.086	0.734	-0.648
3.206	0.510	2.696
4.322	2.587	1.735
3.032	0.080	2.952
2.477	0.508	1.970
6.636	0.572	6.064
0.784	0.782	0.002
1.578	0.571	1.008
0.811	0.414	0.397

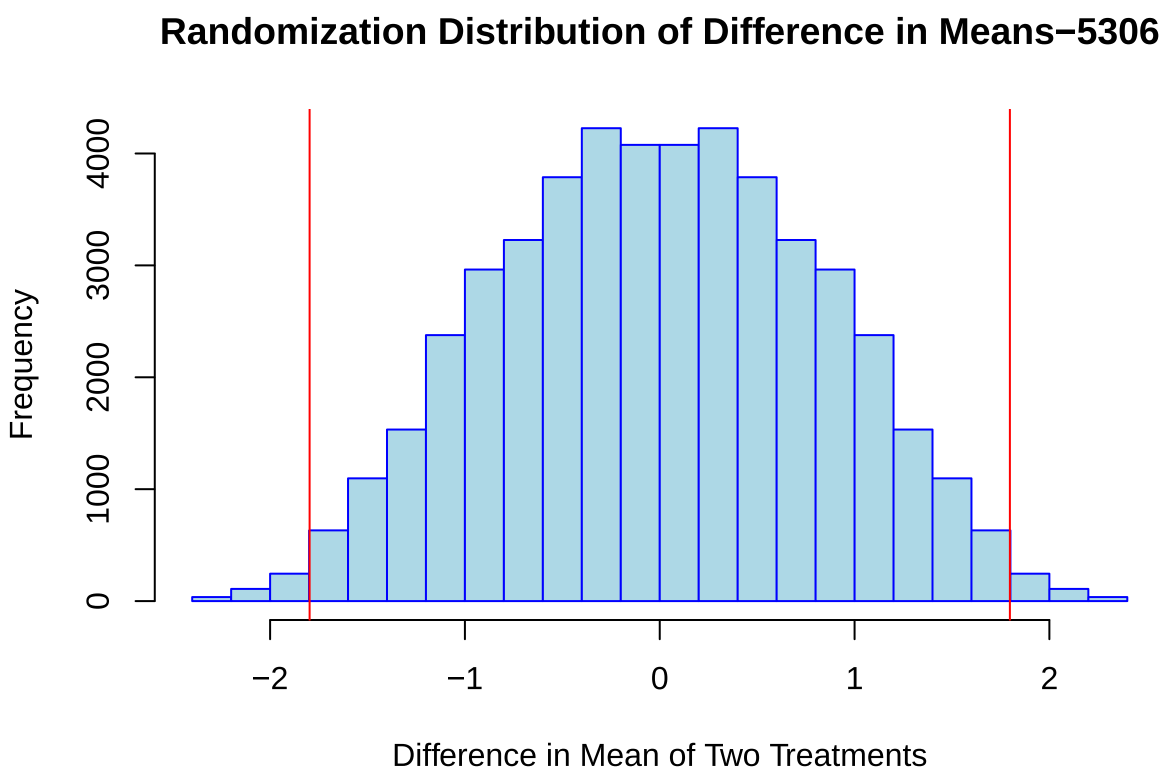
Part II

Two-Sided Randomization Test

i. In two-sided randomization test, we randomly create 9 observations in group S and T, so that we have 18 observations in total. It is an effective way to compare the means of two groups. The number of values that the distribution contains is $\binom{18}{9}$=48620, and the probability of the observed treatment is $1/\binom{18}{9}$.

ii.

```r
my_data <- c(treS, treT)                    # pool data
observed_data <- mean(treS)-mean(treT)  # store observed difference of mean
N1 <- choose(18, 9)
result <- numeric(N1)                       # store the results
# Generate N treatment assignments
index <- combn(1:18, 9)
for (i in 1:N1)
{
  result[i] <- mean(my_data[index[,i]]) - mean(my_data[-index[,i]])
}
# Create histogram plot of two-sample randomization test
hist(result, xlab = "Difference in Mean of Two Treatments",
     main = "Randomization Distribution of Difference in Means-5306",
     border = "blue", col = "light blue", freq = TRUE)
# add vertical lines at observed mean difference on the both sides
abline(v = observed_data, col = "red")
abline(v = -observed_data, col = "red")
```




```r
# Calculate the two-sided p-value for randomization test
tbar_1 <- mean(result)
pvalue_1 <- sum(abs(result-tbar_1) >= abs(observed_data-tbar_1))/N1
round(pvalue_1, 4)                         # Round the p-value to 4 decimal places
```

```
## [1] 0.0162
```

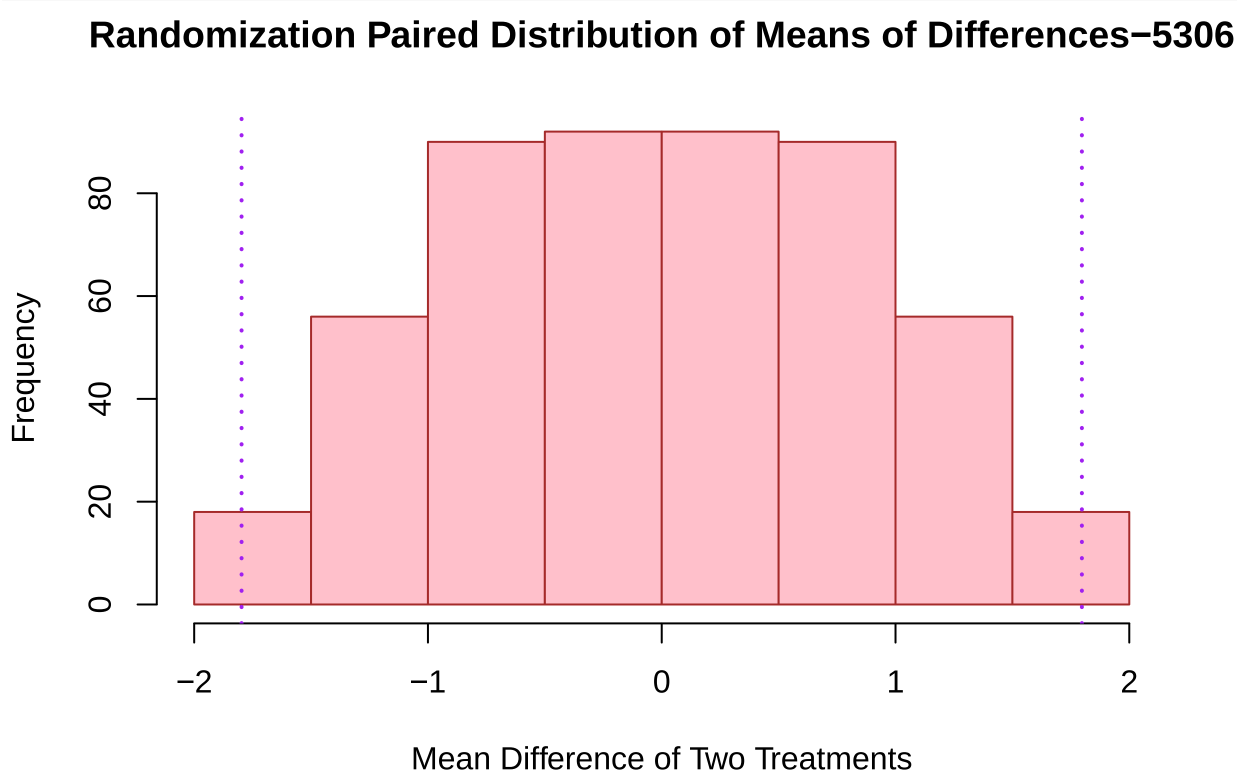
iii. The null hypothesis is the average life expectancy of LED bulbs under treatment S and T is equal (H0: $\mu_S = \mu_T$), and the alternative hypothesis is the average lifetime of LED bulbs under treatment S and T is not equal (Ha: $\mu_S \neq \mu_T$). The result of the p-value we calculated above is 0.0162, which means there is no difference in the average lifetime of LED bulbs under treatment S and T then the proportion of randomization would produce an observed mean difference between treatment S and T of at most 1.79727 is 0.0162. In other words, it is implausible that an observed mean difference as large as 1.79727 would be detected if the expected lifetime of LED bulbs is the same under treatment S and T. Besides, the p-value is less than 0.05, which means very significant. Therefore, we reject the null hypothesis, and there is strong evidence that the average lifetime of LED bulbs is different under treatment S and T.

Randomized Paired Comparison

i. In the paired randomization test, 18 observations will be classified into 9 pairs. In this case, the distribution has number of values of 2^9, and the probability of the observed treatment is $1/2^9$.

ii.

```r
N2 <- 2^(9)                              # total number of treatment assignments
result_paired <- numeric(N2)             # use to store the results
LR <- list(c(-1,1))                      # difference is multiplied by -1 or 1
trtassign <- expand.grid(rep(LR,9))      # generate all treatment assignments
for (i in 1:N2)
{
  result_paired[i] <- mean(as.numeric(trtassign[i,])*difference)
}
# Create histogram plot of paired randomization test
hist(result_paired, xlab = "Mean Difference of Two Treatments", main =
"Randomization Paired Distribution of Means of Differences-5306",
border = "brown", col = "pink")
abline(v = observed_data, lty = 'dotted', col = "purple", lwd = 2)
abline(v = -observed_data, lty = 'dotted', col = "purple", lwd = 2)
```



```r
# Calculate the two-sided p-value for paired randomization test
tbar_2 <- mean(result_paired)
pvalue_2 <- sum(abs(result_paired-tbar_2) >= abs(observed_data-tbar_2))/N1
round(pvalue_2, 5)
```

```
## [1] 0.00021
```

iii. The null hypothesis is the average life expectancy of LED bulbs under treatment S and T is equal (H0: $\mu_S = \mu_T$), and the alternative hypothesis is the average lifetime of LED bulbs under treatment S and T is not equal (Ha: $\mu_S \neq \mu_T$). From part ii, the result of the p-value is 0.00021, which shows that nearly zero paired randomization would produce an extreme difference than the observed difference under the null hypothesis. It is unlikely to detect an observed absolute mean difference as large as or larger than 1.79727 if the expected lifetime of LED bulbs under treatment S and T are identical. Also, the p-value is significant at a 5% significance level. In this case, we reject the null hypothesis as our data supports the average lifetime of LED bulbs is different under treatment S and T.

Part III

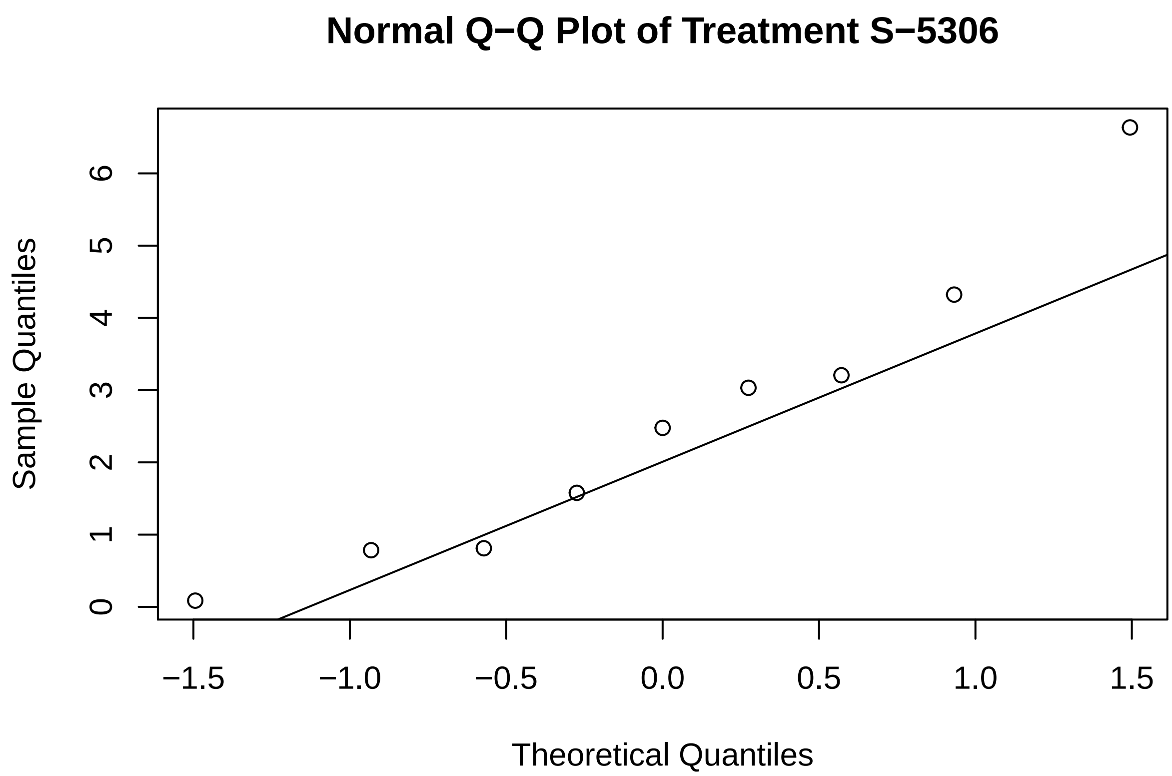
Two-sample T-test

```r
t.test(treS, treT, var.equal = FALSE, alternative = "two.sided")
```

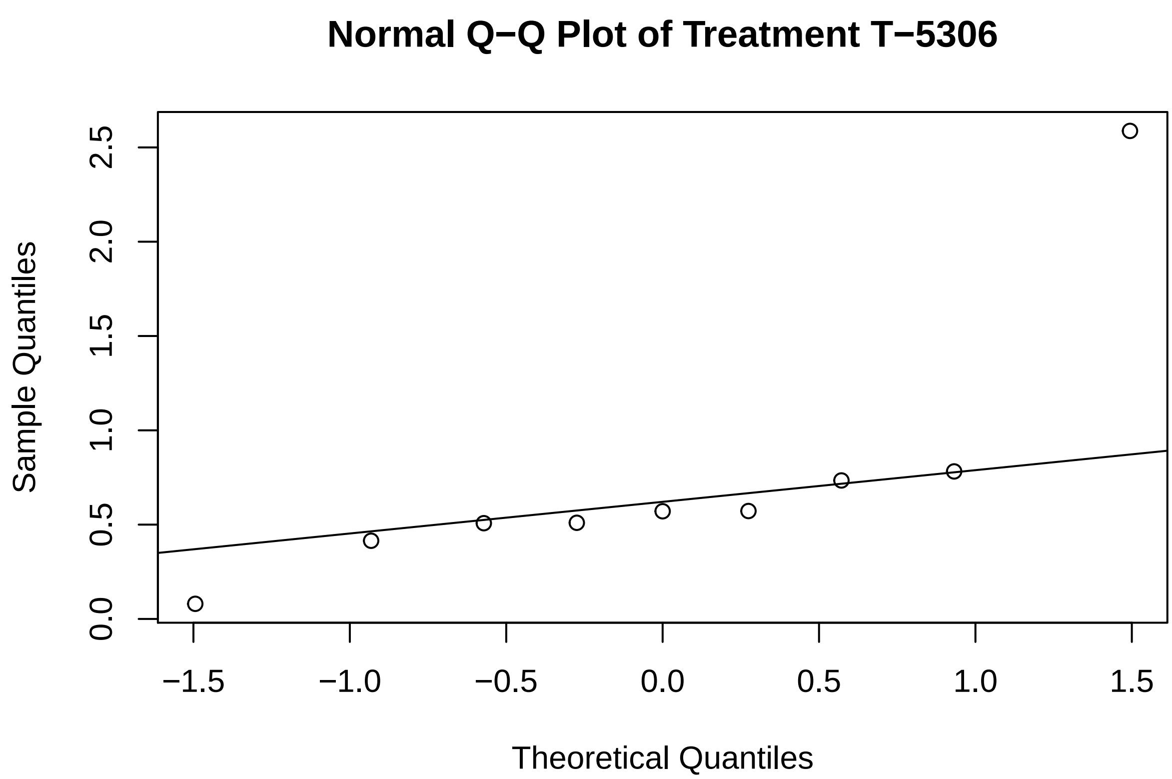
```
##
##  Welch Two Sample t-test
##
## data:  treS and treT
## t = 2.4813, df = 9.9291, p-value = 0.03263
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.1817795 3.4127654
## sample estimates:
## mean of x mean of y
##  2.548255  0.750982
```

i. When conducting the two-sample t-test, the null hypothesis is the average lifetime of LED bulbs under treatment S and T is equal (H0: $\mu_S = \mu_T$), and the alternative hypothesis is the average lifetime of LED bulbs under treatment S and T is not equal (Ha: $\mu_S \neq \mu_T$). From the result table, the p-value is 0.033, which is smaller than 0.05. Therefore, we reject the null hypothesis, and we have strong evidence to conclude that the two means of the lifetime of LED bulbs under two treatments are different ($\mu_S \neq \mu_T$).

```r
qqnorm(treS, main = "Normal Q-Q Plot of Treatment S-5306")
qqline(treS)
```




```r
qqnorm(treT, main = "Normal Q-Q Plot of Treatment T-5306")
qqline(treT)
```



ii. As we know, the t-test is a useful tool to determine the significant difference between the means of the two groups' data. When conducting a t-test, two main assumptions need to be satisfied:

independence and normality. Firstly, as we generate the sample data separately, the observations in each group are independent. Secondly, from the normal Q-Q plots for treatment S and T above, it is not hard to notice that most points are not on the fitted line, which indicates that it does not follow the normal distribution. Therefore, the assumptions behind the two-sample t-test are not satisfied.

iii. The p-value of the two-sided randomization test equals 0.016, while the p-value of the two-sample t-test is 0.033. The two results are significant at a 5% significance level; we have strong evidence to reject the null hypothesis. As the conclusions correspond, the result of the t-test agrees with the result of the randomization test.

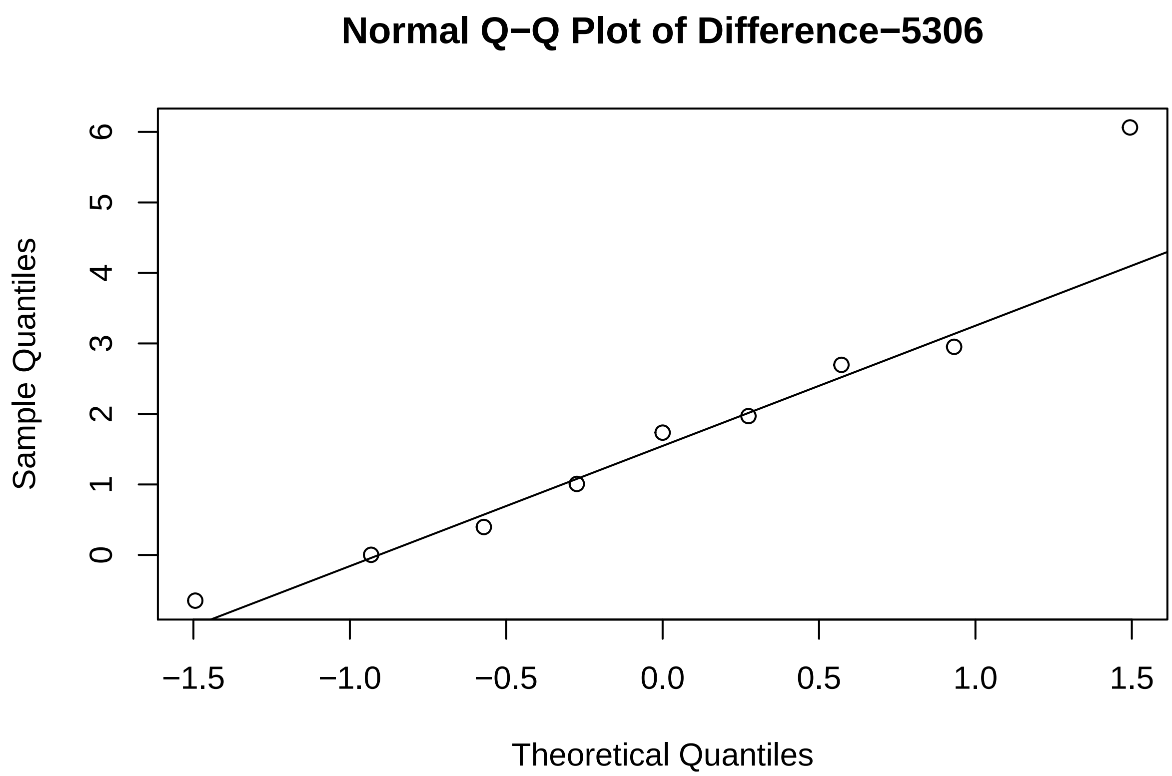
Paired T-test

```r
t.test(treS, treT, alternative = "two.sided", paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  treS and treT
## t = 2.6891, df = 8, p-value = 0.02754
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.2560181 3.3385269
## sample estimates:
## mean of the differences
##                1.797272
```

i. When conducting the paired t-test, the null hypothesis is the average lifetime of LED bulbs under treatment S and T is equal (H0: $\mu_S = \mu_T$), and the alternative hypothesis is the average lifetime of LED bulbs under treatment S and T is not equal (Ha: $\mu_S \neq \mu_T$). From the result table, the p-value is 0.02754, which means there is no difference in the average lifetime of LED bulbs under treatment S and T then the proportion of randomization would produce an observed mean difference between treatment S and T of at most 1.79727 is 0.02754. Therefore, we reject the null hypothesis and conclude that the average lifetime of LED bulbs under treatment S and T is not equal.

```r
qqnorm(difference, main = "Normal Q-Q Plot of Difference-5306")
qqline(difference)
```



ii. Firstly, as we generate the sample data separately, the independence of observations in each group are satisfied. Secondly, from the normal Q-Q plots for the difference between treatment S and T above, most points are on the fitted line, which indicates that it does follow the normal distribution. Therefore, the assumptions behind the two-sample t-test are satisfied.

iii. The p-value of the two-sided randomization test equals 0.0002, while the p-value of the two-sample t-test is 0.0275. Both the results are significant at a 5% significance level. As the conclusions correspond, the result of the t-test agree with the result of the paired randomization test.

Part IV

i. Completely randomized design and t-test

```
set.seed(1006145306)
pvalue_3 <- replicate(1000, t.test(rexp(20, rate=1/3), rexp(20, rate=1),
                                   alternative = "two.sided")$p.value)
# Power of two-sample t-test by simulation is the number of times the test
# reject null hypothesis at the 5% threshold over total simulation number
sum(pvalue_3<=0.05)/1000
```

```
## [1] 0.868
```

ii. Randomized paired design and t-test

```
pvalue_4 <- replicate(1000, t.test(rexp(20, rate=1/3)-rexp(20, rate=1),
                                   alternative = "two.sided")$p.value)
# Power of paired t-test
sum(pvalue_4<=0.05)/1000
```

```
## [1] 0.852
```

iii. Completely randomized design and Wilcoxon test

```r
pvalue_5 <- replicate(1000, wilcox.test(rexp(20, rate=1/3), rexp(20, rate=1),
                                        alternative = "two.sided")$p.value)
# Power of two-sample Wilcoxon test
sum(pvalue_5<=0.05)/1000
```

```
## [1] 0.818
```

iv. Randomized paired design and Wilcoxon test

```r
pvalue_6 <- replicate(1000, wilcox.test(rexp(20, rate=1/3)-rexp(20, rate=1),
                                        alternative = "two.sided")$p.value)
# Power of paired Wilcoxon test
sum(pvalue_6<=0.05)/1000
```

```
## [1] 0.835
```

From the calculations above, I would recommend the t-test since it has larger power than the Wilcoxon test. Although the Wilcoxon test is non-parametric and does not require the data to follow the normal distribution, it has lower power comparatively. Focus on the t-test; although the normality assumption of the t-test is not satisfied since the bulb's lifetime follows an exponential distribution, the t-test is robust towards non-normality. The considerable power provides higher accuracy, the t-test is more likely to detect a statistically significant difference in the average lifetime of LED bulbs under treatment S and T.

Reference

1. Hayes, Adam, "Wilcoxon Test", Investopedia, Apr 2020, https://www.investopedia.com/terms/w/wilcoxon-test.asp
2. Bartlett, Jonathan, "the T-test and Robustness to Non-normality", The Stats Geek, Sep 2013, https://thestatsgeek.com/2013/09/28/the-t-test-and-robustness-to-non-normality/
3. Shivon Sue-Chee, "STA305/1004-Class 3: Comparing Two Treatments", https://q.utoronto.ca/courses/206595/files/11952346?wrap=1
4. Shivon Sue-Chee, "STA305/1004-Week 3: Paired to Power", https://q.utoronto.ca/courses/206595/files/11999091?wrap=1
5. Shivon Sue-Chee,"STA305/1004-Week 4: Finding Power, Intro to Causal Inference", https://q.utoronto.ca/courses/206595/files/12143630?wrap=1