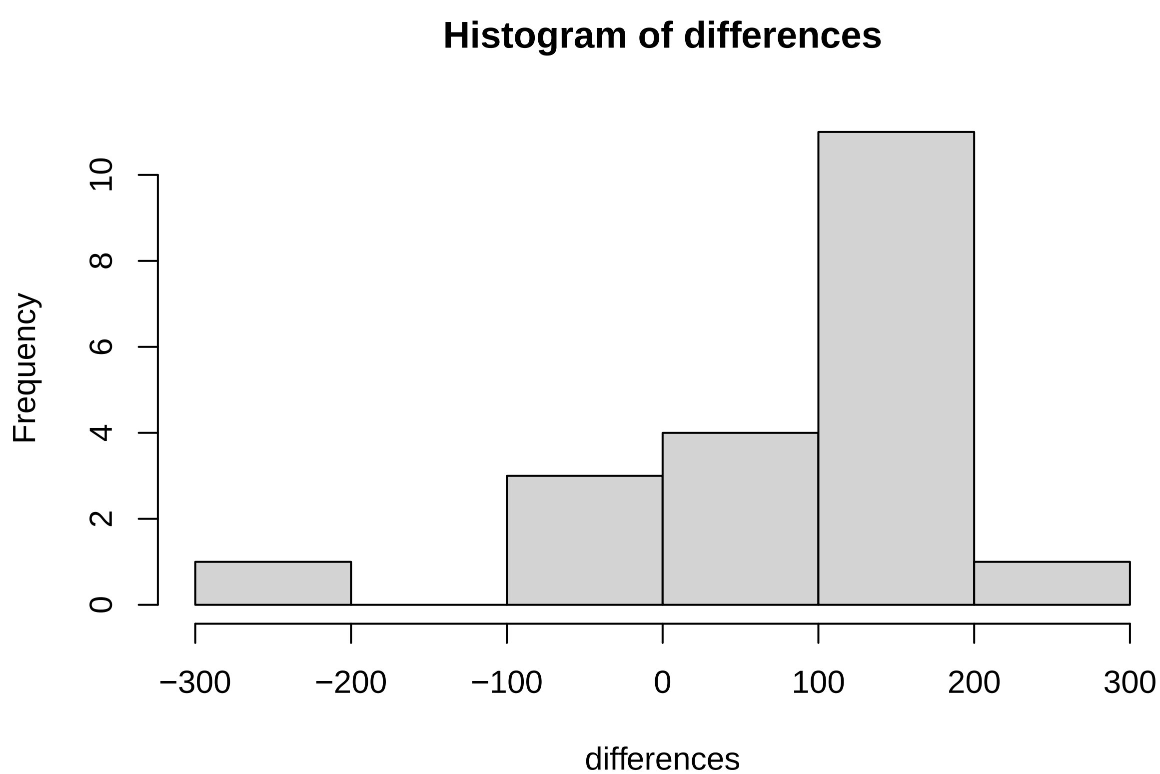
Assignment 4 407

Ryan Klughart

01/03/2021

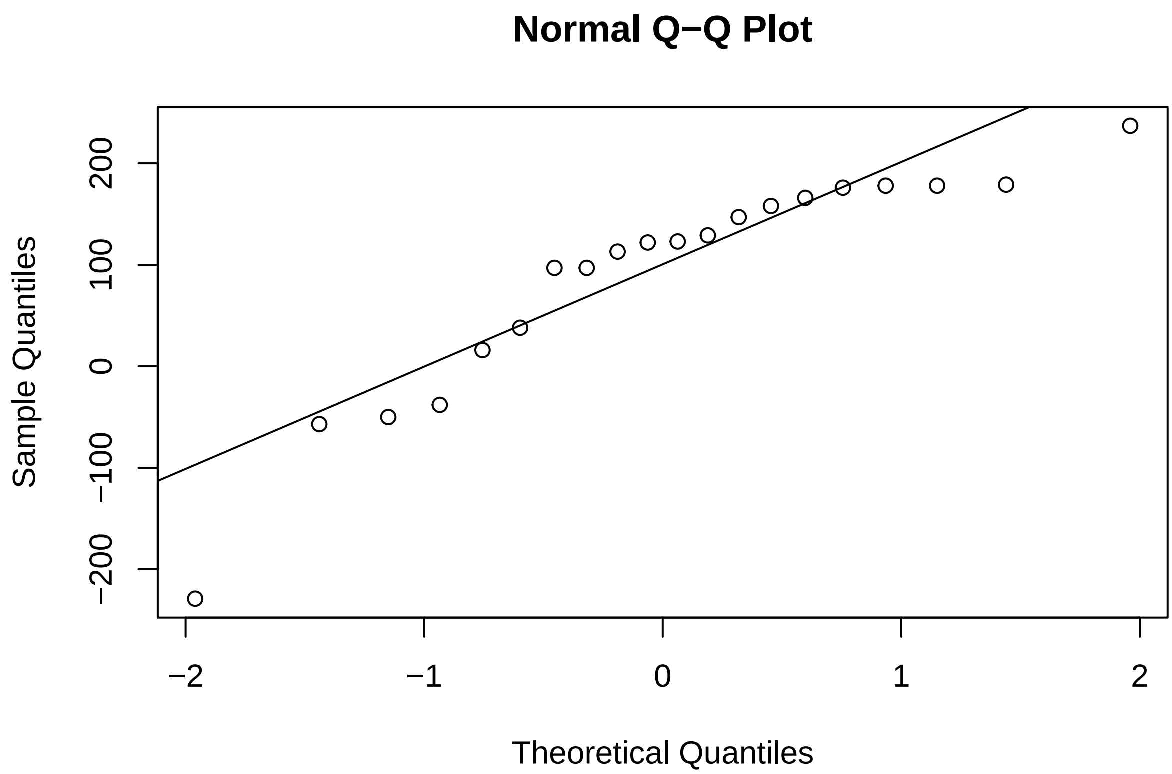
Question 1

```r
options(digits = 3)
before = c(342,582,251,647,620,535,559,522,610,506,
           568,619,650,666,579,532,531,464,499,417)

after = c(304,695,201,776,742,551,656,619,757,743,
          726,795,773,845,757,698,569,407,677,188)

differences = after-before

hist(differences)
```




```r
qqnorm(differences)
qqline(differences)
```

c)



From the histogram of the differences we can see that the distribution has a fairly strong left skewness. Additionally, from the qqnorm plot we can see that there is quite a lot of deviation from the line at both tails which suggests outliers. Since there is skewness and outliers, the normality assumption is violated and we cannot assume the sample of differences are independant normal random variables. Without normal random variables we wont be able to use a t-test.

Question 2

a) The populations of interest are the population of university students one month before the joint offer, and the population of university students one month after the joint offer. The population parameters of interest are the mean amount of money spent by students one month before the joint offer μ_1, and the mean amount of money spent by students one month after the joint offer μ_2

b) The null hypothesis suggests that there is no difference in the mean amount of money the students spend one month before and one month after the joint offer. In other words, the difference in mean amount of money spent before and after is equal to 0. The alternative hypothesis suggests that the mean amount of money spent by the students one month after the joint offer will be greater than the mean amount of money the university students spend one month before the joint offer.

μ_1 = mean amount of money spent by the population of students 1 month before the joint offer.
μ_2 = mean amount of money spent by the population of students 1 month after the joint offer.

$H_0 : \mu_1 = \mu_2$ $H_A : \mu_2 > \mu_1$

c)

```r
dbar = mean(differences)
dbar
```

```
## [1] 89
```

```r
n=20
sampleVariance = (1/(n-1))*sum((differences-dbar)^2)
standardError = sqrt(sampleVariance)/sqrt(n)
standardError
```

```
## [1] 25
```

```r
tratio = dbar/(standardError)
tratio
```

```
## [1] 3.56
```

```r
pvalue = pt(tratio, n-1, lower.tail = FALSE)
pvalue
```

```
## [1] 0.00105
```

```r
#t.test(after,before, paired=T, alternative = "greater" )
```

d) The p-value found is 0.001. Since the p-value is less than our significance level of 5% (0.05), we reject the null hypothesis.

e) After performing tests, we have found that the population of university students one month after the joint offer would spend a significant amount more money on average than the population of university students one month before the the joint offer.

Question 3

```r
UB = round(mean(differences) + 2.021*standardError,2)
LB = round(mean(differences) - 2.021*standardError,2)
UB
```

```
## [1] 140
```

```r
LB
```

```
## [1] 38.4
```

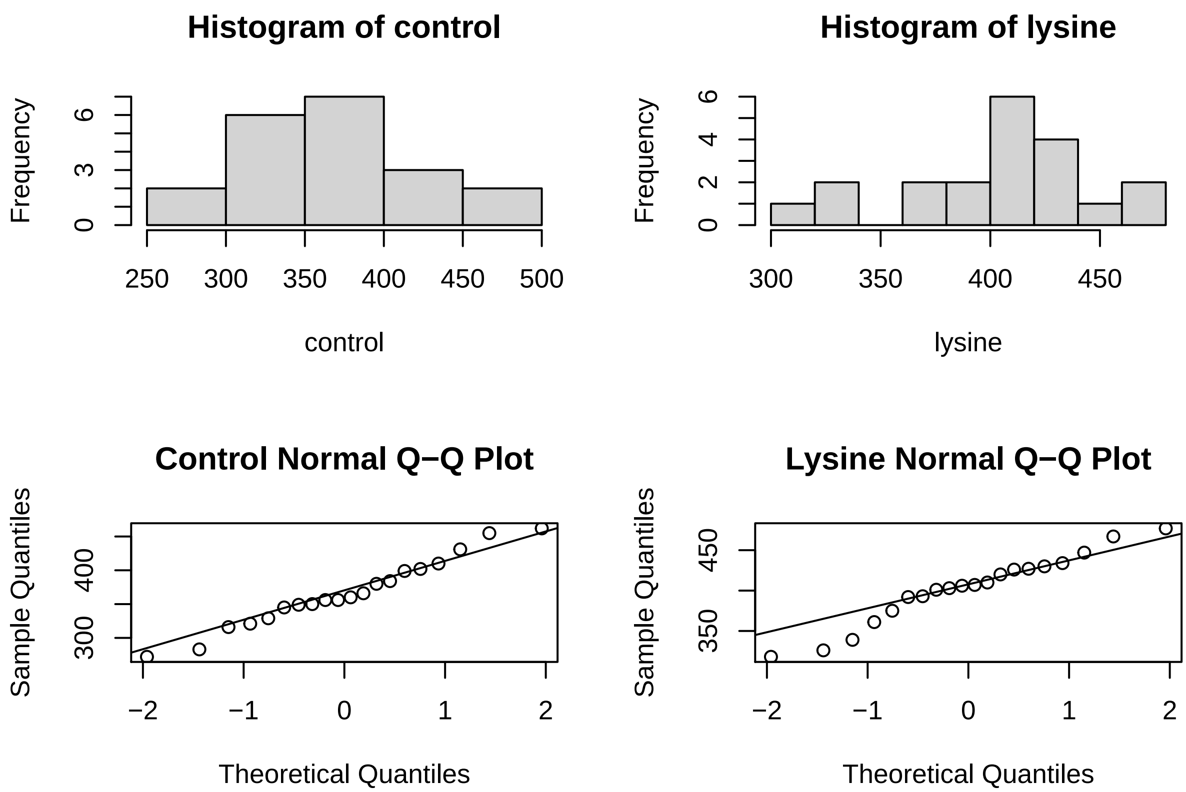
```r
EffectSize = mean(differences)
standardError = sqrt(sampleVariance)/sqrt(n)
```

A 95% confidence interval for the increase in money spent by the population of students one month after the joint offer compared to the population of students one month before the joint offer is (38.44$, 139.56$). This suggests that the bank can be 95% confident that each student will have an average increase of money spent by 38.44$ to 139.56$ the first month after they implement their joint offer, compared to one month before the joint offer. The effect size is 89$ and the standard error is 25.02$

Quesstion 4

```r
control = c(380,321,366,356,283,349,402,462,356,410,329,399,
            350,384,316,272,345,455,360,431)
lysine = c(361,447,401,375,434,403,393,426,406,318,467,407,
           427,420,477,392,430,339,410,326)
par(mfrow=c(2,2))
```

```r
hist(control)
hist(lysine)
qqnorm(control,main = "Control Normal Q-Q Plot")
qqline(control)
qqnorm(lysine,main = "Lysine Normal Q-Q Plot")
qqline(lysine)
```



The histogram of the control group has a small amount of skewness to the right. The histogram for the lysine group seems to be normally distributed with outlier to the left. From the Q-Q plot of the control group we can see slight deviations from the line but not significant enough for outliers. From the Q-Q plot for the lysine group we can see a large deviation from the line at the lower tail. This suggests that there are outliers. This violation of the normality assumption would prevent the use of a t-test.

Question 5

a) The two populations of interest are the population of 1-day old chick fed with normal corn and the population of 1-day old chicks fed with a corn that has increased amounts of lysine. The population parameters of interest are the mean weight gain from normal corn fed chicks μ_1 and the mean weight gain of chicks fed with high lysine corn μ_2.

b) The null hypothesis suggests that the mean amount of weight gain from the population of chicks fed norm corn wont be different from the mean amount of weight gain from the population of chicks fed corn that is high in lysine in 21 days. The alternative hypothesis suggests that the population of chicks fed corn that is high in lysine will gain more weight in 21 days than the population of chicks fed normal corn.

$\mu_1 =$ the mean weight gain for the population of normal corn fed chicks. $\mu_2 =$ the mean weight gain for the population of chicks fed with high lysine corn. $H_0 : \mu_1 = \mu_2$ $H_A : \mu_2 > \mu_1$

c)

```
controlMean = mean(control)
lysineMean = mean(lysine)
controlSD = sqrt(var(control))
lysineSD = sqrt(var(lysine))
```

Data Summary:

Treatment Group	Mean	Standard Deviation
Control	366.3g	50.805g
Lysine	402.95g	42.729g

d) Since the larger standard deviation 50.805g is less than 2 * 42.729g = 85.457g the assumption of equality of population variances is reasonable

e)

```
n=20
pooledVar = ((n-1)*(var(control)+var(lysine)))/(2*n-2)
pooledVar
```

```
## [1] 2203
```

```
t.ratio = (lysineMean-controlMean)/sqrt(pooledVar*(2/n))
t.ratio
```

```
## [1] 2.47
```

```
pval = pt(t.ratio,df = 2*n-2,lower.tail=FALSE)
pval
```

```
## [1] 0.00908
```

f) The p-value is 0.009, since the p-value is less than the significance level of 0.05 we reject the null hypothesis.

g) After performing tests, we found that the mean amount of weight gain for the population of chicks fed corn high in lysine μ_2 is significantly greater than the mean weight gain from the population of chicks fed normal corn μ_1. ## Question 6

```
SE = sqrt(pooledVar*(2/n))
UB = lysineMean-controlMean + 2.021*SE
UB
```

```
## [1] 66.6
```

```
LB = lysineMean-controlMean - 2.021*SE
LB
```

```
## [1] 6.65
```

```
sdpooled = sqrt((controlSD+lysineSD))
Effect = (lysineMean-controlMean) #cohens d
Effect
```

```
## [1] 36.6
```

```
CohensD = (lysineMean-controlMean)/SE
CohensD
```

```
## [1] 2.47
```

A 95% confidence interval for the mean increase in weight gain for the population of chicks fed high lysine corn compared to the population of chicks fed normal corn after 21 days is (6.65g, 66.65g). This means that farmers can be 95% confident that 1 day old chicks will gain an additional 6.65g to 66.65g in 21 days if they feed them with high lysine corn instead of normal corn. The effect size is 36.65g and the standard error is 14.844g

Question 7

why did the chicken cross the road? Probably because it was trying to get away from humans that wanted to chop it up and turn it into a Mc chicken sandwich, like we do to the other 60 billion chickens a year