Homework 2

sbsambado

10/11/2021

PSTAT220A

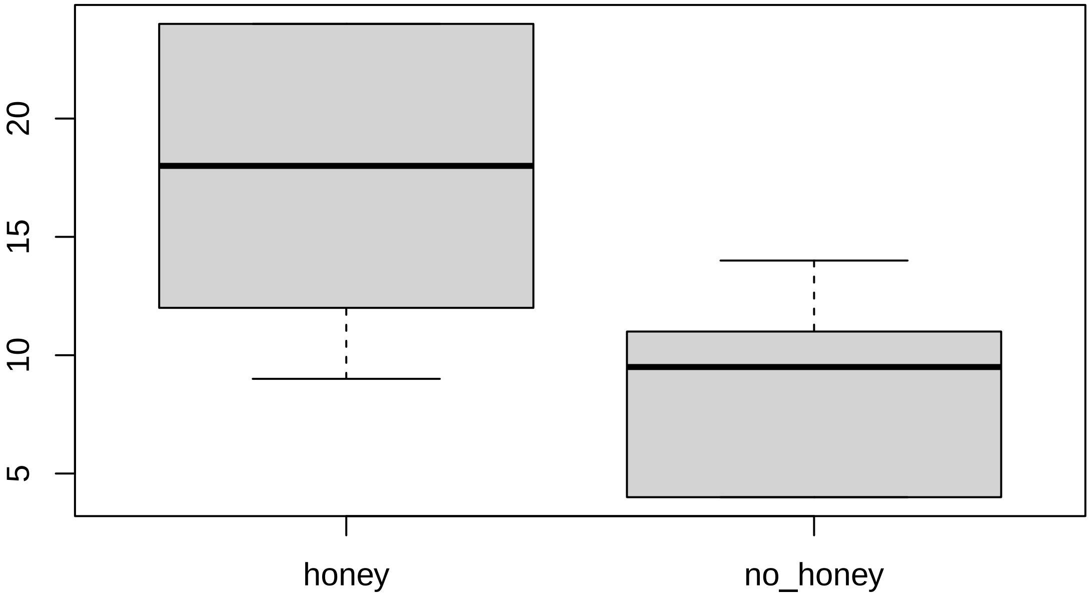
Homework 2: Due 10/20/21

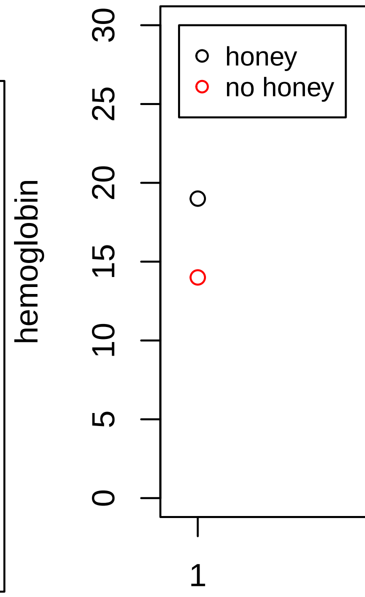
Question 1 paired t-test, singed rank test, two sample t-test (equal variances or unequal variances), Wilcoxon rank test, other

A) ANSWER: Because the researcher expected twinship to eliminate external differences to hemoglobin, a *paired t-test* would be conducted (although I personally think twins are too different to compare metabolisms).

- H0 (null hypothesis): There is no difference in the mean hemogloblin levels between honey and non-honey drinker in twin groups (the difference in the mean = 0).
- HA (alternative hypothesis): There is a difference in the mean hemogloblin levels between honey and non-honey drinker in twin groups.

In conclusion, *there is a significant difference* in the mean hemogloblin levels between honey and non-honey drinker in twin groups. The *p value is 0.005218*, which is < 0.05 and means we can reject our H0. Also, because our 95% CI (4.028033 - 13.638633) does not include 0, then we know the mean difference between the pair is not equal to 0, it is 8.83.





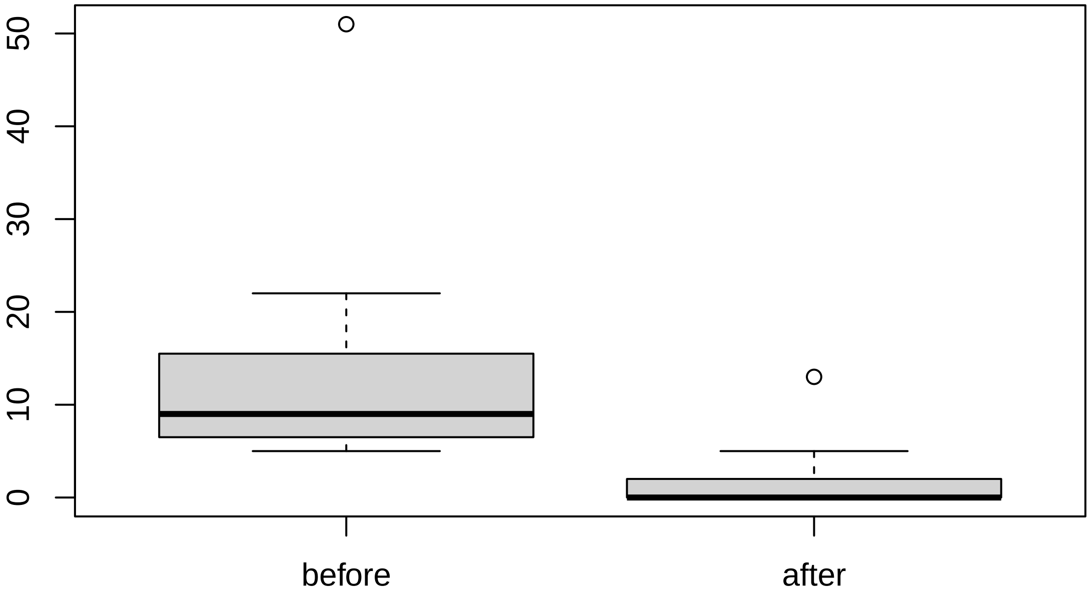
```
##
##  Paired t-test
##
## data:  honey and no_honey
## t = 4.7254, df = 5, p-value = 0.005218
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   4.028033 13.638633
## sample estimates:
## mean of the differences
##                8.833333
```

```
##
##  F test to compare two variances
##
## data:  honey and no_honey
## F = 2.274, num df = 5, denom df = 5, p-value = 0.3883
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##   0.318203 16.250872
## sample estimates:
## ratio of variances
##              2.274
```

B) ANSWER: I would run a *paired t-test* because we are comparing the difference of premature beats in the same patient.

- H0 (null hypothesis): There is no difference in the mean number of premature beats per minutes between the before and after groups (the difference in the mean = 0).
- HA (alternative hypothesis): There is a difference in the mean number of premature beats per minutes between the before and after groups.

In conclusion, *there is a significant difference* in the mean number of premature beats per minutes between the before and after groups. The *p value is 0.01763*, which is < 0.05 and means we can reject our H0. Also, because our 95% CI (2.423084 20.576916) does not include 0, then we know the mean difference between the pair is not equal to 0, it is 11.5.



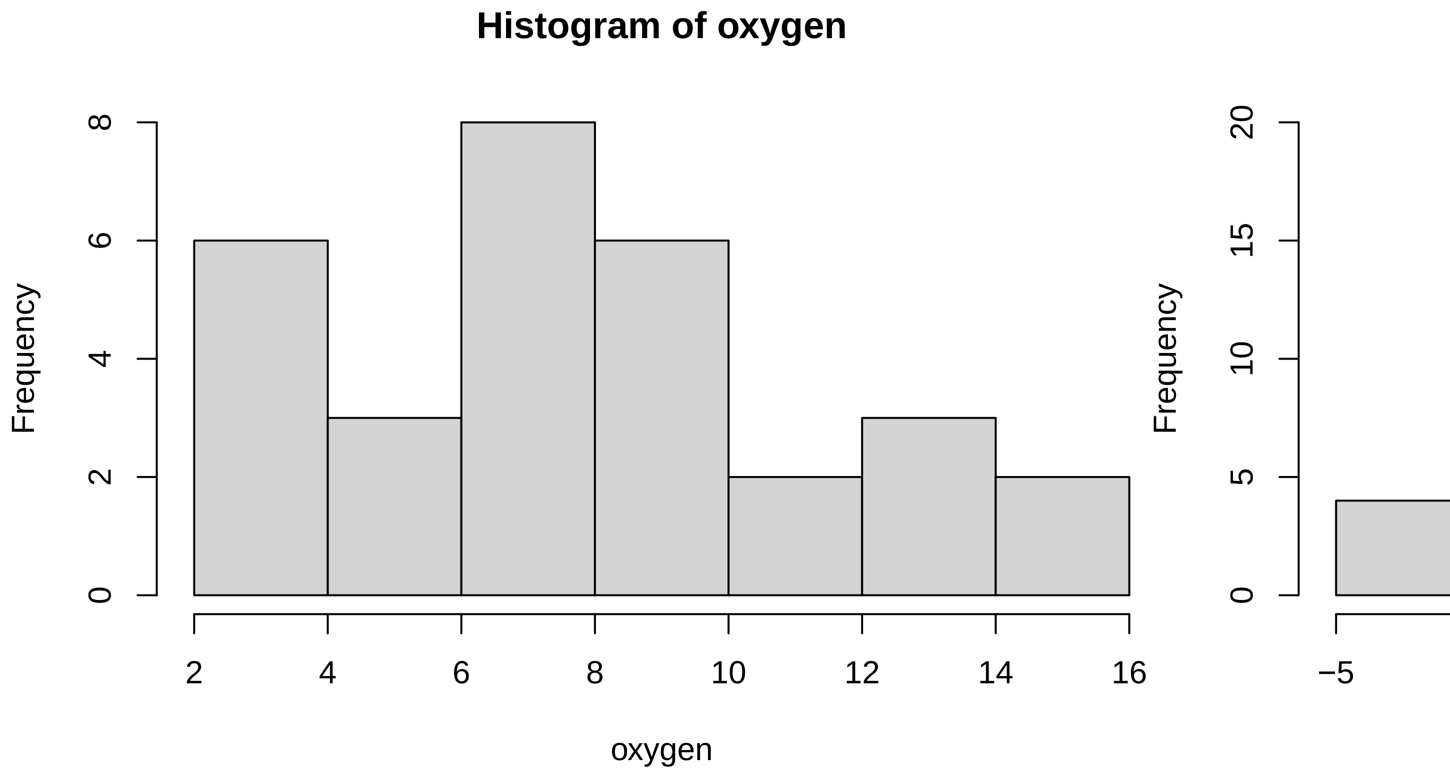

```
##
##  Paired t-test
##
## data:  before and after
## t = 2.7885, df = 11, p-value = 0.01763
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   2.423084 20.576916
## sample estimates:
## mean of the differences
##                    11.5
```

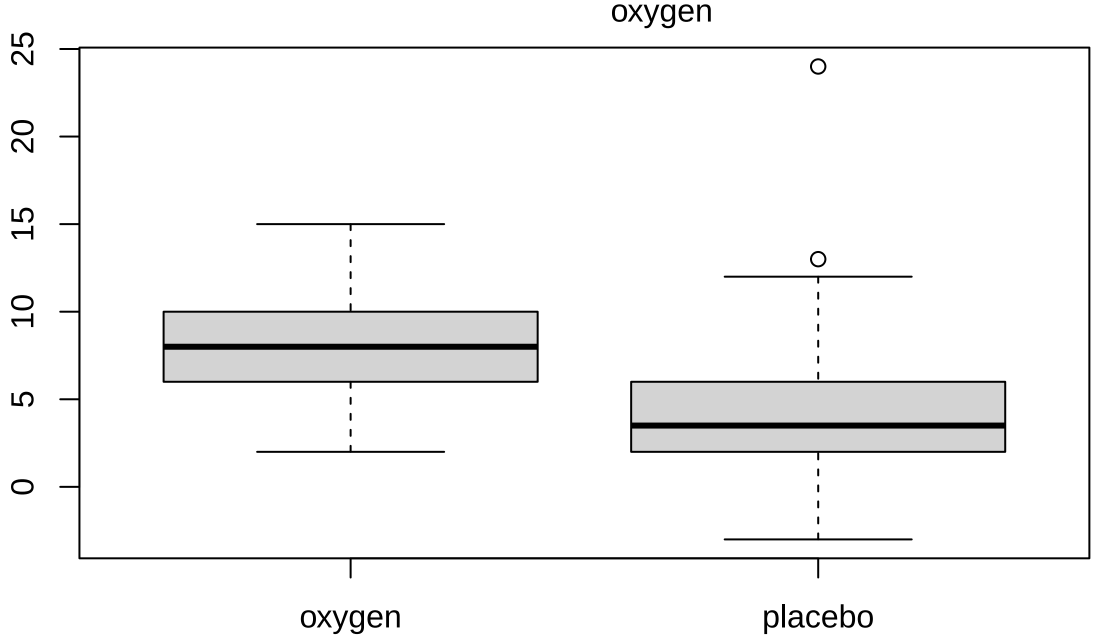
D) ANSWER: I would run a *two Sample t-test (equal variance)* because we are comparing the mean between two groups (oxygen vs placebo gel). The distribution of both groups is trending normal, although there are two outliers in the placebo group which skews it a bit. And the variances are equal (p = 0.2261).

- H0 (null hypothesis): There is no difference in the mean oral hygiene index between the oxygen gel and placebo gel. (mu_oxygen = mu_placebo)
- HA (alternative hypothesis): There is a difference in the mean oral hygiene index between the oxygen gel and placebo gel.

In conclusion, *there is a significant difference* in the mean oral hygiene index between the oxygen gel and placebo gel. The *p value is 0.001077*, which is < 0.05 and means we can reject our H0. Also, because our 95% CI (1.519477 - 5.762876) does not include 0, then we know the mean of both groups are not equal to 0. The mean oral index for oxygen gel and placebo gel, is 8.20 and 4.56, respectively.

```
## Warning in cbind(oxygen, placebo): number of rows of result is not a multiple of
## vector length (arg 1)
```






```
##
##  Welch Two Sample t-test
##
## data:  oxygen and placebo
## t = 3.4787, df = 61.449, p-value = 0.0009321
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.548450 5.733903
## sample estimates:
## mean of x mean of y
```

```
##  8.200000  4.558824
```

```
##
##  F test to compare two variances
##
## data:  oxygen and placebo
## F = 0.64022, num df = 29, denom df = 33, p-value = 0.2261
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.3146534 1.3247171
## sample estimates:
## ratio of variances
##          0.6402219
```

```
##
##  Two Sample t-test
##
## data:  oxygen and placebo
## t = 3.4306, df = 62, p-value = 0.001077
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.519477 5.762876
## sample estimates:
## mean of x mean of y
##  8.200000  4.558824
```

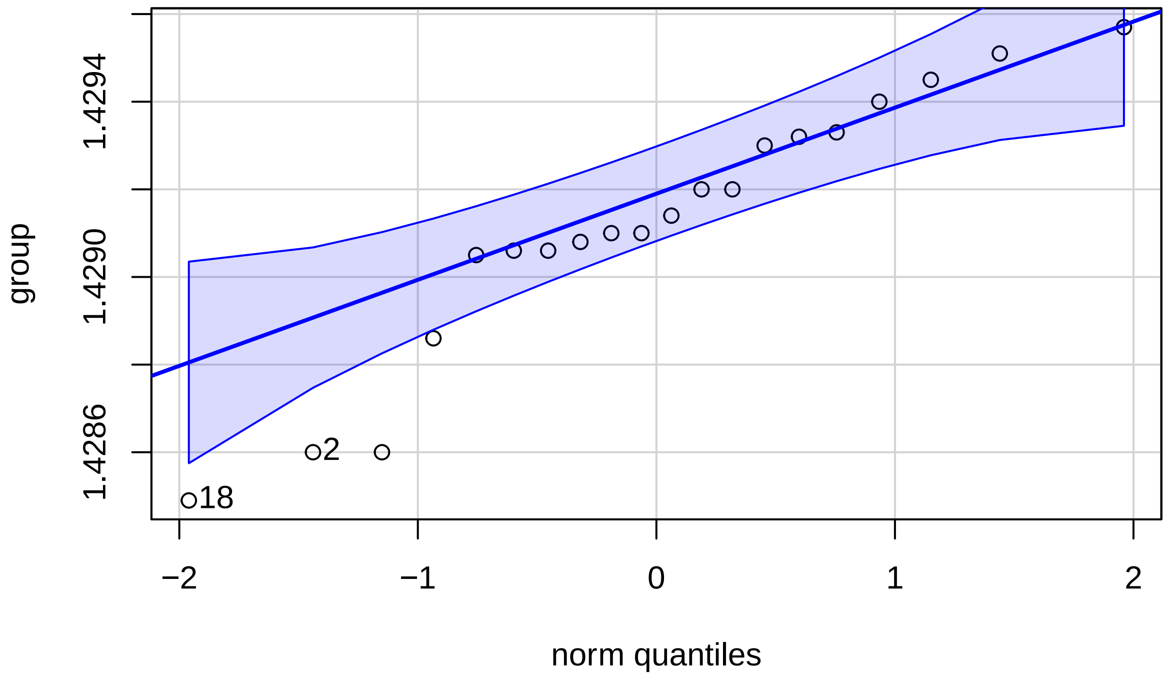
E) ANSWER: I would run a *Wilcoxon rank sum test* because we are comparing the mean between two groups (potassium vs electrolytic oxygen) but their distributions are not normal with some outliers.

- H0 (null hypothesis): The true location shift is equal to 0 (median) for the potassium vs electrolytic oxygen groups. (M_d = 0)
- HA (alternative hypothesis): The true location shift is not equal to 0 (median) for the potassium vs electrolytic oxygen groups. (M_d /= 0)

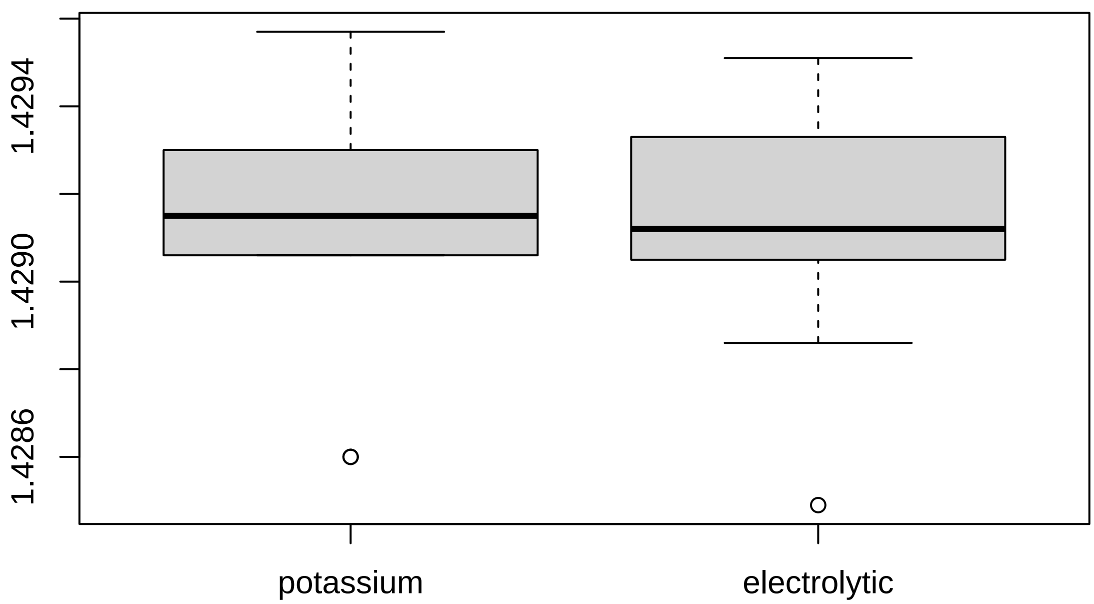
In conclusion, *there is no significant difference* in the median between the potassium vs electrolytic oxygen groups. The *p value is 0.6961*, which is > 0.05 and means we fail to reject our H0.

```
## Warning in cbind(potassium, electrolytic): number of rows of result is not a
## multiple of vector length (arg 1)
```

```
## Loading required package: carData
```




```
## [1] 18  2
```

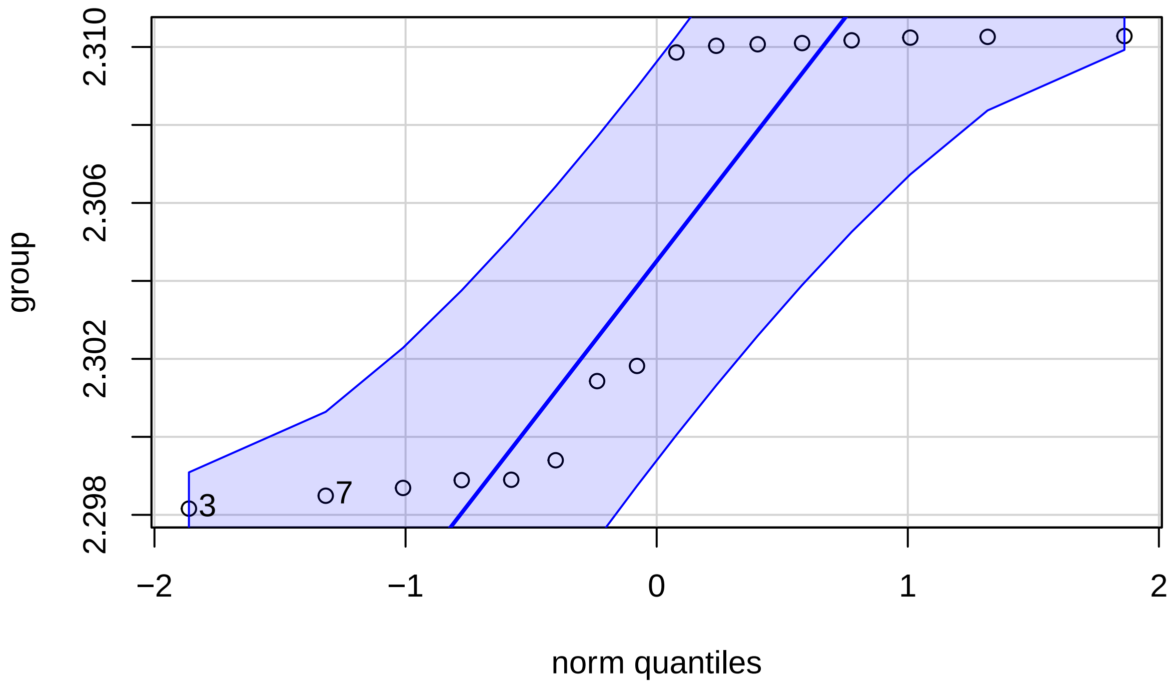



```
## Warning in wilcox.test.default(potassium, electrolytic): cannot compute exact p-
## value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  potassium and electrolytic
## W = 39.5, p-value = 0.6961
## alternative hypothesis: true location shift is not equal to 0
```

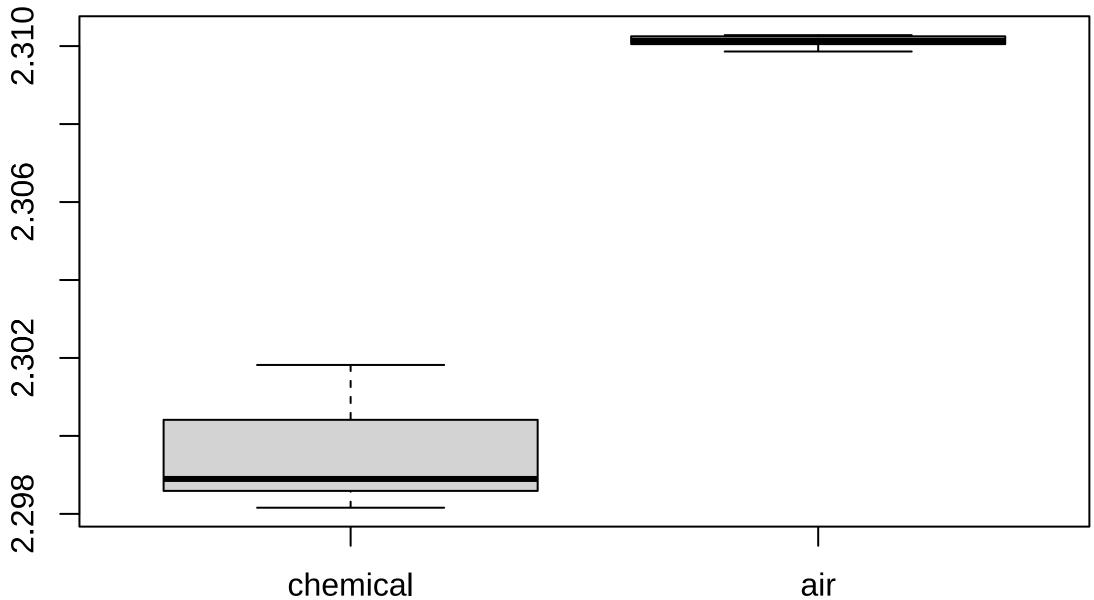
F) ANSWER: I would run a *two Sample t-test (unequal variance)* because we are comparing the mean between two groups (chemical vs air). The distribution of both groups are trending normal, although there are two slight outliers. And the variances are un-equal (p = 4.234e-06).

- H0 (null hypothesis): There is no difference in the mean density of nitrogen between the chemical compounds and air groups. (mu_chemical = mu_air)
- HA (alternative hypothesis): There is a difference in the mean density of nitrogen between the chemical compounds and air groups.

In conclusion, *there is a significant difference* in the mean density of nitrogen between the chemical compounds and air groups. The *p value is 8.594e-08*, which is < 0.05 and means we can reject our H0. Also, because our 95% CI (-0.01180801 - -0.00949949) does not include 0, then we know the mean of both groups are not equal to 0. The mean oral index for chemical compounds and air, are 2.29 and 2.31, respectively.




```
## [1] 3 7
```



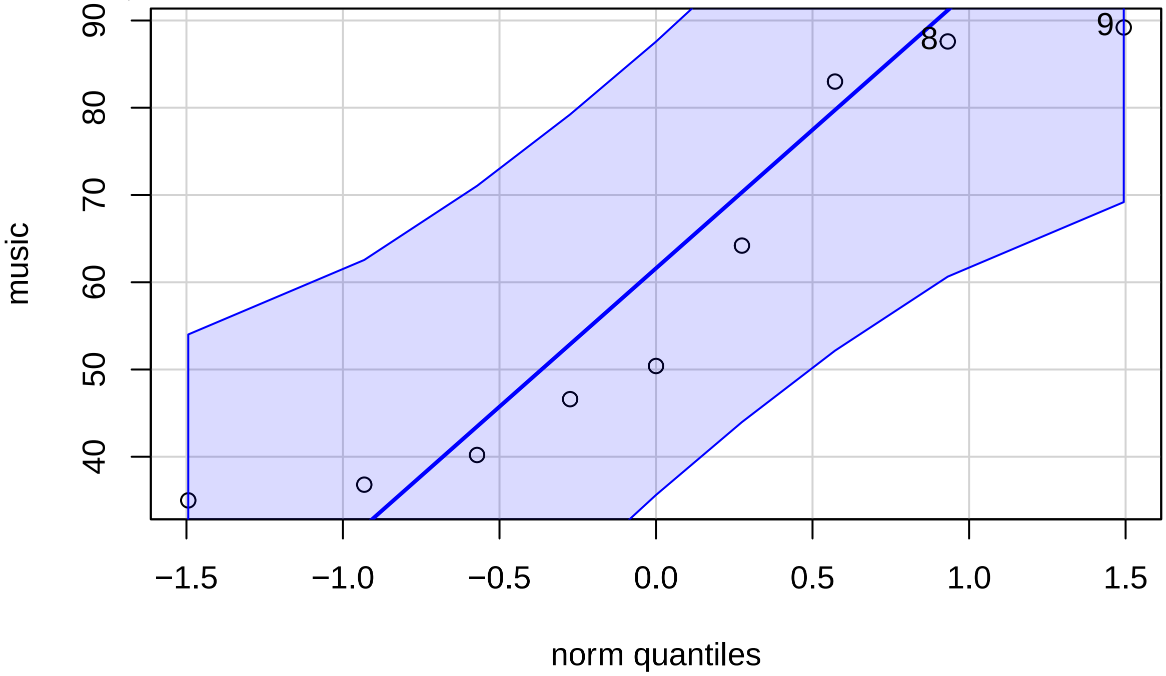

```
##
##  F test to compare two variances
##
## data:  chemical and air
## F = 94.846, num df = 7, denom df = 7, p-value = 4.234e-06
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##   18.98847 473.74564
## sample estimates:
## ratio of variances
##           94.84569
```

```
##
##  Welch Two Sample t-test
##
## data:  chemical and air
## t = -21.734, df = 7.1476, p-value = 8.594e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01180801 -0.00949949
## sample estimates:
## mean of x mean of y
##  2.299473  2.310126
```

G) ANSWER: I would run a *two Sample t-test (equal variance)* because we are comparing the mean between two groups (music vs no music). The distribution of both groups is trending normal. And the variances are equal (p = 0.3048).

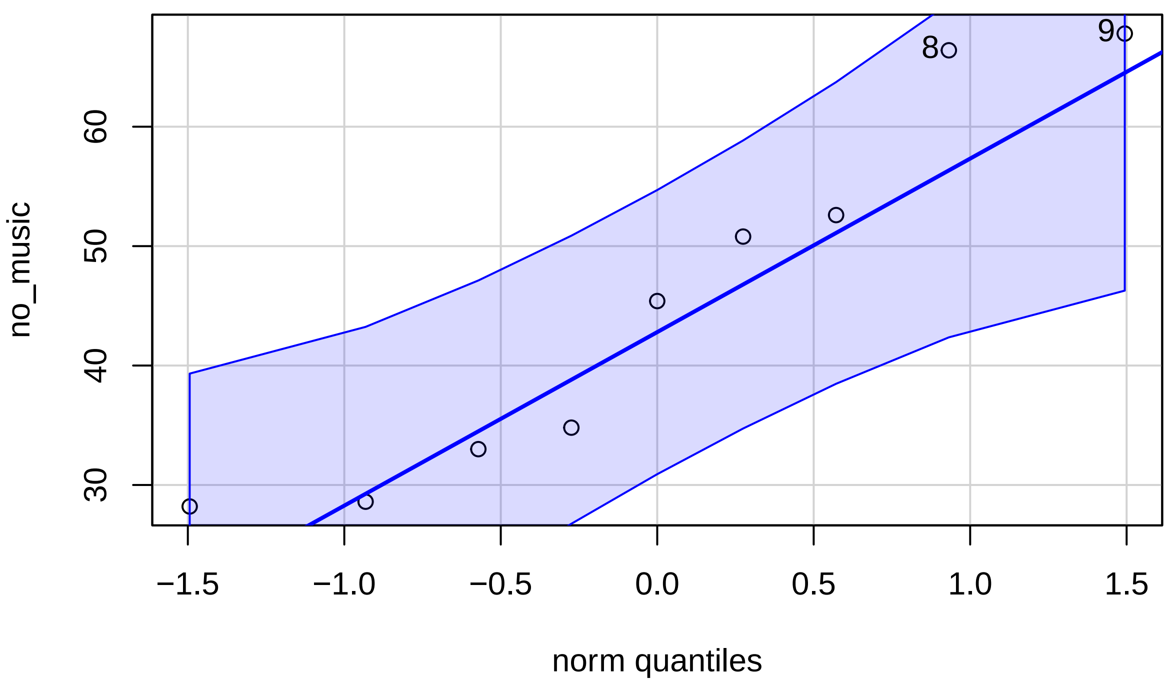
- H0 (null hypothesis): There is no difference in the mean productivity between the music and non music groups. (mu_music = mu_nomusic)
- HA (alternative hypothesis): There is a difference in the mean productivity between the music and non music groups.

In conclusion, *there is not a significant difference* in the mean oral hygiene index between the oxygen gel and placebo gel. The *p value is 0.1417*, which is > 0.05 and means we fail to reject our H0. Also, because our 95% CI (-5.175961 33.042627) does include 0. The mean oral index for oxygen gel and placebo gel, is 59.2 and 49.2,

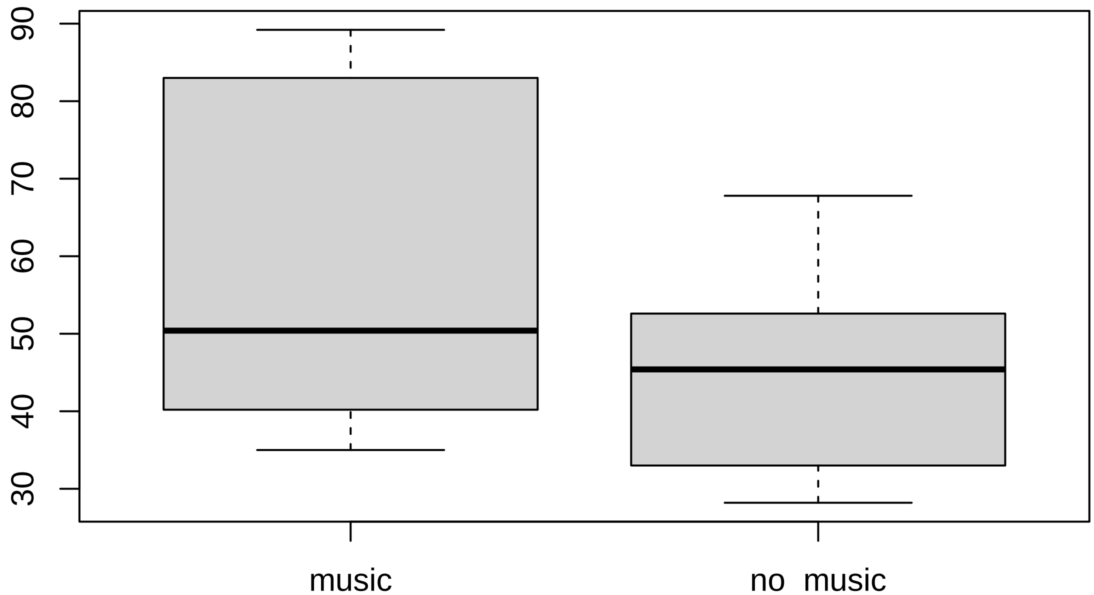


respectively.

```
## [1] 9 8
```




```
## [1] 9 8
```




```
##
##  F test to compare two variances
##
## data:  music and no_music
## F = 2.1322, num df = 8, denom df = 8, p-value = 0.3048
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.4809515 9.4525212
## sample estimates:
## ratio of variances
##           2.132183
```

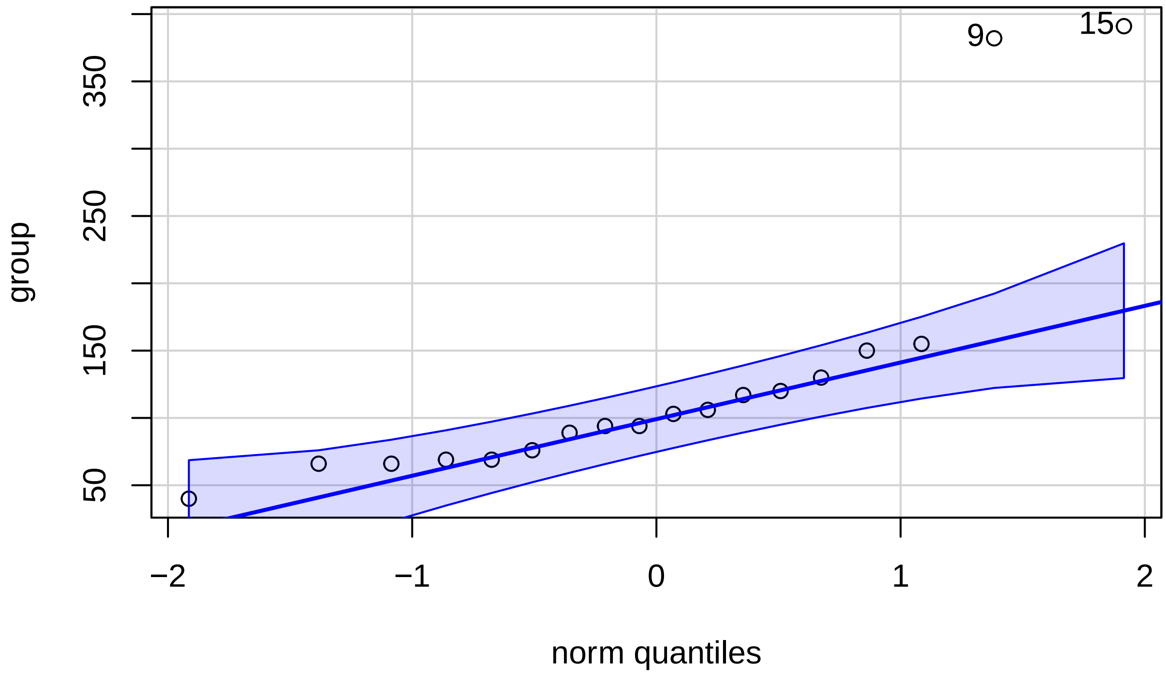
```
##
##  Two Sample t-test
##
## data:  music and no_music
## t = 1.5457, df = 16, p-value = 0.1417
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -5.175961 33.042627
## sample estimates:
## mean of x mean of y
##  59.22222  45.28889
```

H) ANSWER: I would run a *Wilcoxon rank sum test* because we are comparing the mean between two groups (male vs female) but their distributions are not normal with some outliers.

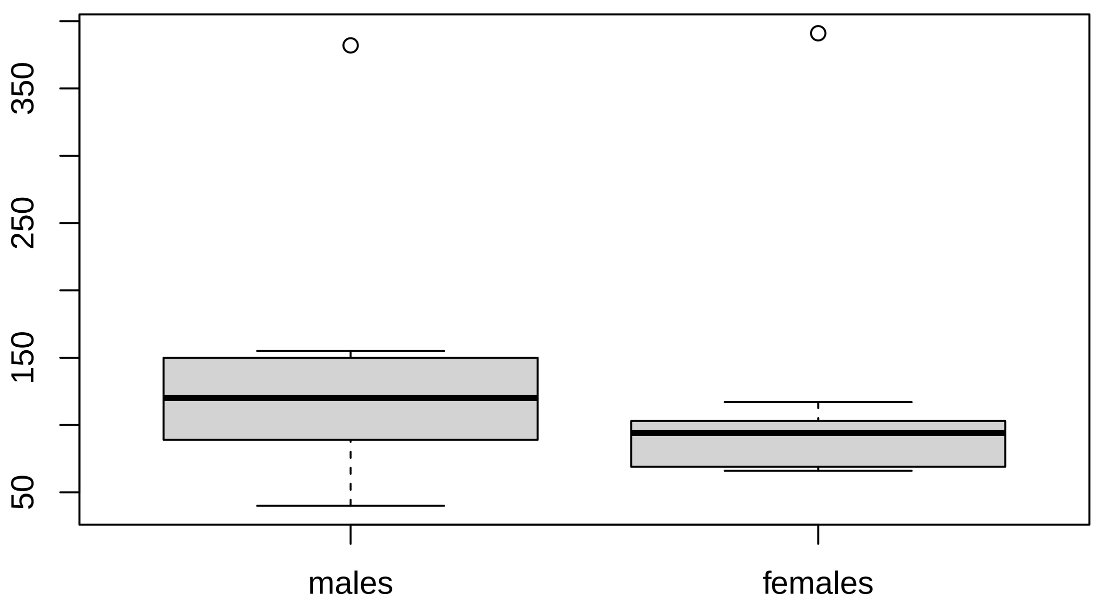
- H0 (null hypothesis): The true location shift is equal to 0 (median) for the male vs females groups. (M_d = 0)
- HA (alternative hypothesis): The true location shift is not equal to 0 (median) for the male vs females groups. (M_d /= 0)

In conclusion, *there is no significant difference* in the median between the potassium vs electrolytic oxygen groups. The *p value is 0.5287*, which is > 0.05 and means we fail to reject our H0.

```
## Warning in cbind(males, females): number of rows of result is not a multiple of
## vector length (arg 2)
```




```
## [1] 15  9
```

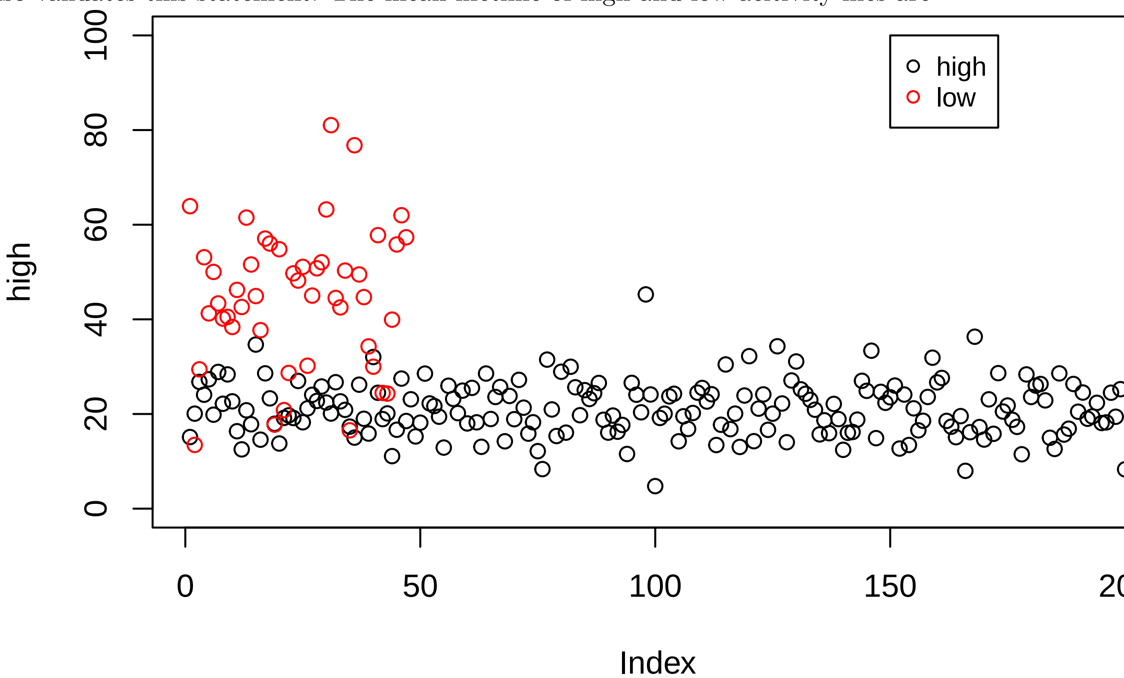



```
##
##  Wilcoxon rank sum exact test
##
## data:  males and females
## W = 33, p-value = 0.5287
## alternative hypothesis: true location shift is not equal to 0
```

I) ANSWER: I will use a two-sample t-test since I generated/simulated two populations (high and low) with a normal distribution.

- H0 (null hypothesis): The mean lifetime of the population of high activity flies is equal to the mean lifetime of the population of low flies. (mu_high = mu_low)
- HA (alternative hypothesis): The mean lifetime of the population of high activity flies is not equal to the mean lifetime of the population of low flies

In conclusion, with a *p-value = < 2.2e-16* ($p < 0.05$) I can *reject the null hypothesis* which means the mean lifetime of high activity and low activity flies are significantly different. The *95 CI (-33.69829 - -25.68222)* does not contain 0 which also validates this statement. The mean lifetime of high and low acitivity flies are *20.9* and *50.6*, respectively.

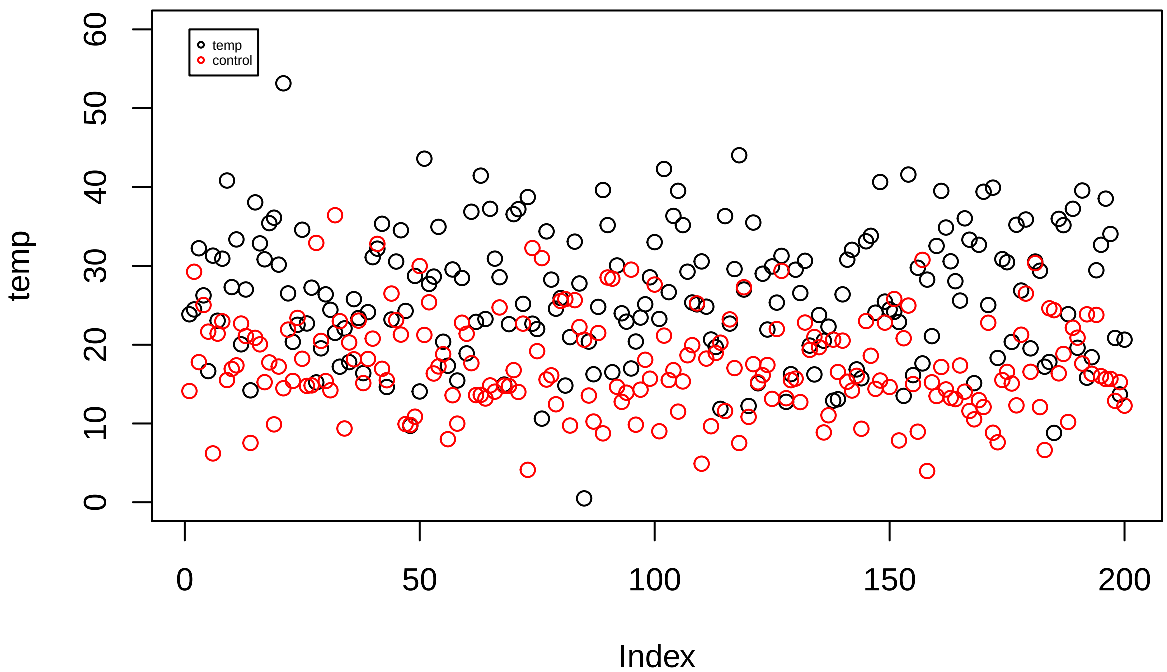



```
##
##  Welch Two Sample t-test
##
## data:  high and low
## t = -10.96, df = 49.37, p-value = 7.972e-15
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -28.30172 -19.53239
## sample estimates:
## mean of x mean of y
##  21.09437  45.01142
```

J) ANSWER: I will use a two-sample t-test since I generated/simulated two populations (temp and control) with a normal distribution.

- H0 (null hypothesis): The mean lifetime of the population of temperature flies is equal to the mean lifetime of the population of control flies. (mu_temp = mu_control)
- HA (alternative hypothesis): The mean lifetime of the population of temperature flies is not equal to the mean lifetime of the population of control flies.

In conclusion, with a *p-value = < 2.2e-16* ($p < 0.05$) I can *reject the null hypothesis* which means the mean lifetime of temperature and control flies are significantly different. The *95 CI (6.466135 - 9.220216)* does not contain 0 which also validates this statement. The mean lifetime of temperature and control flies are *25.4* and *17.6*, respectively.




```
##
##  Welch Two Sample t-test
##
## data:  temp and control
## t = 12.192, df = 366.02, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   7.462622 10.332781
## sample estimates:
## mean of x mean of y
##  26.40069  17.50299
```

Question 2 ANSWER: If I had to apply one of the above simple tests from question 1, I would conduct a *two sample t-test* given normal distribution and equal variance because you are comparing the mean cortex weight between two groups - the treatment and control. However a two sample t-test would not include the factor, experiment number, and because the experiementers wanted to repeat their experiments to test an effect of toys on cortex weight, a two sample t-test would be insufficient.

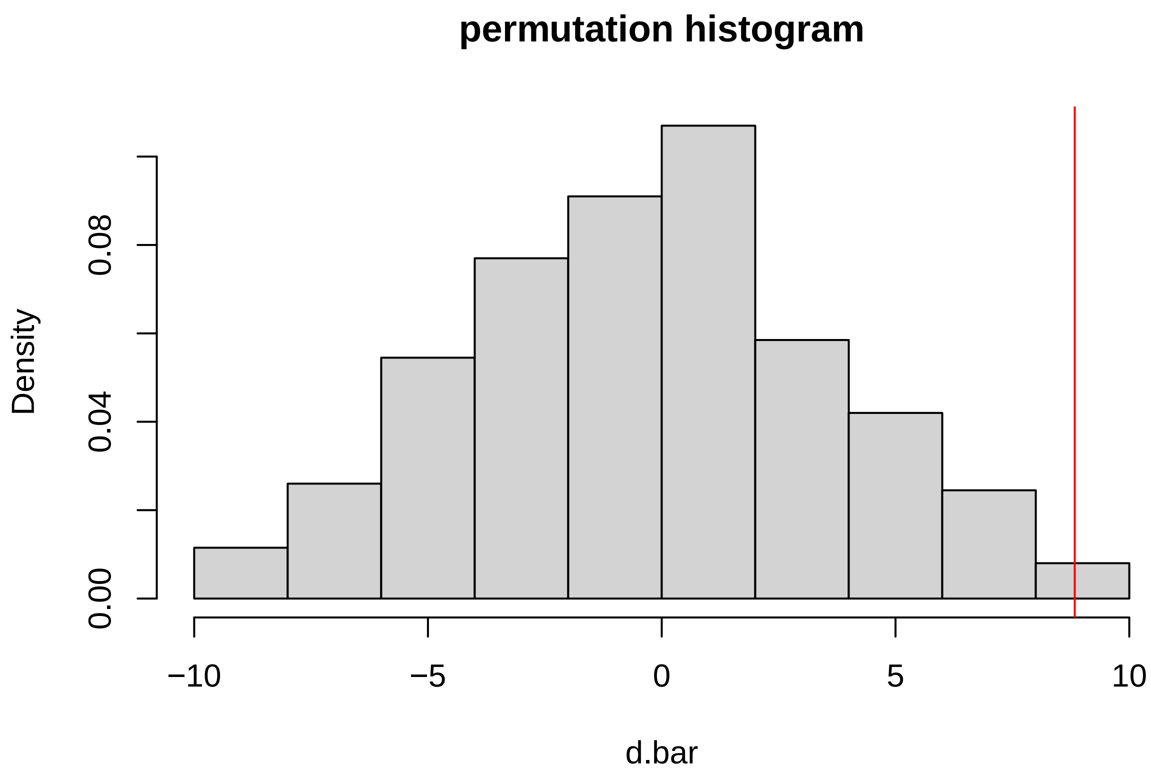
I think a more reasonable test to conduct would be a *linear regression model* with the dependent variable (Y = cortex weight) and independent variables ($X1$ = toy allowance and $X1$ = experiment #) to test the association between toys and cortex development. The variable toy allowance would be a factor with two levels - given toys (treatment) or not given toys (control). However other variables, or confounders, in this study could be individual rat weight, overall rat health, etc.

Question 3 ANSWER: The permutation test was ran with the below hypotheses and the results are summarized with a histogram.

- H0 (null hypothesis): The amount of hemogloblin associated with honey group is equal to amount of hemogloblin associated with non honey group (mu_difference = 0)
- HA (alternative hypothesis): The amount of hemogloblin associated with honey group is not equal to amount of hemogloblin associated with non honey group.

The *bootstrap confidence intervals* for the difference of population means are 95% CI: *5 - 13*.

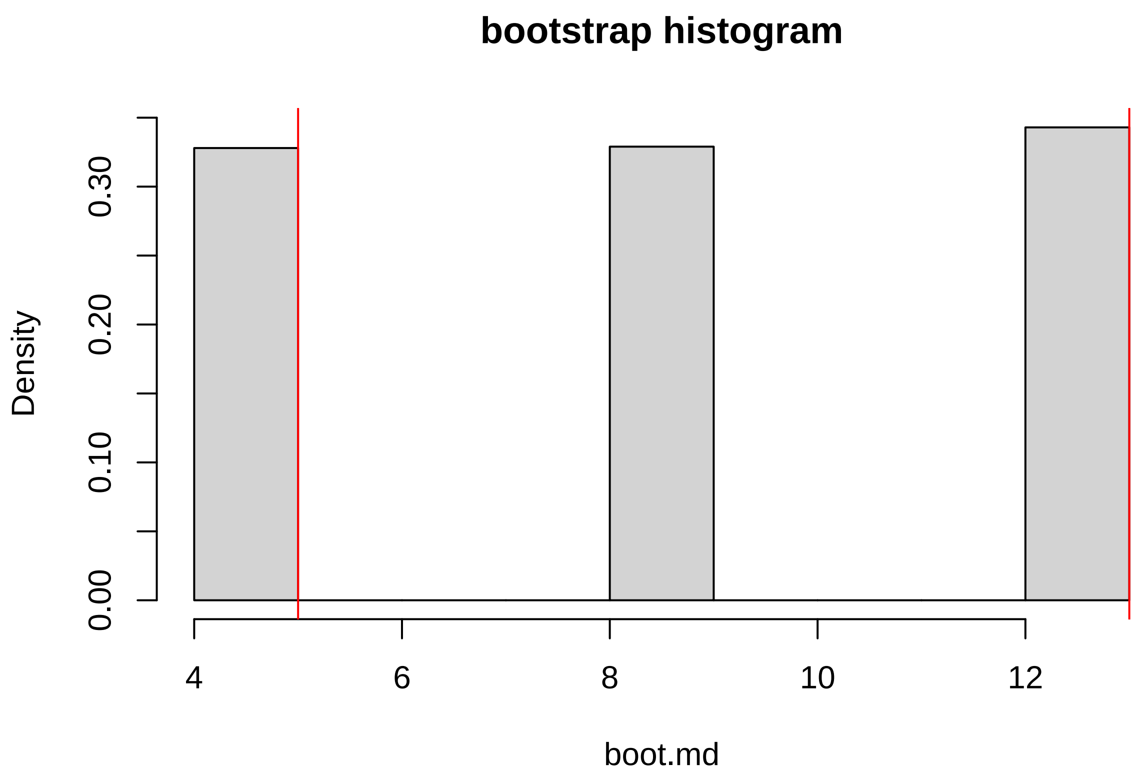
```
## Loading required package: survival
```




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

```
## [1] 10.99321
```

```
##   25% 97.5%
##     5    13
```

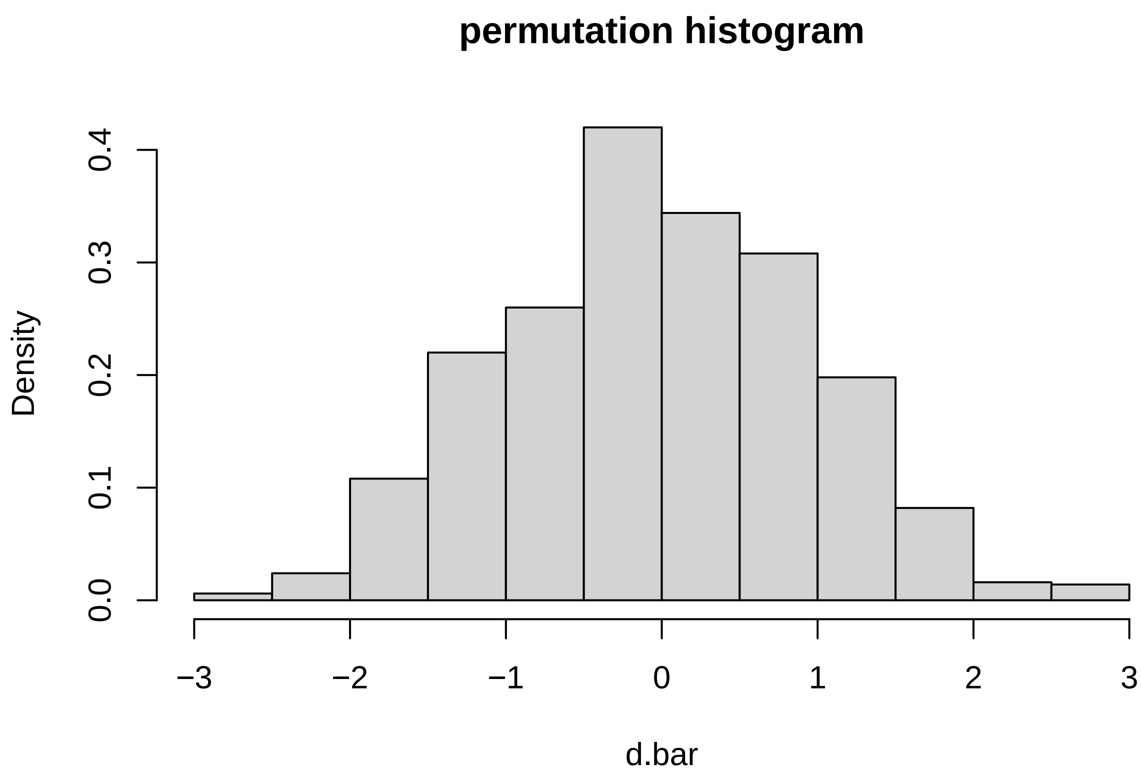


Question 4 ANSWER: The permutation test was ran with the below hypotheses and the results are summarized with a histogram.

- H0 (null hypothesis): The amount of oral hygiene index associated with oxygen group is equal to amount of oral hygiene index associated with placebo group (mu_difference = 0).
- HA (alternative hypothesis): The amount of oral hygiene index associated with oxygen group is not equal to amount of oral hygiene index associated with placebo group

The *bootstrap confidence intervals* for the difference of population means are 95% CI: *3.5 - 6*.

```
## Warning in oxygen - placebo: longer object length is not a multiple of shorter
## object length
```

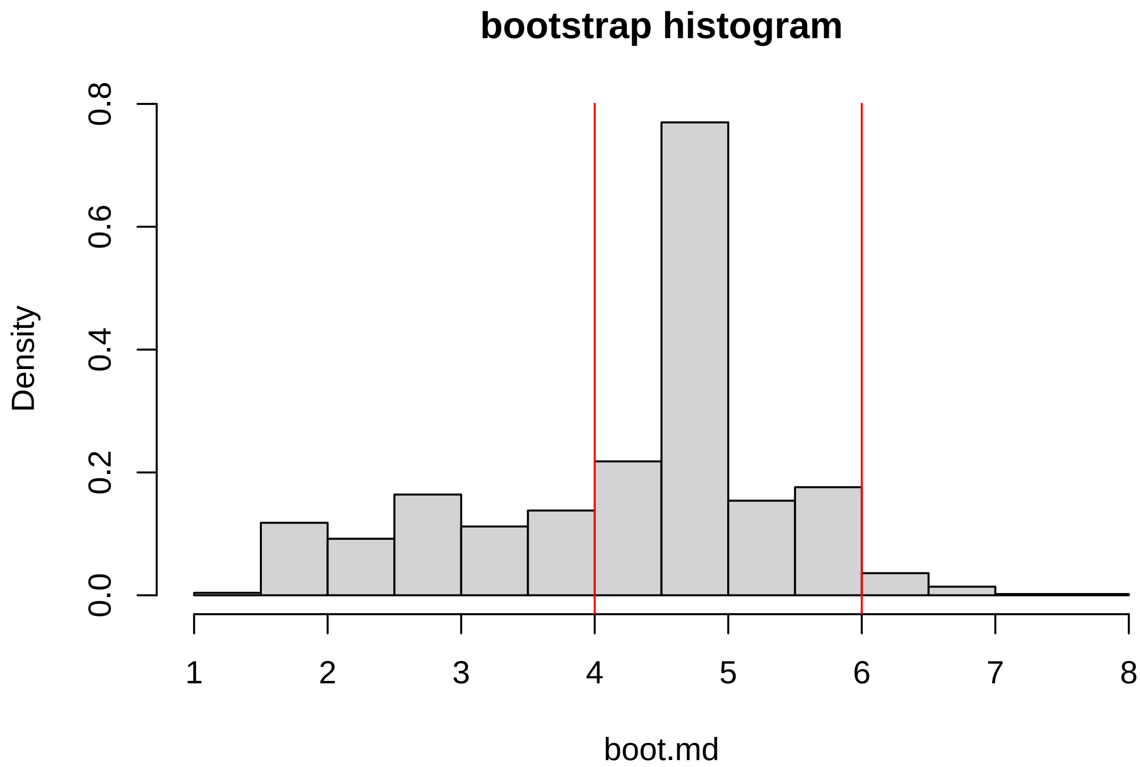



```
## [1] 0
```

```
## Warning in oxygen - placebo: longer object length is not a multiple of shorter
## object length
```

```
## [1] 1.341337
```

```
##   25% 97.5%
##   4.0   6.5
```



Question 5 ANSWER: The **t-test** is used to compare *means*, whereas the **wilcoxon's** test the *ordering* of the data. The **t-test** is a *parametric* test whereas the **wilcoxon's** is a *non-parametric* test, meaning the **t-test** follows *more assumptions* than the wilcoxon's (i.e. t-test assumptions assumes normal distribution and equal variance). Because the **t-test** is the parametric version and follows strict assumptions it is more conservative and a stronger test to perform for significance when comparing two groups. With my gaussian distribution, simulations we can see that the t-test (black line) has a higher power than the wilcoxon test (red line), although the lines were very close. However, with my non-gaussian distribution (poisson), there is a larger difference between t-test and wilcoxon with the t-test having higher power than wilcoxon.

*note: I try running the simulation for rpois but the output was 20+ pages long, so I inserted the photo of the simulation to reduce the number of pages in my knitted document

```
##         [,1]  [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]  [,9] [,10]
## powert 0.037 0.097 0.254 0.436 0.691 0.869 0.953 0.988 0.999     1
## powerw 0.038 0.095 0.242 0.429 0.666 0.852 0.944 0.985 0.997     1
```

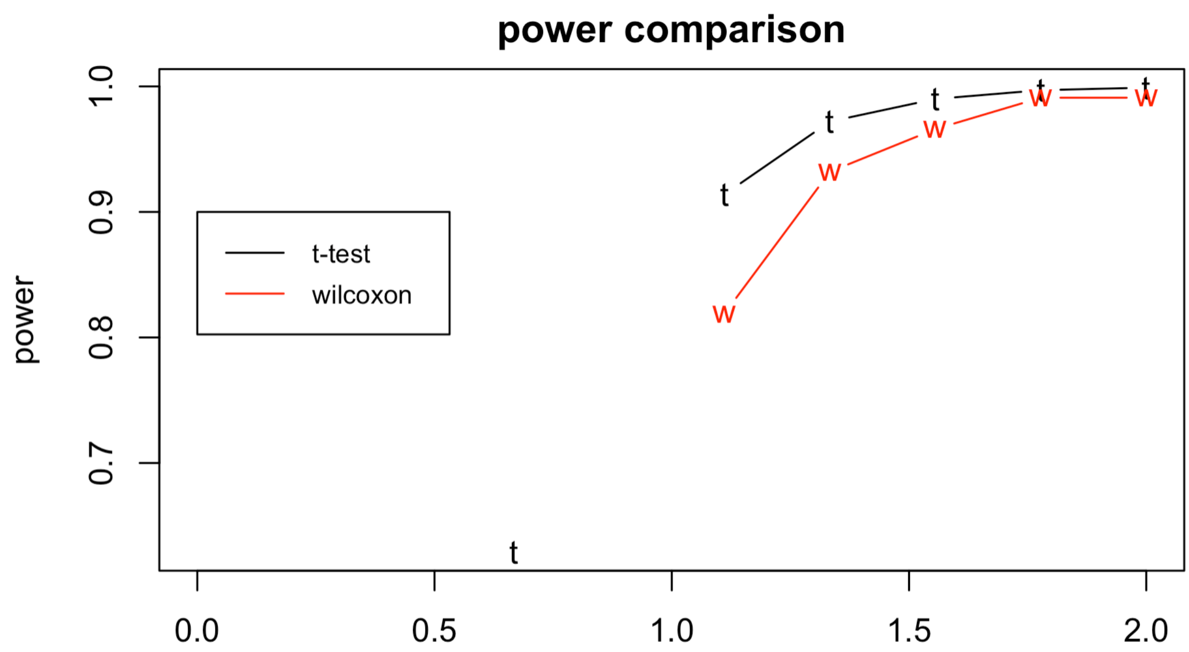
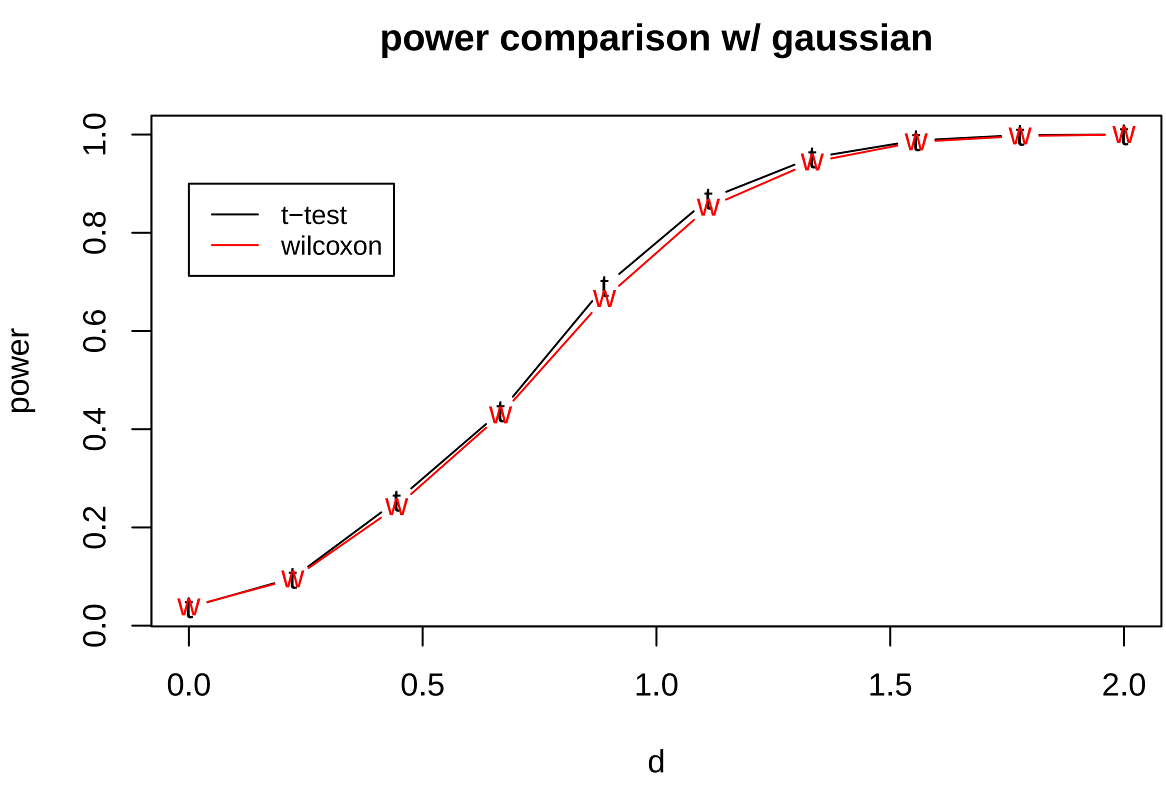


Figure 1: poisson distribution power graph



Question 6 ANSWER: To compare two proportions of recieving dealth penalty from two independent groups (white and black), we will conduct a 2-sample test for equality of proportions with continuity correction.

- H0 (null hypothesis): The proportion of white population is equal to proportion of black population that would receive the death penalty (p_white = p_black).
- HA (alternative hypothesis): The proportion of white population is not equal to proportion of black population that would receive the death penalty.

In conclusion, with a *p-value = 0.7689* ($p > 0.05$), we fail to reject our null hypothesis meaning the proportion of white pop. *is equal* to the proportion of black pop. that receives the death penalty with a *95 CI of -0.05791211 - 0.09059283*. The *estimates of proportion for white and black populations is 0.1187500 and 0.1024096*, respectively. The Pearson's chi-squared test statistic is *X-squared = 0.086343*.

```
##
##  2-sample test for equality of proportions with continuity correction
##
## data:  x out of n
## X-squared = 0.086343, df = 1, p-value = 0.7689
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.05791211  0.09059283
## sample estimates:
##    prop 1    prop 2
## 0.1187500 0.1024096
```

Question 7 ANSWER: To test of goodness of fit establishes whether or not an observed frequency distribution differs from a theoretical distribution. I have written a chi-square test for given probabilities for two scenarios: tomatoes and birth hour.

For **tomatoes**,

- H0 (null hypothesis): The observed frequency of tomato phenologies is the same as a theoretical distribution
- HA (alternative hypothesis): The observed frequency of tomato phenologies is not the same as a theoretical distribution

In conclusion, my **p-value = 0.6895** ($p > 0.05$), so we fail to reject the null hypothesis which means tomoato phenologies are the same as the theoretical distribution. AKA, the **tomato data is consistent with Mendel's law**.

For **birth hour**,

- H0 (null hypothesis): The observed frequency of birth hour is the same as a theoretical distribution
- HA (alternative hypothesis): The observed frequency of birth hour is not the same as a theoretical distribution

In conclusion, my **p-value = < 2.2e-16** ($p < 0.05$), so we reject the null hypothesis which means birth hours are not the same as the theoretical distribution. AKA, the **birth hour data is not consistent throughout the night** and is more likely to commence very late at night.

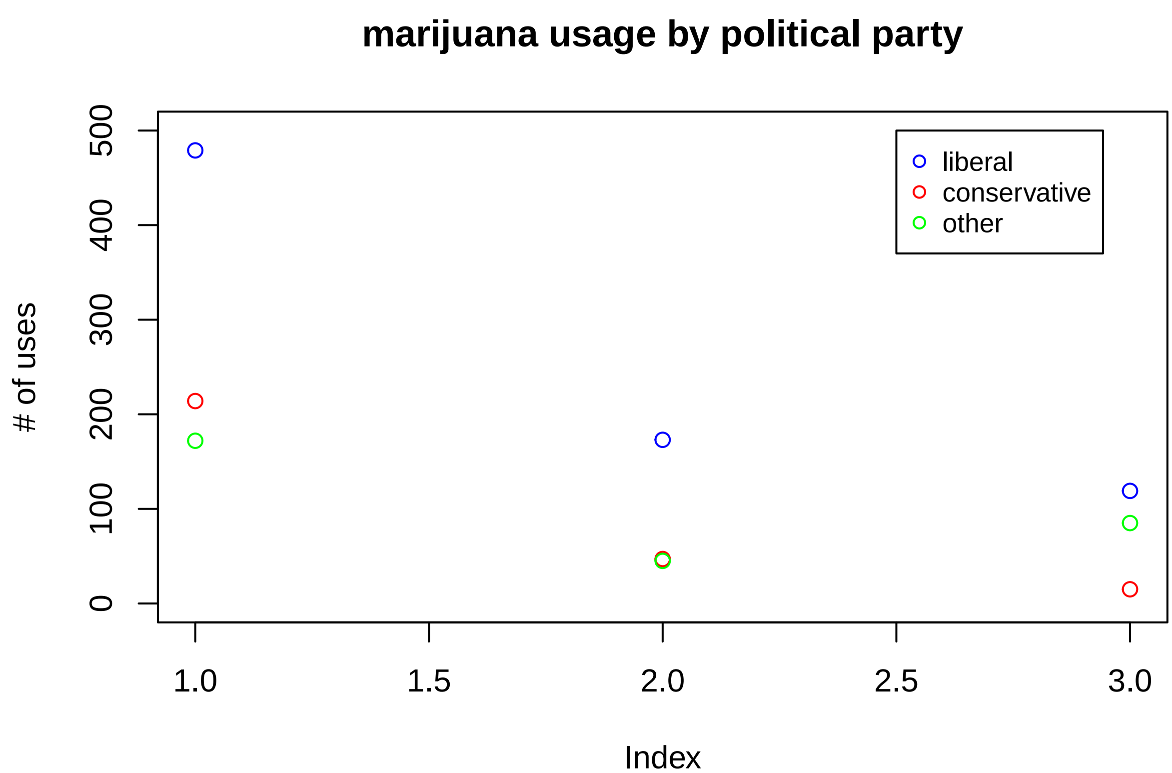
```
##
##  Chi-squared test for given probabilities
##
## data:  x
## X-squared = 1.4687, df = 3, p-value = 0.6895
```

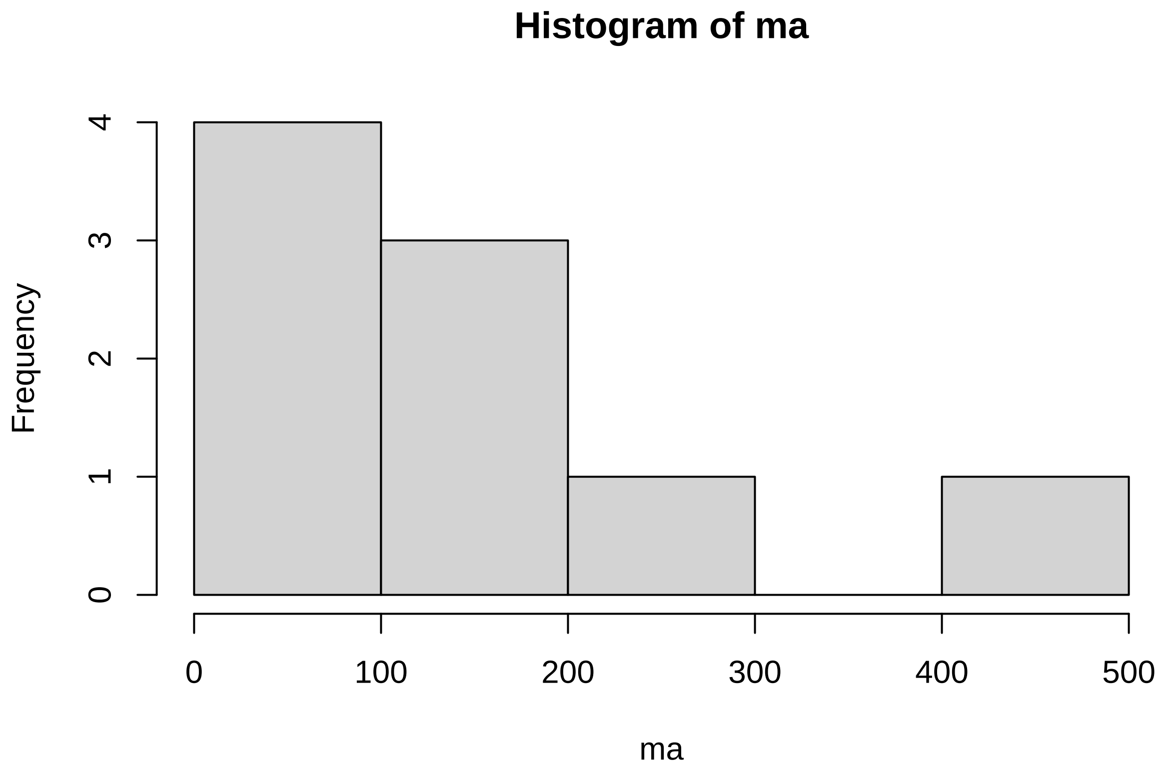
```
## [1] 4

##
##  Chi-squared test for given probabilities
##
## data:  x
## X-squared = 162.78, df = 23, p-value < 2.2e-16
```

Question 8 ANSWER: I will use a *contingency table* to answer questions regarding the independence of both political views and marijuana usage. To test for independency I will test that an individual's category with respect to variable (political party) is independent of the category with respect to variable (marijuana usage).

- H0 (null hypothesis): In terms of the number of times individuals used marijuana, the effect of usage level is independent of the effect of political views. (p_ij = p_i + p + j, where j = usage level and i = political party).
- HA (alternative hypothesis): In terms of the number of times individuals used marijuana, the effect of usage level is not independent of the effect of political views.

In conclusion, our p-value from a pearson's chi-square test is $p = 3.043e\text{-}13$ ($p < 0.05$), meaning we reject our null hypothesis and can say signficantly that political party and marijuana usage level *are not two independent variables.*






```
##
##  Pearson's Chi-squared test
##
## data:  ma
## X-squared = 64.654, df = 4, p-value = 3.043e-13
```

Question 9 ANSWER: To get inference on a proportion, I will use an *exact binomial test* where *success is defined by finding a defective battery (n = 124)* and failure (n = 26) is defined by finding a non-defective battery in a random sample of 150 flashlights.

- H0 (null hypothesis): the true probability of success (selecting defective battery) is equal to .75
- HA (alternative hypothesis): the true probability of success (selecting defective battery) is not equal to .75

In conclusion, with a *p-value* $<$ *2.2e-16* ($p < 0.05$) we can reject the null hypothesis and state that the true probability of selecting a defective battery (success) is not equal to .75 and the true probability of selecting a defective battery is *0.4525547* with the 95% CI 0.3925903-0.5135564.

```
##
##  Exact binomial test
##
## data:  c(success, total)
## number of successes = 124, number of trials = 274, p-value < 2.2e-16
## alternative hypothesis: true probability of success is not equal to 0.75
## 95 percent confidence interval:
##  0.3925903 0.5135564
## sample estimates:
## probability of success
##              0.4525547
```

Question 10 ANSWER: The populations (present/absent) and each population is divided into the same categories of insulin dose (<.25, .25-.49, .50-.74, .75-99, >=1). I will run a two-way tables to calculate the pearson's chi-squared test.

- H0 (null hypothesis): the proportion of individuals who belong to different categories of insulin dose is the same as individuals part of present and absent variables.
- HA (alternative hypothesis): the proportion of individuals who belong to different categories of insulin dose is not the same as individuals part of present and absent variables.

In conclusion, with a **p-value = 0.01182** ($p < 0.05$) we reject the null hypothesis and state that the presence/absence of **hypoglycemia is not independent of insulin dosage**.

```
##
##  Pearson's Chi-squared test
##
## data:  x
## X-squared = 12.892, df = 4, p-value = 0.01182
```