

DATA 602 HW 3

Kane Smith

2022-10-09

Contents

Question 1	1
Question 2	2
Question 3	4
Question 4	7
Question 5	8
Question 6	9
Question 7	10
Question 8	12
Question 9	14
Question 10	15
Question 11	17

Setting seed so that my explanations match my outputs from bootstrapping and simulations using random sampling:

```
set.seed(555)
```

Question 1

a)

```
smoker_no <- subset(NCBirths2004, Smoker == "No")
smoker_yes <- subset(NCBirths2004, Smoker == "Yes")
head(NCBirths2004)
```

##	ID	MothersAge	Tobacco	Alcohol	Gender	Weight	Gestation	Smoker
## 1	1	30-34	No	No	Male	3827	40	No
## 2	2	30-34	No	No	Male	3629	38	No
## 3	3	35-39	No	No	Female	3062	37	No
## 4	4	20-24	No	No	Female	3430	39	No
## 5	5	25-29	No	No	Male	3827	38	No
## 6	6	35-39	No	No	Female	3119	39	No

We will create 2 vectors, one with the baby weights of non smokers and one with the baby weights of smokers. We then sample from vectors these to calculate our sample means for each category respectively.

```
smoker_vec <- smoker_yes$Weight
non_smoker_vec <- smoker_no$Weight
q1a_bootstrap = do(1000)*(mean(resample(non_smoker_vec)) - mean(resample(smoker_vec)))
```

b)

```
qdata(~result, c(0.025, 0.975), data=q1a_bootstrap)
```

```
##      2.5%    97.5%
## 115.6006 321.9095
```

Our 95% confidence interval from bootstrapping is **(115.6006, 321.9095)**. This means we can say with 95% confidence that the on average the difference between the population means of baby weights from non-smokers and smokers is between 115.6006 and 321.9095.

c)

```
t.test(~Weight | Smoker, data = NCBirths2004, var.equal = FALSE )$conf
```

```
## [1] 112.3161 317.6881
## attr(,"conf.level")
## [1] 0.95
```

95% confidence interval from t.test: **(112.3161, 317.6881)**. This means we can say with 95% confidence that the on average the difference between the population means of baby weights from non-smokers and smokers is between 112.3161 and 317.6881.

d)

We can say with 95% confidence that babies born to birth mothers who did not smoke during pregnancy weigh more on average than babies born to birth mothers who did smoke during pregnancy. This is because both 95% confidence intervals calculated for the mean weight of non-smoke - mean weight of smoke in b) and c) is positive.

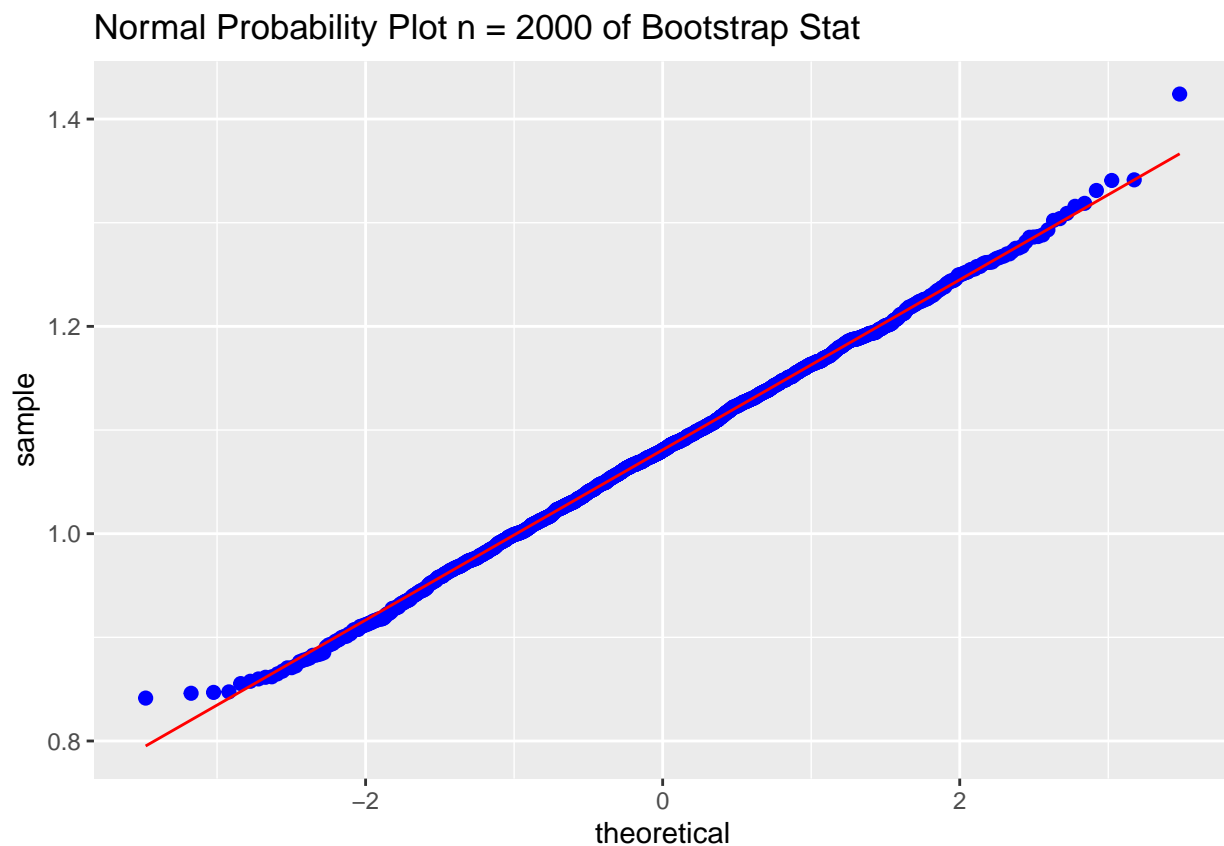
Question 2

a)

```
q2a_bootstrap = do(2000)*(sd(resample(smoker_vec)) / sd(resample(non_smoker_vec)))
```

b)

```
ggplot(q2a_bootstrap, aes(sample = result)) + stat_qq(size=2, col="blue") + stat_qqline(col="red") + ggtitle("Normal Probability Plot n = 2000 of Bootstrap Stat")
```



Looking at the normal probability plot, the ratio of the sample standard deviations does appear to follow an approx. normal distribution. The majority of the points fall on, or very close to the line.

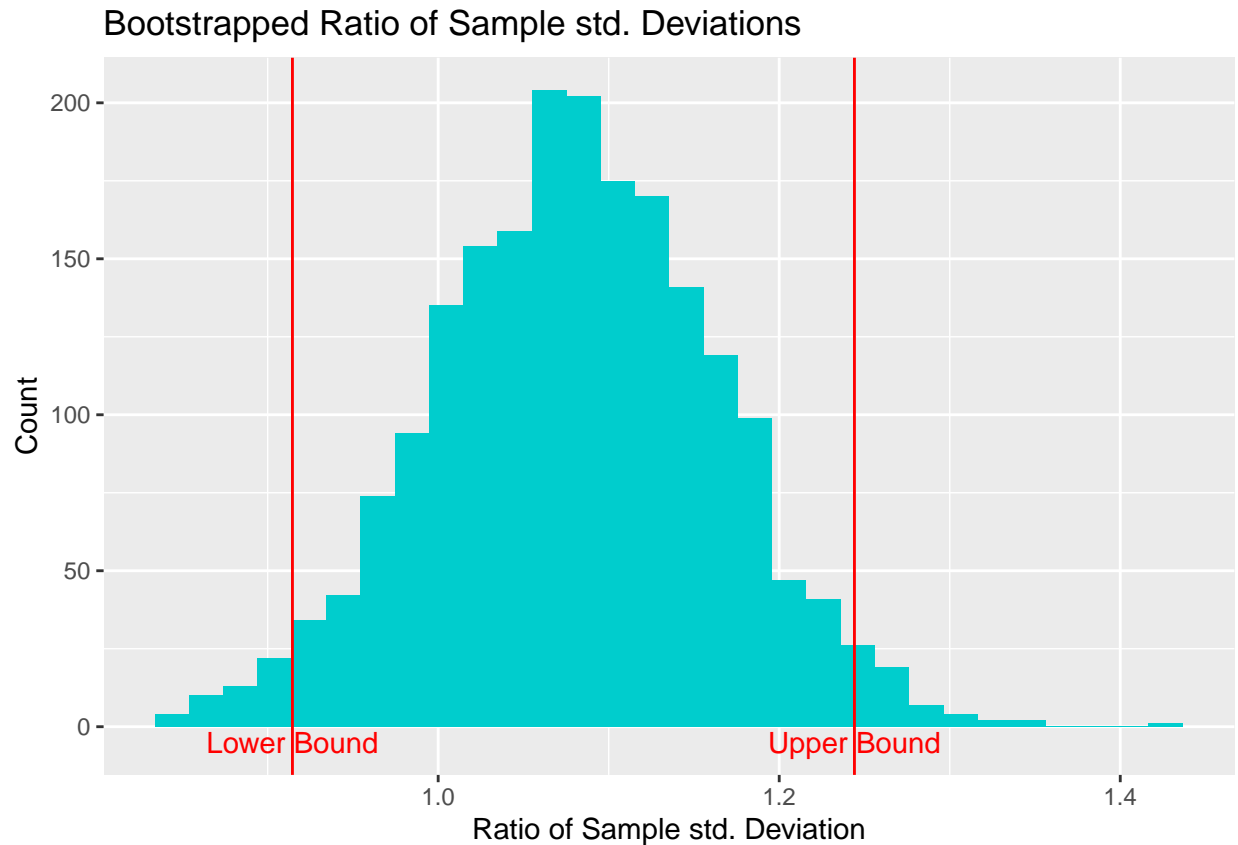
c)

```
qdata(~result, c(0.025, 0.975), data=q2a_bootstrap)
```

```
##      2.5%      97.5%
## 0.9145372 1.2441046
```

```
ggplot(q2a_bootstrap, aes(x = result)) + geom_histogram(fill = "cyan3") + ggtitle("Bootstrapped Ratio of Standard Deviations")
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



The 95% bootstrap confidence interval is **(0.9145, 1.2441)**. This means we can say with 95% confidence that on average the ratio between the standard deviation of baby weights from non-smokers and smokers is between 0.9145 and 1.2441.

d)

A practical result of considering the ratio of sample standard deviation of smoker's baby weights to non-smoker's baby weights would be to see if on average we observe more variability in smoker's baby weights compared to non-smoker's baby weights. Based on our 95% confidence interval calculated in c), there is no difference in the variability of babies weight based on what mother they were born from. This is because the confidence intervals contain 1.

Question 3

```
q3_vec <- c(1.2,1.1,1.0,1.0,1.1,1.0,1.0,1.0,0.9,1.1,1.1,1.2,1.0,1.1,1.0,1.1,1.0,0.9,1.0,1.0,1.1,1.0,1.0)
q3_vec1<- rep("ppm", 31)
q3.df <- data.frame(q3_vec1, q3_vec)
```

a)

Action level is set to mercury levels **over** 1 ppm.

Null hypothesis: $H_0 : \mu = 1.0$

Alternative hypothesis: $H_A : \mu > 1.0$

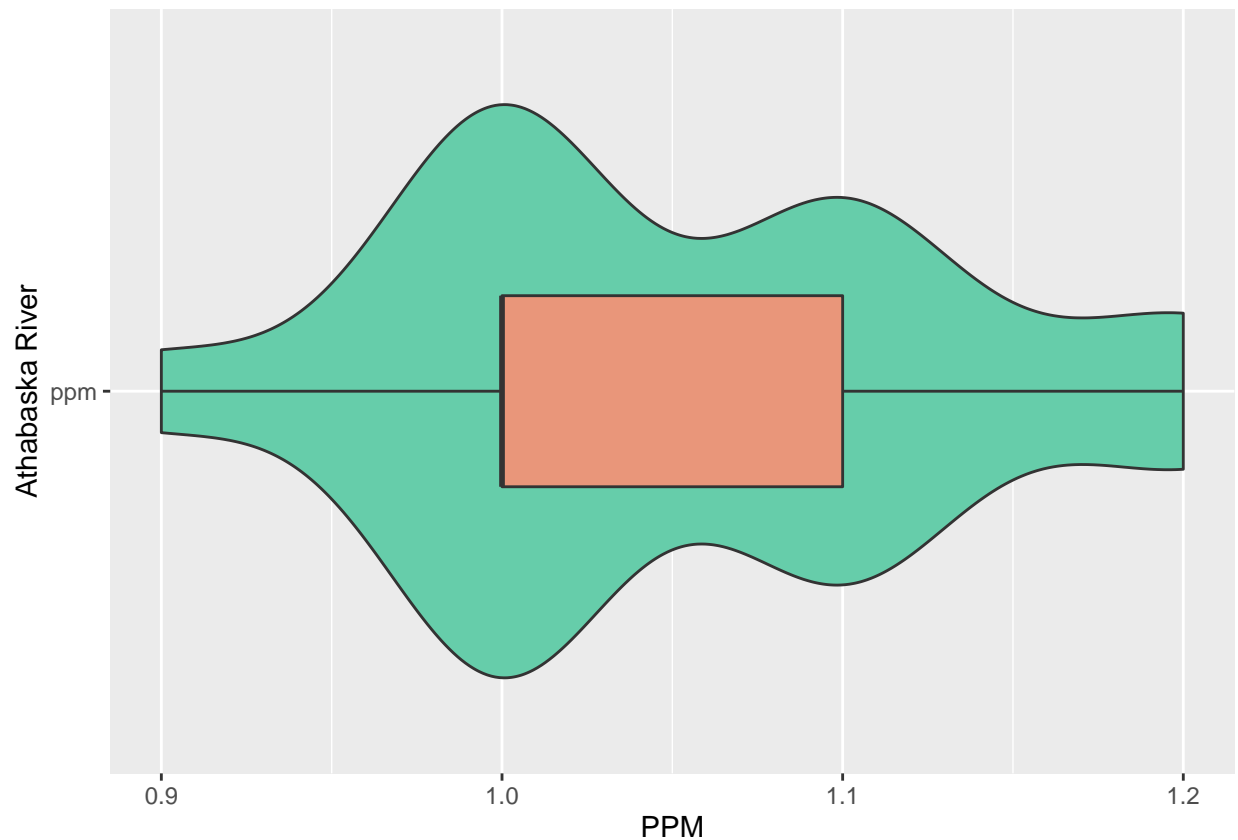
b)

Type I Error: Reject H_0 when H_0 is true. This would mean concluding that the PPM of mercury in Athabaska River is not less than or equal to 1.0, when it actually is.

Type II Error: Fail to reject H_0 when H_0 is false. This would mean concluding that the PPM of mercury in Athabaska River is less than or equal to 1.0 when it is actually greater than 1.0.

c)

```
ggplot(q3.df, aes(x = q3_vec, y = q3_vec1)) + geom_violin(fill = "aquamarine3")+ geom_boxplot(width = 0.5)
```



q3.df

##	q3_vec1	q3_vec
## 1	ppm	1.2
## 2	ppm	1.1
## 3	ppm	1.0
## 4	ppm	1.0
## 5	ppm	1.1
## 6	ppm	1.0
## 7	ppm	1.0
## 8	ppm	1.0
## 9	ppm	0.9

```
## 10    ppm    1.1
## 11    ppm    1.1
## 12    ppm    1.2
## 13    ppm    1.0
## 14    ppm    1.1
## 15    ppm    1.0
## 16    ppm    1.1
## 17    ppm    1.0
## 18    ppm    0.9
## 19    ppm    1.0
## 20    ppm    1.0
## 21    ppm    1.1
## 22    ppm    1.0
## 23    ppm    1.0
## 24    ppm    1.1
## 25    ppm    1.2
## 26    ppm    1.0
## 27    ppm    1.1
## 28    ppm    1.0
## 29    ppm    1.0
## 30    ppm    1.2
## 31    ppm    1.1
```

The PPM of mercury in fish located in the Athabaska River ranged from 0.9 to 1.2 with the most frequent level being 1.0. Most of the data lies above 1.0 PPM rather than below 1.0 PPM.

Because our sample size is greater than 25, we do not need to check for normality before using `t.test`.

d)

```
t.test(q3_vec, mu = 1, alternative = "greater")

##
## One Sample t-test
##
## data:  q3_vec
## t = 3.5425, df = 30, p-value = 0.0006595
## alternative hypothesis: true mean is greater than 1
## 95 percent confidence interval:
##  1.026884      Inf
## sample estimates:
## mean of x
##  1.051613
```

Using `t.test`, we get a p value of 0.0007. Comparing this value to a set alpha value of 0.05, we reject our null hypothesis that the mercury levels are not greater than 1.0 ppm with a significance level of 0.05. This means Health Canada should place a moratorium on commercial fishing on walleye fish in Athabaska River.

Our p-value means that we have a 0.0007 chance of obtaining test results at least as extreme as the ones we actually observed. This means that it is extremely unlikely that our observation from our sample was a random outlier.

From our t.test we get a 95% confidence interval of **(1.0269, inf)**. The reason the upper bound of the interval is infinity is because we were conducting a right sided test. The meaning of this interval is that we can say with 95% confidence that the mean mercury levels in Athabaska River is greater than 1.0269.

Question 4

We are trying to test if the proportion of environmentally certified coffee growers in Southern Mexico is **greater than 0.6**. We begin testing by setting up the null and alternative hypothesis:

Null hypothesis: $H_0 : P_{certified} = 0.6$

Alternative hypothesis: $H_A : P_{certified} > 0.6$ From this, we know we will be conducting a right tailed test.

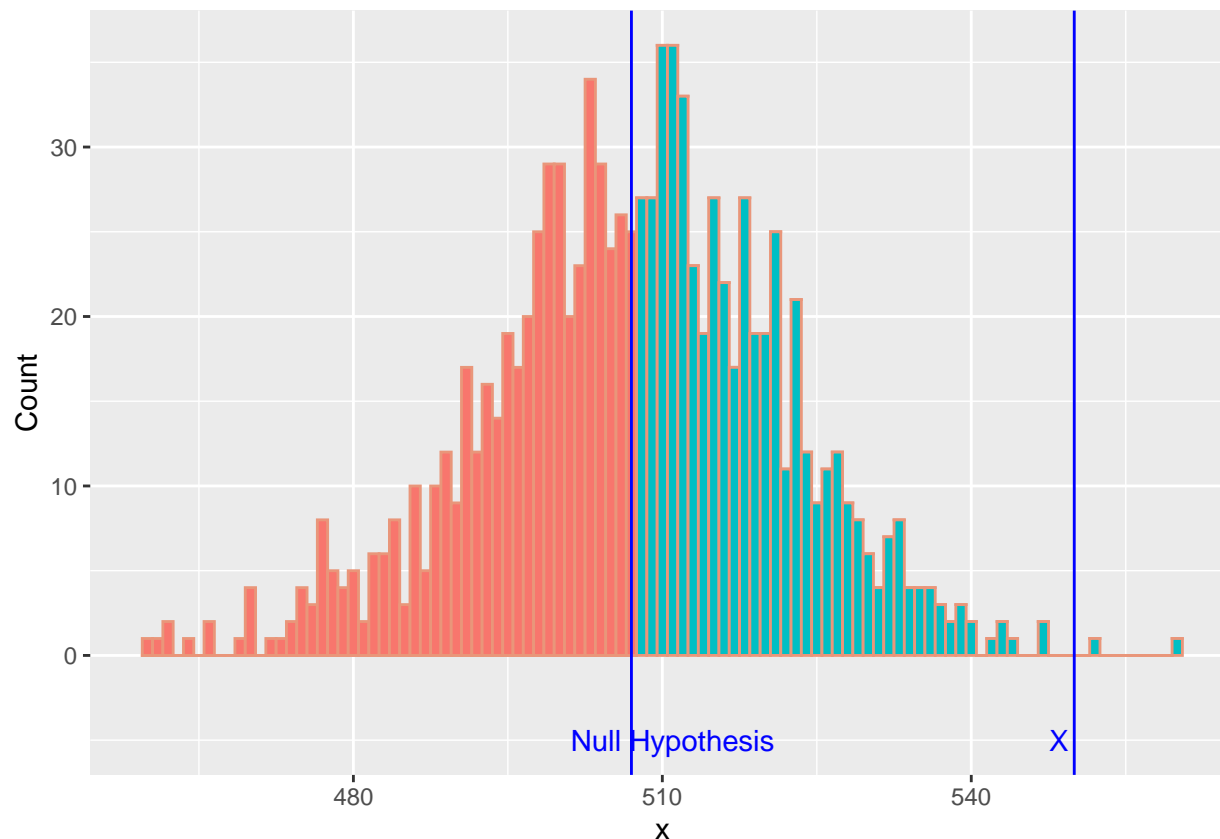
From our representative sample of $n = 845$, we get a test statistic of $X = 475 + 75 = 550$ environmentally friendly coffee growers. We can now create a distribution of the test statistic by re-sampling $n = 845$. We will do 1000 re-samples.

```
q4_vec <- c(rep(1, 845*0.6), rep(0, 845*0.4))

q4_samples <- numeric(1000)
for(i in 1:1000){
  pst.simulation <- resample(q4_vec)
  q4_samples[i] <- sum(pst.simulation)
}
q4.df <- data.frame(x = q4_samples, below = q4_samples <= 845*0.6, above = q4_samples > 845*0.6)
```

Plotting the distribution:

```
ggplot(data=q4.df, aes(x=x, fill = above)) + geom_histogram(binwidth = 1, col = "darksalmon") + ylab("C
```



As seen from the distribution, our observed amount of environmentally friendly coffee growers is on the far-right tail, very far from our assumed amount of coffee growers, the null hypothesis, at $x = 507$. To conclude our test, we will calculate the area under the curve to the right of 550 (X), known as the p-value, and compare this to our alpha value which is set to 0.05.

```
1 - pbinom(550, 845, 0.60)
```

```
## [1] 0.001045019
```

Using the pbinom function, we get a p-value of 0.0010. Because this is less than our alpha of 0.05, we can reject our null hypothesis and conclude that the proportion of Southern Mexican coffee growers who are either certified or in the process of being certified is more than 60% with probability of type 1 error of 0.05.

Question 5

To calculate the 95% pivotal quantity confidence interval, we can use prop.test. We will make sure to use the Agresti-Coull version of the confidence interval by adding 1 to the numerator and 2 to the denominator to each proportion.

```
prop.test(c(274+1, 348+1), c(376+2, 670+2), conf.level = 0.95, correct=FALSE)$conf
```

```
## [1] 0.149503 0.266833
## attr("conf.level")
## [1] 0.95
```


The 95% confidence interval is **(0.1495, 0.2668)**. This means we can conclude that with 95% confidence that the proportion of people with at least a undergraduate degree that do not believe in the science behind the vaccine is larger than the proportion of people with at most a high school diploma that do not believe in the science behind the vaccine. This is because both our upper and lower bound in our 95% confidence interval of $p_{Uni} - p_{HS}$ is positive.

Question 6

a)

Null hypothesis: $H_0 : P_{vote} = 0.45$

Alternative hypothesis: $H_A : P_{vote} < 0.45$. This tells us that we will be conducting a left sided test.

b)

In order to calculate α , we compute the probability $P(RejectH_0|H_0isTrue)$.

$P(H_0isTrue)$ is 0.45.

$P(RejectH_0)$ is $P(X < 20)$

We can use pbinom to calculate alpha. This will give us the probability that we reject the null hypothesis if it actually true.

```
pbinom(19,50, 0.45)
```

```
## [1] 0.1973679
```

We get an alpha value of 0.19737.

c)

To get the probability that we will decide that the candidate should run, we compute the area to the right of $x = 20$, given that the candidate gets 42% of the votes.

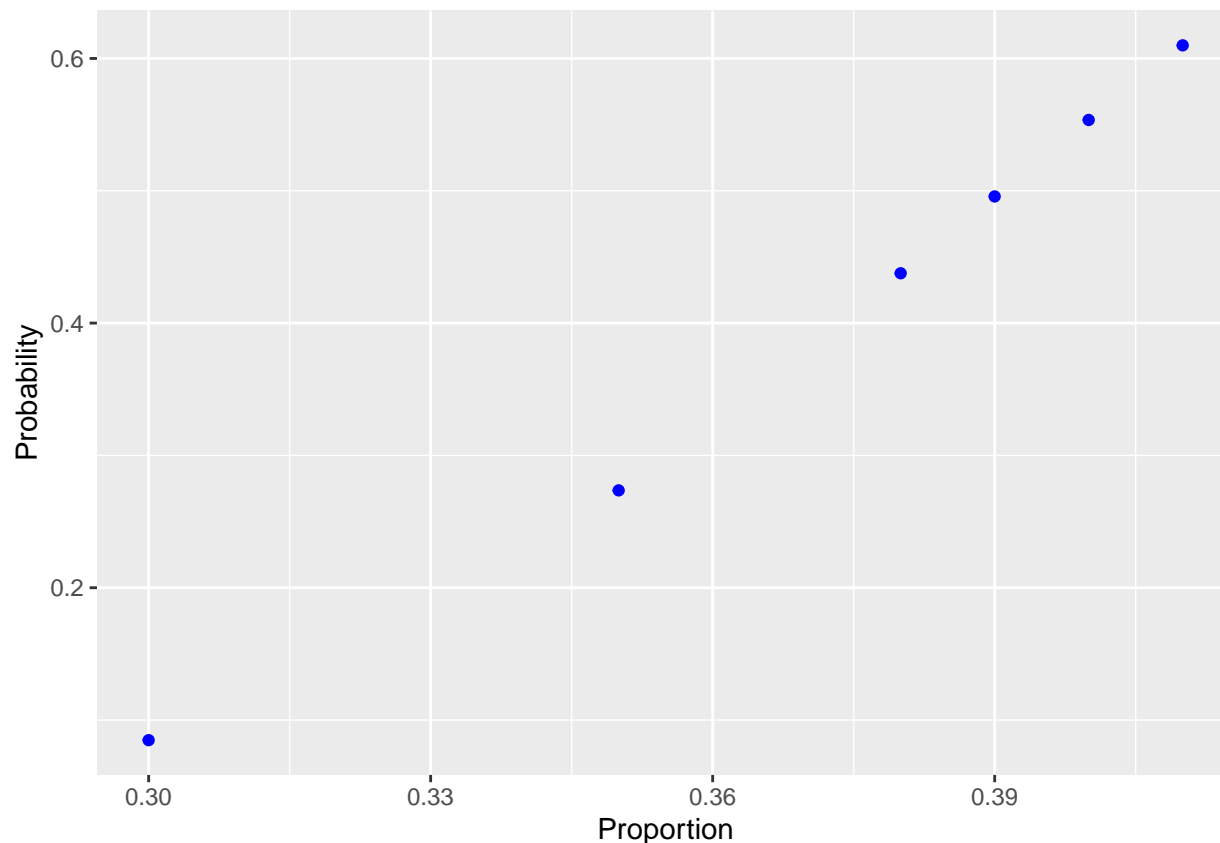
```
1 - pbinom(19, 50, 0.42)
```

```
## [1] 0.663807
```

We get a probability of 0.6638.

d)

```
x <- c(0.41, 0.40, 0.39, 0.38, 0.35, 0.30)
y = numeric(6)
for (vals in 1:length(x)){
  y[vals] = 1 - pbinom(19, 50, x[vals])
}
q6.df <- data.frame(x, y)
ggplot(q6.df, aes(x = x, y = y)) + geom_point(stat="identity", color = "blue") + xlab("Proportion") + ylab("Probability")
```



e)

As the proportion of people that vote for the candidate increases, the probability that we conclude that the candidate should run increases as well.

Since we are plotting proportion of people that vote for the candidate that are less than 0.45, these are also beta/type 2 error. ($P(\text{Fail to Reject } H_0 | H_0 \text{ is False})$). One way to improve the test would be to increase the threshold of successes observed in the sample to reject the null hypothesis. This is because we assume the proportion is 0.45, however we advise the candidate to run when we observe at least 20 vote ($20/50 = 0.4$).

Question 7

a)

To test the claim that the price of used textbooks from the University of Calgary are more than from Amazon, we will create a new variable variable called “Diff” that is price of the used textbook at UofC minus the price at Amazon. We will then do our statistical test on this new variable. The reason we are able to do this is because the “object” (textbook) is the same at both the UofC and Amazon and we can therefore call these two samples dependent. The null and alternative hypothesis are below:

$$H_0 = \mu_d = 0$$

$$H_A = \mu_d > 0$$

```
# Read the data file into R
q7_data <- read.csv("http://people.ucalgary.ca/~jbstall/DataFiles/bookprices.csv")
q7_data["Diff"] = q7_data["UsedBkStore"] - q7_data["UsedAmazon"]
```

```
t.test(q7_data$Diff, mu=0, alternative="greater", data=q7_data)
```

```
##
## One Sample t-test
##
## data: q7_data$Diff
## t = 3.0357, df = 14, p-value = 0.004449
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
## 10.69172 Inf
## sample estimates:
## mean of x
## 25.468
```

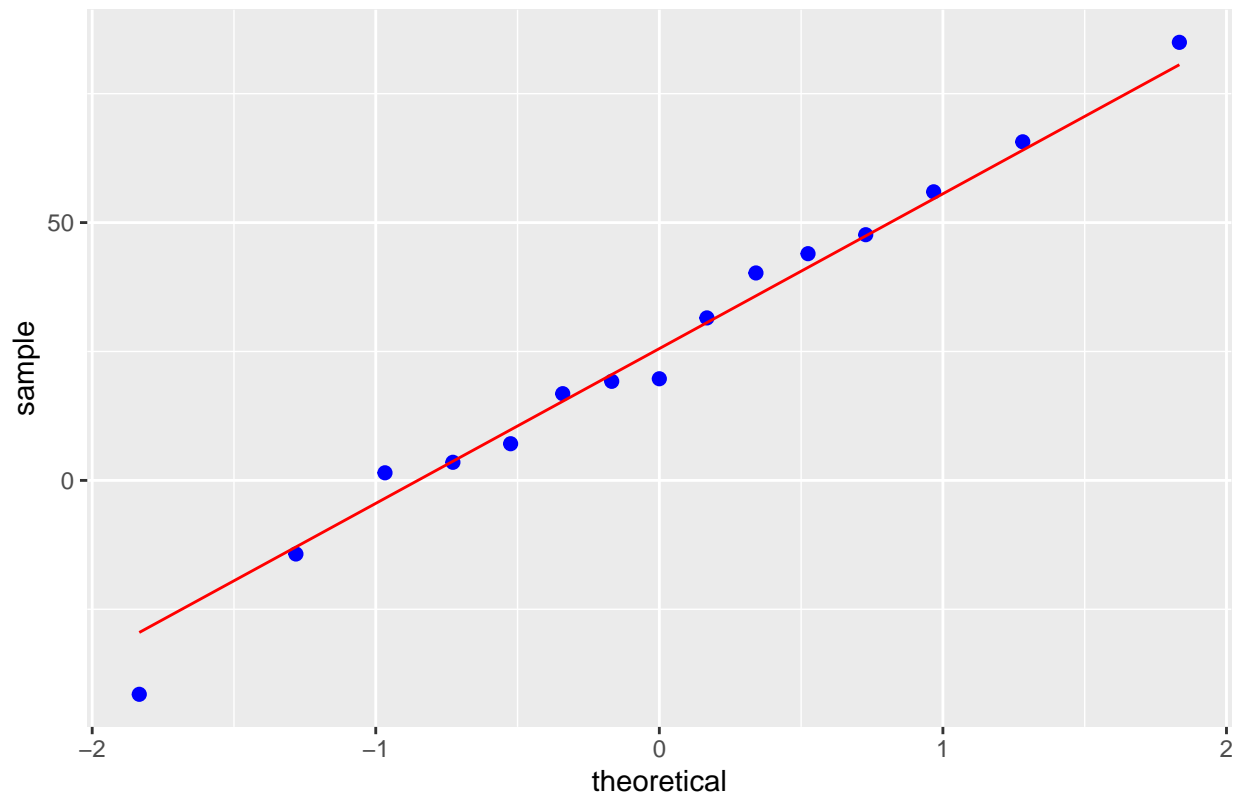
From our t-test, we get a p-value of 0.0044. Using an alpha value of 0.05, means we can reject our null hypothesis that the mean of the difference between used textbook price at UofC and Amazon is 0. So we can conclude that on average, the price of the textbooks sampled are greater at the University of Calgary's bookstore compared to Amazon.

b)

Because both samples have less than 25 observations, we must test that the samples follow an approx. normal distribution using stat_qq.

```
ggplot(q7_data, aes(sample = Diff))+ stat_qq(size=2, col="blue") + stat_qqline(col="red") + ggtitle("Nor
```

Normal Probability Plot n = 15 Difference in Textbook Price



Looking at the normal probability plot of the difference column, the difference in the textbook prices at UofC and Amazon seem to follow an approx. normal distribution, so we can move on to testing out hypothesis using t.test.

Question 8

We want to test whether the proportion of Albertans with some post-secondary education is different than the claim of the Government of Alberta of 0.62. To start, we will set up our null and alternative hypothesis.

Null hypothesis: $H_0 : P_{PS} = 0.62$

Alternative hypothesis: $H_A : P_{PS} \neq 0.62$. This tells us that we will be conducting a two-sided test.

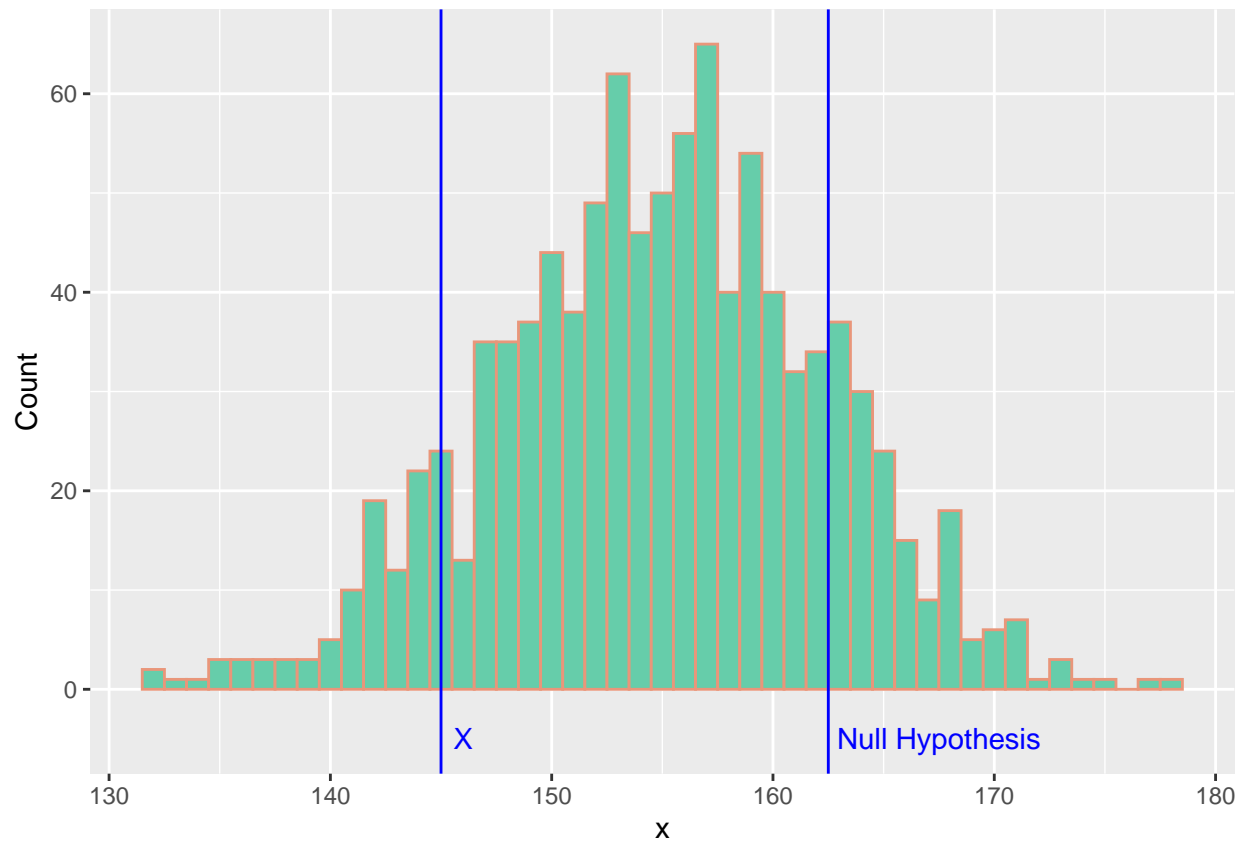
To conduct a two-sided, we will find the p-value of one side and multiply it by 2. For our test, we will use an alpha of 0.05. We can use binom.test to do this.

We will create a distribution by re-sampling 1000 times and then visualize the distribution.

```
q8_vec <- c(rep(1, 250*0.62), rep(0, 250*0.32))

q8_samples <- numeric(1000)
for(i in 1:1000){
  pst.simulation <- resample(q8_vec)
  q8_samples[i] <- sum(pst.simulation)
}
q8.df <- data.frame(x = q8_samples, below = q8_samples <= 250*0.68, above = q8_samples > 250*0.38)
```

```
ggplot(data=q8.df, aes(x=x)) + geom_histogram(binwidth = 1, fill = "aquamarine3", col = "darksalmon") +
```



We will compute the left sided p-value by using two sided binom.test.

```
binom.test(145, 250, 0.62, alternative='two.sided', correct=FALSE)
```

```
##
##
##
## data: 145 out of 250
## number of successes = 145, number of trials = 250, p-value = 0.1932
## alternative hypothesis: true probability of success is not equal to 0.62
## 95 percent confidence interval:
## 0.5161732 0.6419069
## sample estimates:
## probability of success
## 0.58
```

We get a p-value of 0.1932 which is larger than our alpha of 0.05, so we fail to reject the null hypothesis that the proportion of Albertans that have post secondary education is 0.62. Therefore we can conclude that the proportion of Albertans that have post secondary education not is different from 0.62 with a significance level of 0.05.

Question 9

a)

Since this is a two-sided test, we can use either bootstrapping or `prop.test`. We will try both and compare the results.

Bootstrapping: We first set up two representative vectors of the surveys in 2012 and 2019 respectively so we can sample from them while using bootstrapping.

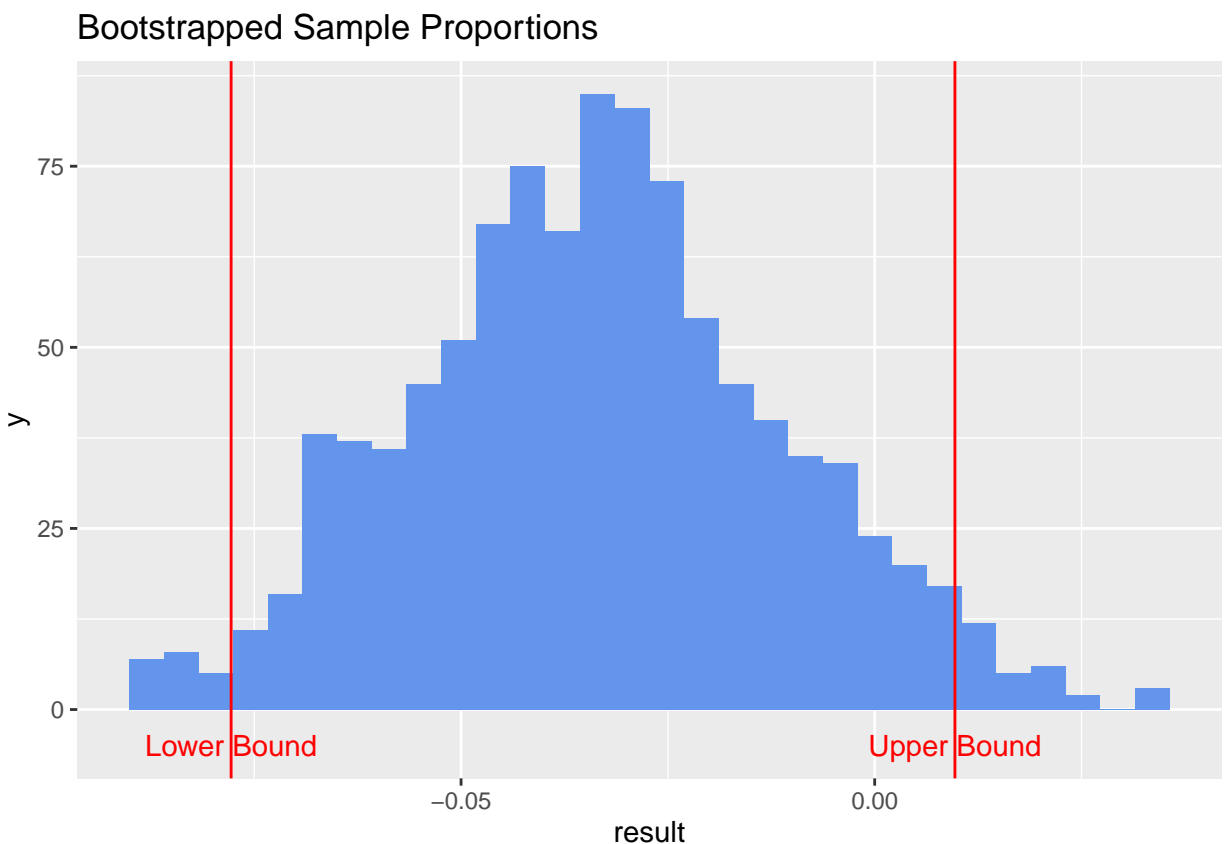
```
plastics_2012 <- c(rep(1, 601), rep(0, 1010-601))
plastics_2019 <- c(rep(1, 561), rep(0, 1000-561))

q9_bootstrap = do(1000)*(mean(resample(plastics_2019))- mean(resample(plastics_2012)))
q9_phat.df <- data.frame(q9_bootstrap)
qdata(~result, c(0.025, 0.975), data=q9_phat.df)
```

```
##      2.5%      97.5%
## -0.0749703  0.0115005
```

```
ggplot(q9_bootstrap, aes(x = result)) + geom_histogram(fill = "cornflowerblue") + ggtitle("Bootstrapped
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



Our 95% confidence interval from bootstrapping is **(-0.0778, 0.0097)**.

```
prop.test(c(561+1, 601+1), c(1000+2, 1010+2), conf.level = 0.95, correct=FALSE)$conf
```

prop.test using Agresti-Coull

```
## [1] -0.077100209 0.009133376
## attr(,"conf.level")
## [1] 0.95
```

Our 95% confidence interval from prop.test is **(-0.0771, 0.0091)**.

Both confidence intervals from bootstrapping and prop.test are very similar.

b)

Since 0 is within our 95% confidence interval, we **cannot** conclude that there is a significant difference between the proportions in 2019 and 2012, with 95% confidence.

Question 10

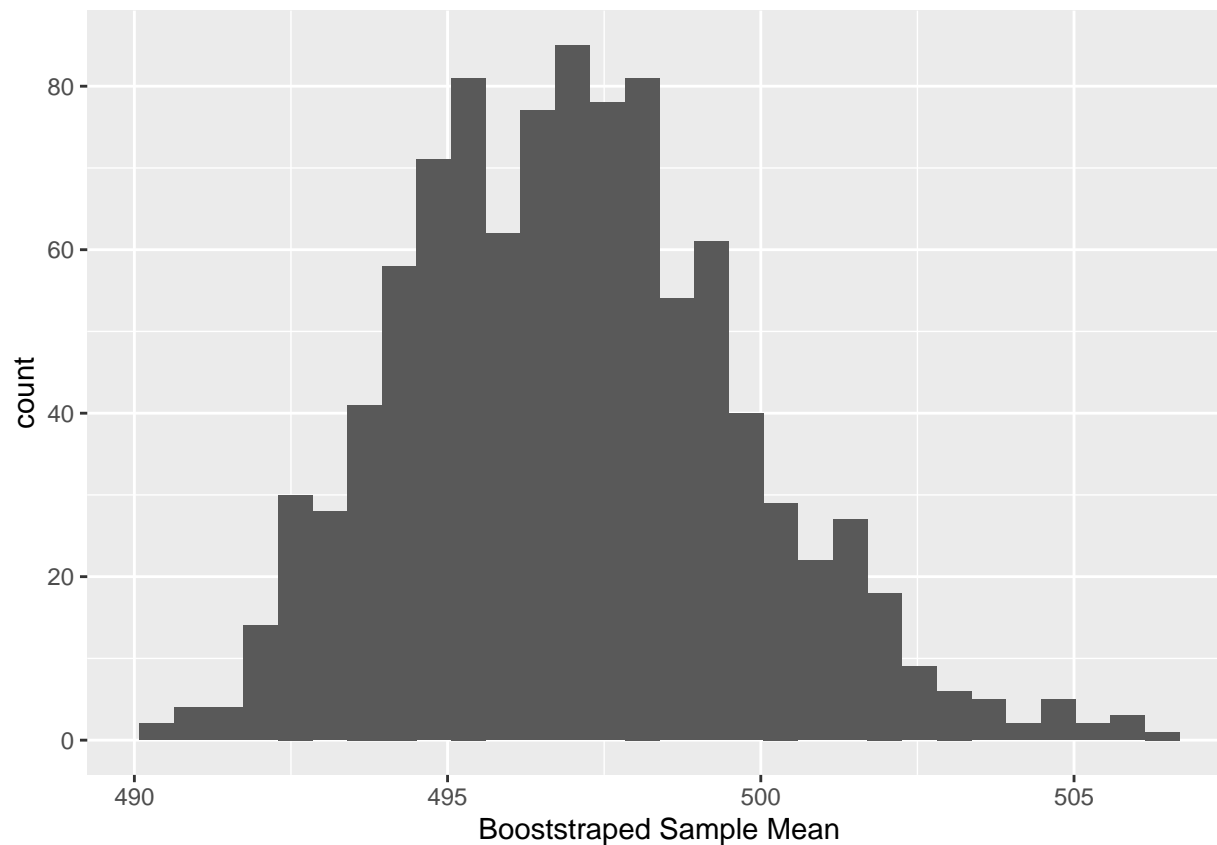
```
q10_vec <- c(497.2, 499.9, 495.8, 514.2, 490.0, 498.3, 495.1, 486.7)
x_bar = mean(q10_vec)
s_sample = sd(q10_vec)
n = length(q10_vec)
q10.df <- data.frame(q10_vec)
```

a)

We can use a statqq plot to check the normality of the data.

```
q10a_bootstrap = do(1000)*(mean(resample(q10_vec)))
q10a.df <- data.frame(q10a_bootstrap)
ggplot(q10a.df, aes(x = result)) + geom_histogram() + xlab("Booststraped Sample Mean")
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



Looking at the plot of bootstrapped sample means of cereal box weights, it seems like the data suggests that Usman is getting around 500g on average in each box.

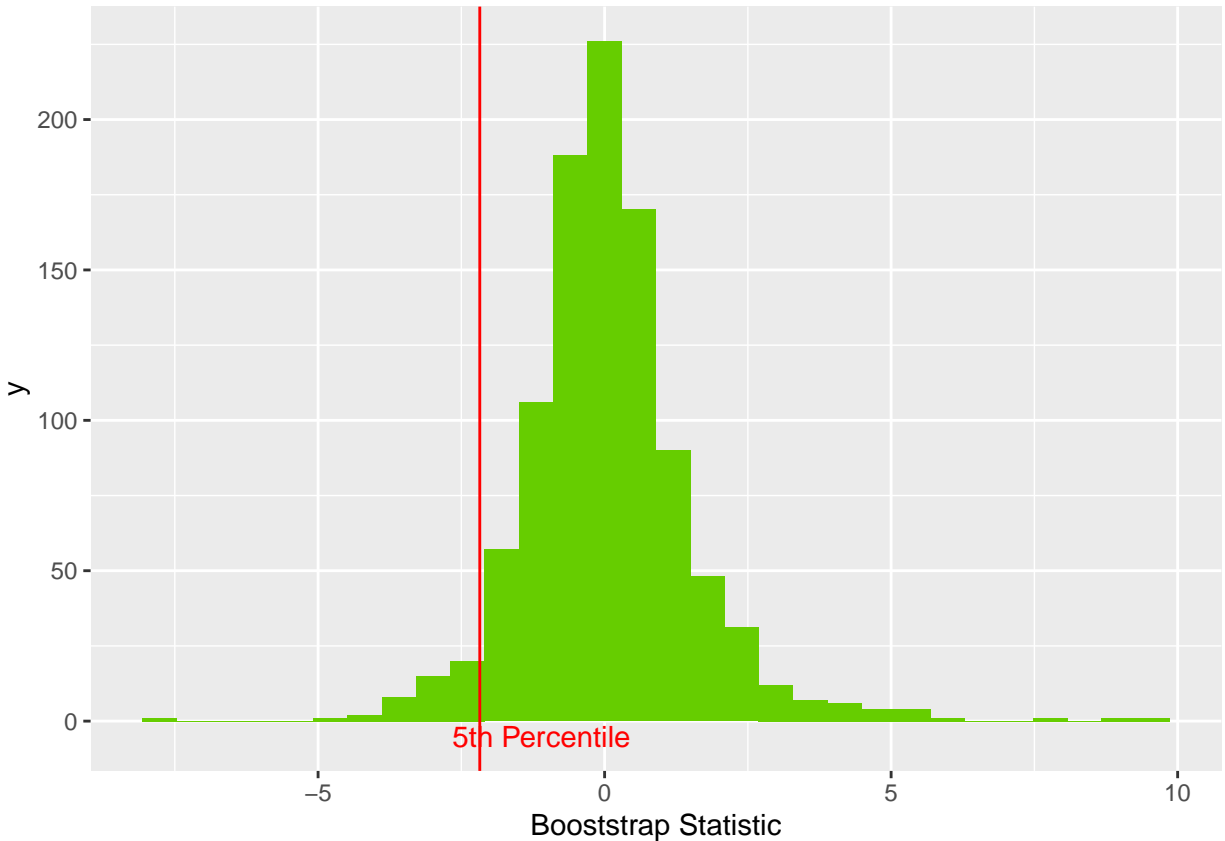
b)

```
q10b_bootstrap = do(1000)*((mean(resample(q10_vec))- x_bar)/(sd(resample(q10_vec))/(sqrt(n))))
q10b.df <- data.frame(q10b_bootstrap)
qdata(~result, c(0.05), data=q10b.df)
```

```
##          5%
## -2.012121
```

```
ggplot(q10b.df, aes(x = result)) + geom_histogram(fill = "chartreuse3") + xlab("Bootstrap Statistic")
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```

The 5th percentile of the distribution of the bootstrap is -2.0121. This means 5% of the distribution lies below -2.0121.

c)

```
x_bar - (-2.1059)*(s_sample/n)
```

```
## [1] 499.2977
```

We get a 5th percentile of \bar{x} of 499.30. We cannot infer the same thing as in part a). This is because if the 5th percentile of sample means is 499.30, which is very close to 500g, the weight of each box on average is not going to be 500g since 95% of the sample means will lie above 499.30.

Question 11

a)

To calculate alpha, we calculate $P(\text{Reject } H_0 | H_0 \text{ is True})$. We assume H_0 is true, so $\beta = 2$. We reject the null hypothesis when $X_{\min} < 2.018$

To find this probability, we must find the area under the curve defined by the function $f(x_{\min}) = ne^{-n(x_{\min} - \beta)}$. The upper bound will be 2.018 because this is highest value of x_{\min} where we can reject the null hypothesis. The lower bound will be 2 because $\beta \leq x_{\min}$, and assuming the null hypothesis is true means beta is 2 and therefore $2 \leq x_{\min}$.

Before we integrate the function, we can plug known values of $n = 6$ and $\beta = 2$ to simplify. We get:

$$\int_2^{2.018} 6e^{-6(x_{min}-2)} dx_{min}.$$

Simplifying the exponent, we get:

$$e^{12} \int_2^{2.018} 6e^{-6x_{min}} dx_{min}$$

It is important to note that $6e^{-6x_{min}}$ follows an exponential distribution with $\lambda = 6$, so we can use the function `pexp` to get the area under the curve.

```
exp(12)*(pexp(2.018, rate = 6) - pexp(2, rate = 6))
```

```
## [1] 0.1023724
```

Multiplying the area under the curve by e^{12} , we get an alpha value of 0.1024.

b)

This sample does support the null hypothesis. We can only reject the null hypothesis is $X_{min} < 2.018$, however the minimum value in the sample is 2.11.

c)

To get beta, we calculate $P(FailRejectH_0|H_0isFalse)$. A useful note is that this is the compliment of alpha. So we can modify the expression in a) to find probability of concluding $\beta = 2$ when $\beta = 1.8$.

Since our actual beta is 1.8, this is our new lower bound.

$$1 - \int_{1.8}^{2.018} 6e^{-6(x_{min}-2)} dx_{min}$$

$$1 - (e^{(6*1.8)} \int_2^{2.018} 6e^{-6x_{min}} dx_{min})$$

We can now use the `pexp` in R to calculate this.

```
1 - (exp(6*1.8)*(pexp(2.018, rate = 6) - pexp(1.8, rate = 6)))
```

```
## [1] 0.2703602
```

We get a probability of 0.2704 that we will conclude that beta is 2 when beta is actually 1.8.

d)

We must solve for x_{min} in the following expression:

$$P(RejectH_0|H_0isTrue) = 0.2$$

$$P(X_{min} < k|\beta = 2) = 0.2$$

We can utilize the expression in a).

$$\int_2^k 6e^{-6(x_{min}-2)} dx_{min} = 0.2$$

$$e^{12} \int_2^k 6e^{-6x_{min}} dx_{min} = 0.2$$

Completing the integration:

$$e^{12}[-e^{-6x_{min}} + e^{-12}]_2^k = 0.2$$

Substituting k:

$$-e^{-6k} + e^{-12} = \frac{0.2}{e^{12}}$$

Simplifying:

$$0.8e^{-12} = e^{-6k}$$

Taking the natural log of both sides:

$$\log(0.8) - 12 = -6k$$

$$\frac{\log(0.8)-12}{-6} = k$$

```
(log(0.8)-12)/(-6)
```

```
## [1] 2.037191
```

k is approx. 2.0372. This means $\leq X_{min} < 2.0372$ would result in a rejection of the null hypothesis.

Session info:

```
sessionInfo()
```

```
## R version 4.1.3 (2022-03-10)
## Platform: x86_64-conda-linux-gnu (64-bit)
## Running under: Ubuntu 22.04.1 LTS
##
## Matrix products: default
## BLAS/LAPACK: /opt/conda/lib/libopenblas-r0.3.21.so
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8      LC_NUMERIC=C
##  [3] LC_TIME=en_US.UTF-8      LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=en_US.UTF-8  LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_US.UTF-8     LC_NAME=C
##  [9] LC_ADDRESS=C             LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods    base
##
## other attached packages:
##  [1] resampledData_0.3.1 mosaic_1.8.3      ggribes_0.5.3      mosaicData_0.20.2
##  [5] ggformula_0.10.1   ggstance_0.3.5    Matrix_1.4-1       lattice_0.20-45
##  [9] dplyr_1.0.9        ggplot2_3.3.6
##
## loaded via a namespace (and not attached):
```

## [1]	ggrepel_0.9.1	Rcpp_1.0.9	tidyr_1.2.0	assertthat_0.2.1
## [5]	digest_0.6.29	utf8_1.2.2	ggforce_0.3.4	R6_2.5.1
## [9]	plyr_1.8.7	backports_1.4.1	labelled_2.9.1	evaluate_0.16
## [13]	highr_0.9	pillar_1.8.1	rlang_1.0.4	rstudioapi_0.14
## [17]	rmarkdown_2.15	labeling_0.4.2	splines_4.1.3	readr_2.1.2
## [21]	stringr_1.4.1	htmlwidgets_1.5.4	polyclip_1.10-0	munsell_0.5.0
## [25]	broom_1.0.0	compiler_4.1.3	xfun_0.32	pkgconfig_2.0.3
## [29]	htmltools_0.5.3	tidyselect_1.1.2	tibble_3.1.8	gridExtra_2.3
## [33]	mosaicCore_0.9.0	fansi_1.0.3	tzdb_0.3.0	withr_2.5.0
## [37]	MASS_7.3-58.1	grid_4.1.3	gtable_0.3.0	lifecycle_1.0.1
## [41]	DBI_1.1.3	magrittr_2.0.3	scales_1.2.1	cli_3.3.0
## [45]	stringi_1.7.8	farver_2.1.1	leaflet_2.1.1	ellipsis_0.3.2
## [49]	ggdendro_0.1.23	generics_0.1.3	vctrs_0.4.1	tools_4.1.3
## [53]	forcats_0.5.2	glue_1.6.2	tweenr_2.0.1	purrr_0.3.4
## [57]	hms_1.1.2	crosstalk_1.2.0	fastmap_1.1.0	yaml_2.3.5
## [61]	colorspace_2.0-3	knitr_1.39	haven_2.5.0	