STAR511: HW 4

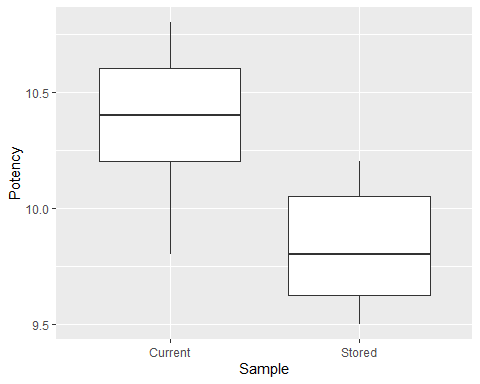
Megan Sears

# Q1 \*\*

A sample size of 15 per group is required to achieve a 90% power using alpha = 0.05.

##   
## Two-sample t test power calculation   
##   
## n = 14.48098  
## delta = 1.5  
## sd = 1.2  
## sig.level = 0.05  
## power = 0.9  
## alternative = two.sided  
##   
## NOTE: n is number in \*each\* group

# Q2



# Q3

| Sample | Mean | StDev |
| --- | --- | --- |
| Current | 10.37 | 0.32 |
| Stored | 9.83 | 0.24 |

# Q4

Since the max standard deviation (Current samples) and min standard deviation (Stored samples) ratio is less than 2, we can assume “equal” variances and use a pooled two-sample t-test. Since the sample sizes are equal, the effect of unequal variances on the t-test is minimal. However, pooled test is preferred since the standard deviation ratio is less than 2.

## [1] 1.333333

# Q5 \*\*

##   
## Two Sample t-test  
##   
## data: Potency by Sample  
## t = 4.2368, df = 18, p-value = 0.0004959  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 0.2722297 0.8077703  
## sample estimates:  
## mean in group Current mean in group Stored   
## 10.37 9.83

## t   
## 4.24

## [1] 0.0004959478

The null hypothesis is mu1 - mu2 (Current standard deviation - Stored standard deviation) equal zero and the alternative hypothesis is that it does not equal 0.

The test statistics is ts. The p-value is pval.  
Based on the p-value, XXXXXXXXXXXXX

# Q6

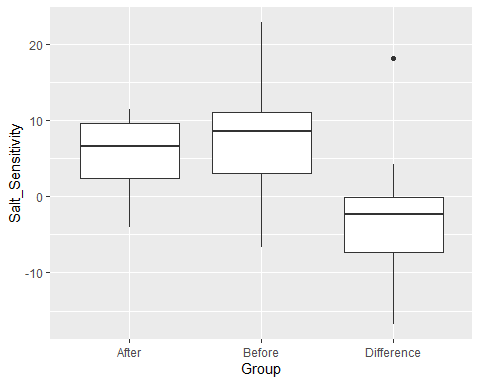
The p-value is the probability of observing the test statistic (4.24) as or more supportive of the alternative hypothesis (u1 - u2 does not equal 0) than the actual observed value, given the null hypothesis is true. In other words, the probability;ity that if H0: u1 - u2 = 0 was true, a new sample of data would give a test statistic at least as large as 4.24 in absolute magnitude.

# Q7

The 95% CI for the difference between the two population means is (0.27, 0.81). Since the CI does not include 0, we have evidence that there is a difference between the population means.

##   
## Two Sample t-test  
##   
## data: Potency by Sample  
## t = 4.2368, df = 18, p-value = 0.0004959  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 0.2722297 0.8077703  
## sample estimates:  
## mean in group Current mean in group Stored   
## 10.37 9.83

# Q8



# Q9

A. The statistical hypotheses are the null hypothesis is the mean of the differences are equal to 0 and the alternative hypothesis is they are not equal to 0.

##   
## Paired t-test  
##   
## data: salt$After and salt$Before  
## t = -0.86098, df = 9, p-value = 0.4116  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -9.373261 4.205261  
## sample estimates:  
## mean of the differences   
## -2.584

## t   
## -0.8609794

## [1] 0.4115991

B. T-test output above. The test statistic is ts and the p-value is pval.

C. The confidence interval is -9.37 to 4.21.

# Q10

##   
## Exact Wilcoxon-Pratt Signed-Rank Test  
##   
## data: y by x (pos, neg)   
## stratified by block  
## Z = -1.1722, p-value = 0.2754  
## alternative hypothesis: true mu is not equal to 0

#Q11

The p-value (0.41) is greater than alpha, which indicates failing to reject the null hypothesis (difference of means = 0). Therefore, we did not find evidence that there is a difference between means. Additionally, the confidence interval of -9.37 to 4.21 includes 0, indicating that we can conclude that there is not a difference between means. \*\*\*

## Continue in this manner for remainder of questions

# Appendix

#Retain (and do not edit) this code chunk!!!  
library(knitr)  
knitr::opts\_chunk$set(echo = FALSE)  
knitr::opts\_chunk$set(message = FALSE)  
  
library(tidyverse)  
library(here)  
library(coin)  
  
#Q1  
pool\_sd <- 1.2  
delta <- 1.5  
alpha <- 0.05  
power <- 0.9  
  
power.t.test(delta = delta, sd = pool\_sd,   
 sig.level = alpha, power = power,   
 type = 'two.sample', alternative = 'two.sided')  
  
  
#Q2  
potency <- read\_csv(here('./HW4/ex6-59.txt'), quote = "'")  
  
pot\_long <- potency %>%  
 rename(Current = Sample1,  
 Stored = Sample2) %>%  
 pivot\_longer(everything(), names\_to = 'Sample',  
 values\_to = 'Potency')  
  
boxplot <- ggplot(pot\_long, aes(x=Sample, y=Potency)) + geom\_boxplot()  
  
boxplot  
  
#Q3  
pot\_long\_summ <- pot\_long %>%  
 group\_by(Sample) %>%  
 summarize(Mean = mean(Potency),  
 StDev = round(sd(Potency),2))  
  
kable(pot\_long\_summ)  
  
#Q4  
sd\_ratio <- 0.32/0.24  
sd\_ratio  
  
#Q5  
q5test <- t.test(data = pot\_long, Potency ~ Sample, var.equal = T)  
q5test  
  
ts <- round(q5test$statistic, 2)  
ts  
  
pval <- q5test$p.value ## do I need to multiple by 2, how to use to reject or ftr the Ho  
pval  
  
#Q7  
q5test  
  
#Q8  
salt <- read\_csv(here('./HW4/ex6-28.txt'), quote = "'")  
  
salt\_long <- salt %>%  
 mutate(Difference = After - Before) %>%  
 pivot\_longer(everything(), names\_to = 'Group',  
 values\_to = 'Salt\_Sensitivity')  
  
salt\_boxp <- ggplot(salt\_long, aes(x=Group, y=Salt\_Sensitivity)) +  
 geom\_boxplot()  
  
salt\_boxp  
  
#Q9  
salt <- salt %>%  
 mutate(Difference = After - Before)  
  
q9 <- t.test(salt$After, salt$Before, paired = T)  
q9  
  
ts <- q9$statistic  
ts  
  
pval <- q9$p.value  
pval  
  
#Q10  
wilcoxsign\_test(After ~ Before, data = salt, distribution = 'exact')