Homework Assignment-8 (S- 520)

FNU ANIRUDH

1. (10.5 Problem Set A)

a) 1- sample t-test

$$t_n = \frac{X_n - \mu_0}{s_n/\sqrt{n}} = \frac{3.194887 - 0}{\sqrt{104.0118}/\sqrt{400}} = \frac{3.194887}{10.19862 / 20} = 6.265332$$

b) Expression iv= 1 – pt (1.253607, df = 399), best approximates the significance probability.

c) True : p = 0.03044555 < 0.05 = α -→ reject the null hypothesis

2. To build a confidence interval with confidence 0.95, the following needs to hold: 1 − α = 0.95 ==⇒ α/2 = 0.025.

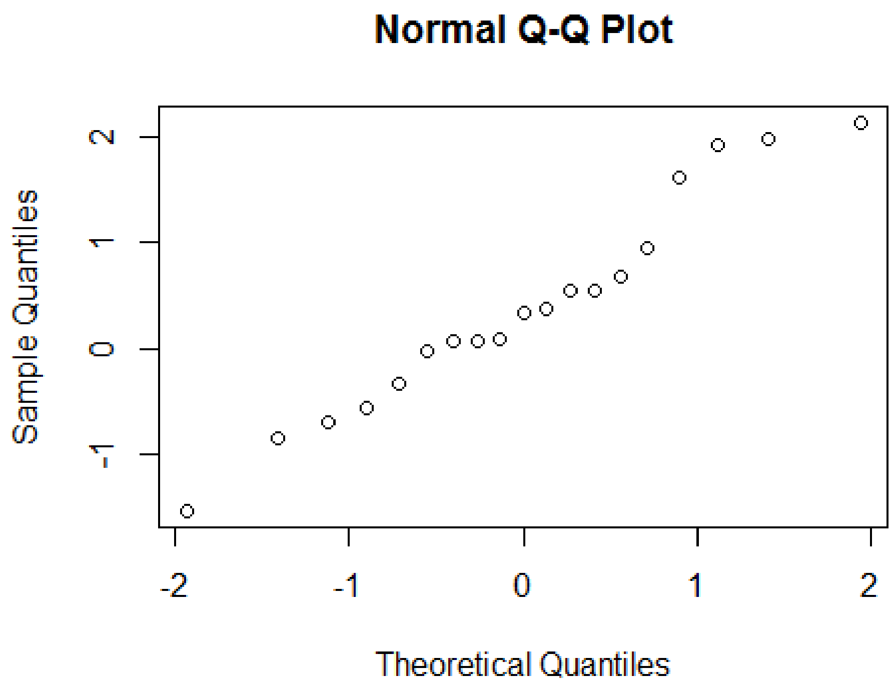
k= qbinom(0.025, 20, 0.5)
k=6
After sorting values in R
The form of interval is (sorting the values) : $(x_{(k+1)}, x_{(n-k)}) = (x_7, x_{14})$
= (239, 251)

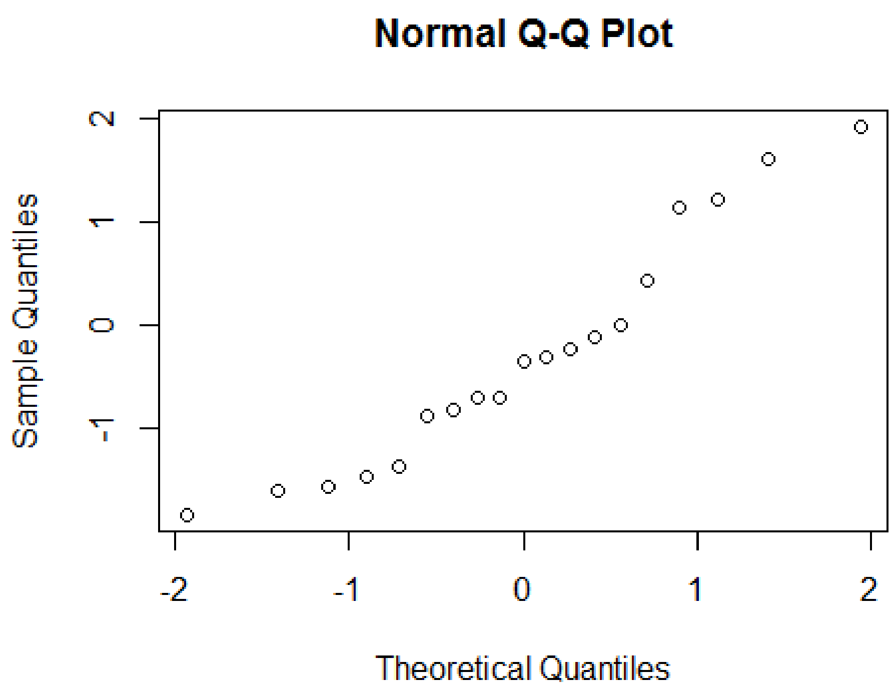
solution 2

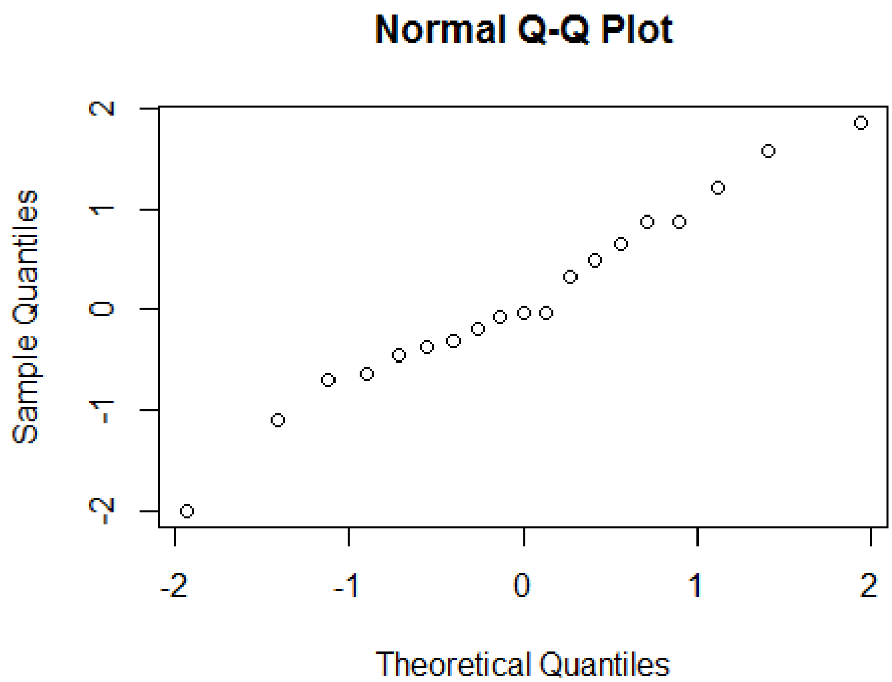
```
# a) Generating 4 samples from Normal Distribution with n=19
d <- data.frame(matrix(ncol = 4, nrow = 19))
i=4
while (i>0){
  x<-rnorm(19)
  d[,i]<-x
  i=i-1
}
d
```

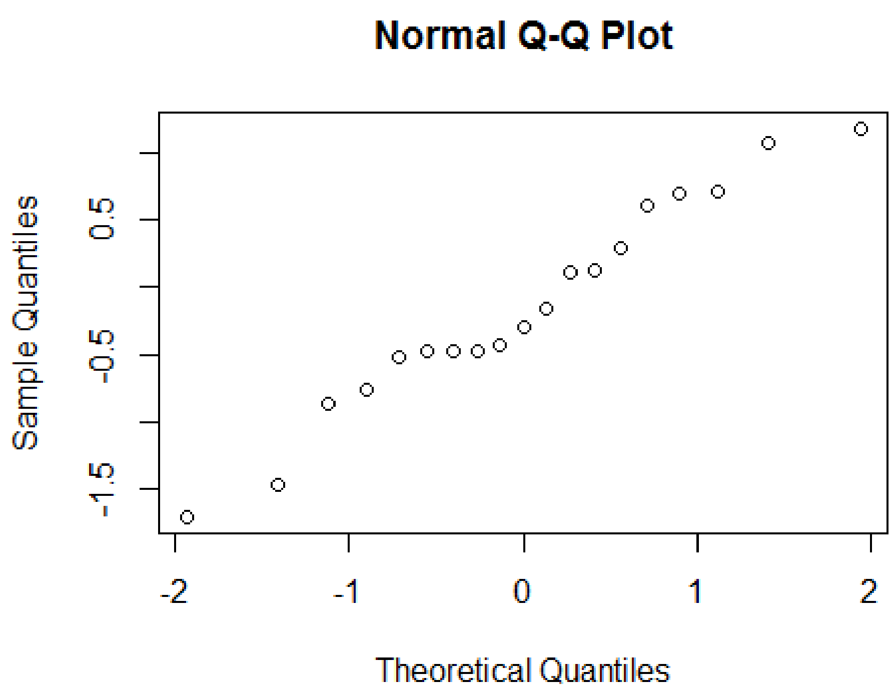
```
##             X1          X2          X3         X4
## 1  -1.53272317  1.60487944 -0.07541228 -0.4826547
## 2  -0.03600853 -0.23529556 -0.44803204 -0.5241822
## 3   0.08169431 -0.87929606 -2.00627100 -0.4352119
## 4   0.67205432  0.41749550 -0.36380116 -1.7078537
## 5  -0.32609883 -0.32315114  0.87156747 -1.4645963
## 6  -0.69586327 -1.37959405 -0.31727082 -0.1662459
## 7  -0.84653693 -0.70344290  1.22584503 -0.7629906
## 8   0.07150582 -1.46845106  0.33187508  0.2922941
## 9   0.54902297 -1.84562924  0.87079217 -0.2923694
## 10  2.12771897  1.14077821 -0.19817311  0.7008000
```

```
## 11  0.54511893 -0.11901312 -0.03060624 -0.4778269
## 12  0.93908104 -1.61172032 -1.10585205  1.0730649
## 13  0.36897691 -1.57236175 -0.69217617  0.5968755
## 14  1.96777958  1.91901840 -0.02815739  0.7047139
## 15 -0.56233223 -0.01202092  1.86839482 -0.8704041
## 16  0.33233669 -0.82273798 -0.62637023  0.1209882
## 17  1.90919608 -0.35868639  0.65039706  1.1766840
## 18  0.07366799 -0.70743538  1.58419423 -0.4842665
## 19  1.61712442  1.20458935  0.48915275  0.1102594

# b) QQ Plot for each sample
plot<-function(x) { qqnorm(x)}
apply(d,2,plot )
```


Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
-2
-1
0
1
2


Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
-2
-1
0
1
2



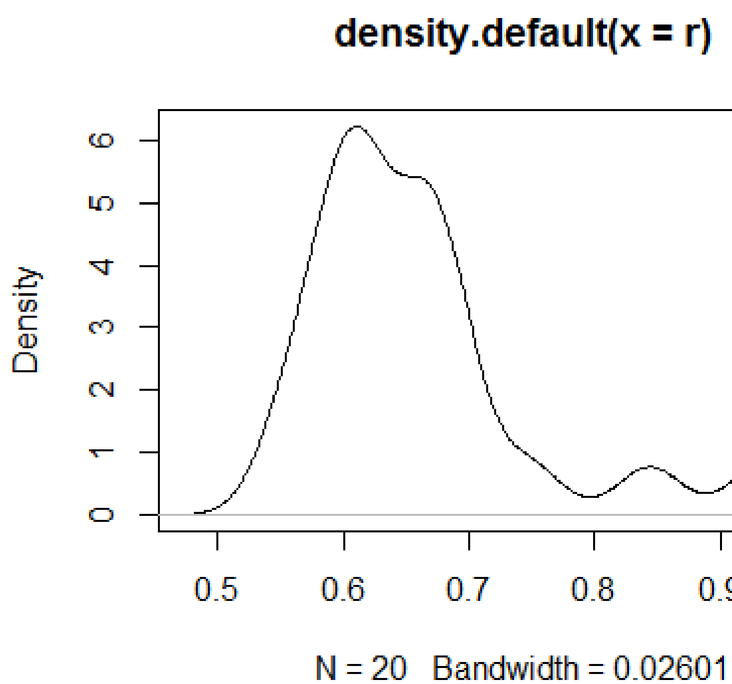



```
## $X1
## $X1$x
##  [1] -1.9379315 -0.5549229 -0.1323129  0.5549229 -0.7164975 -1.1189584
##  [7] -1.4121876 -0.4067243  0.4067243  1.9379315  0.2669941  0.7164975
```

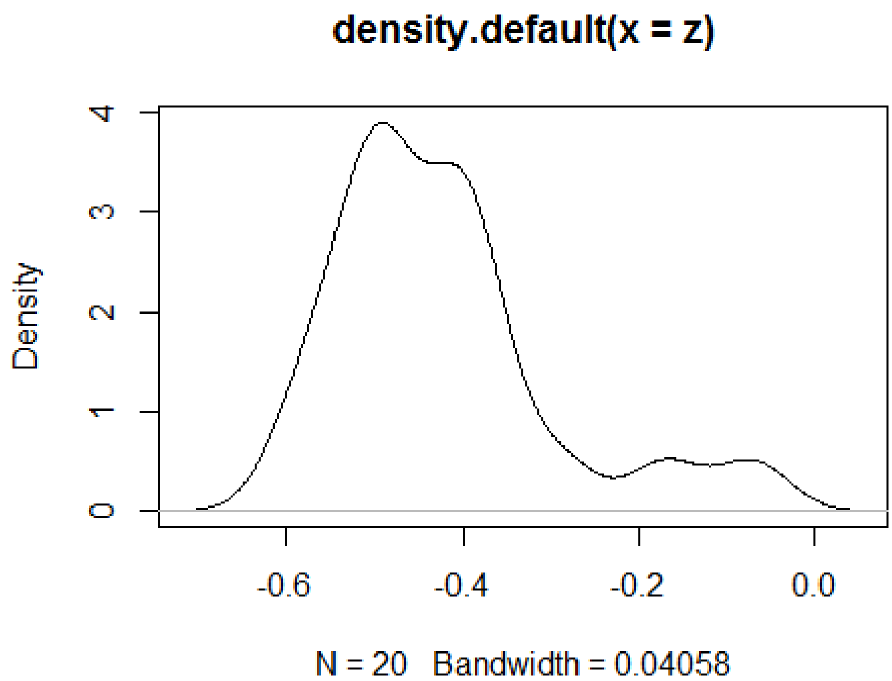
```
## [13]  0.1323129  1.4121876 -0.8994349  0.0000000  1.1189584 -0.2669941
## [19]  0.8994349
##
## $X1$y
##  [1] -1.53272317 -0.03600853  0.08169431  0.67205432 -0.32609883
##  [6] -0.69586327 -0.84653693  0.07150582  0.54902297  2.12771897
## [11]  0.54511893  0.93908104  0.36897691  1.96777958 -0.56233223
## [16]  0.33233669  1.90919608  0.07366799  1.61712442
##
##
## $X2
## $X2$x
##  [1]  1.4121876  0.2669941 -0.5549229  0.7164975  0.1323129 -0.7164975
##  [7] -0.1323129 -0.8994349 -1.9379315  0.8994349  0.4067243 -1.4121876
## [13] -1.1189584  1.9379315  0.5549229 -0.4067243  0.0000000 -0.2669941
## [19]  1.1189584
##
## $X2$y
##  [1]  1.60487944 -0.23529556 -0.87929606  0.41749550 -0.32315114
##  [6] -1.37959405 -0.70344290 -1.46845106 -1.84562924  1.14077821
## [11] -0.11901312 -1.61172032 -1.57236175  1.91901840 -0.01202092
## [16] -0.82273798 -0.35868639 -0.70743538  1.20458935
##
##
## $X3
## $X3$x
##  [1] -0.1323129 -0.7164975 -1.9379315 -0.5549229  0.8994349 -0.4067243
##  [7]  1.1189584  0.2669941  0.7164975 -0.2669941  0.0000000 -1.4121876
## [13] -1.1189584  0.1323129  1.9379315 -0.8994349  0.5549229  1.4121876
## [19]  0.4067243
##
## $X3$y
##  [1] -0.07541228 -0.44803204 -2.00627100 -0.36380116  0.87156747
##  [6] -0.31727082  1.22584503  0.33187508  0.87079217 -0.19817311
## [11] -0.03060624 -1.10585205 -0.69217617 -0.02815739  1.86839482
## [16] -0.62637023  0.65039706  1.58419423  0.48915275
##
##
## $X4
## $X4$x
##  [1] -0.4067243 -0.7164975 -0.1323129 -1.9379315 -1.4121876  0.1323129
##  [7] -0.8994349  0.5549229  0.0000000  0.8994349 -0.2669941  1.4121876
## [13]  0.7164975  1.1189584 -1.1189584  0.4067243  1.9379315 -0.5549229
## [19]  0.2669941
##
## $X4$y
##  [1] -0.4826547 -0.5241822 -0.4352119 -1.7078537 -1.4645963 -0.1662459
##  [7] -0.7629906  0.2922941 -0.2923694  0.7008000 -0.4778269  1.0730649
## [13]  0.5968755  0.7047139 -0.8704041  0.1209882  1.1766840 -0.4842665
## [19]  0.1102594
```

```
# c) ratio of Interquantile range to standard deviation for each sample
x<-c()
ratio<-function(x) { IQR(x)/sd(x)}
x<-c(x,apply(d,2,ratio))
x

##        X1        X2        X3        X4
## 0.9850215 1.1855577 1.2302576 1.2014168

# d) After trying in R, It is quite plausible to say that x bar was drawn
# from normal distribution
# 2.
z<-c(1.1402,-1.8658,0.8520,-1.8251,0.8530,-0.0589,-1.6554,-1.7599,-1.4330,
     -1.3853,2.9794,2.4919,2.1601,2.2670,-0.5479,-0.7164,0.6462,
     -0.8365,1.1997)
tval= ((mean(z)-0)/(sd(z)/sqrt(19)))
tval

## [1] 0.3545188

pval<- 1-pt(tval,18)
pval

## [1] 0.3635348

ci<- mean(z)+c(-1,1)*qt(.950,18)*sd(z)/sqrt(19)
ci

## [1] -0.5131008  0.7768166

# Since P value is so high , we cannot reject null hypothesis HO: μ ≤ 0
# 3
z<- sort(z)
z

##  [1] -1.8658 -1.8251 -1.7599 -1.6554 -1.4330 -1.3853 -0.8365 -0.7164
##  [9] -0.5479 -0.0589  0.6462  0.8520  0.8530  1.1402  1.1997  2.1601
## [17]  2.2670  2.4919  2.9794

k<- qbinom(0.05,19,0.5)
k

## [1] 6

# K=6 and 90% Confidence Interval of median (xk+1, xn-k) = (x7,x13)
z[7]

## [1] -0.8365

z[13]

## [1] 0.853
```

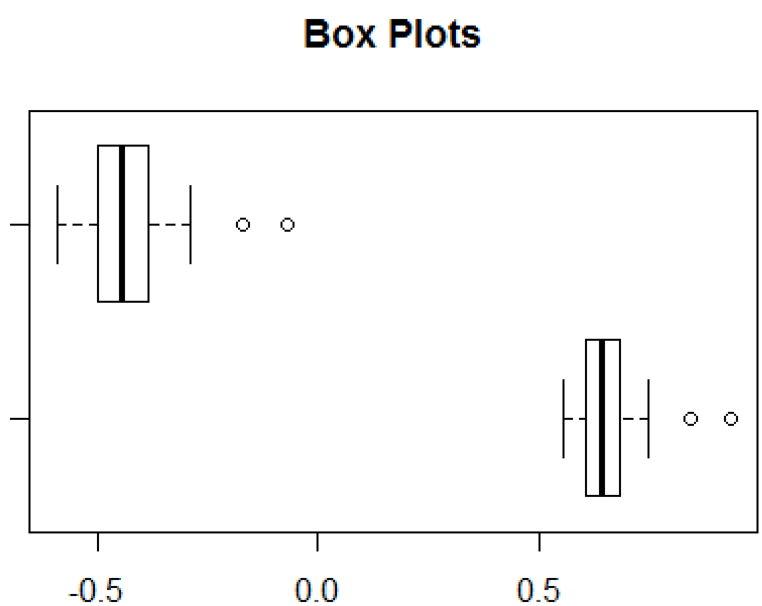
```r
# Question 3 (10.5 Problem Set D)
r<- c(0.693,0.662,0.690,0.606,0.570,0.749,0.672,0.628,0.609,0.844,0.654,
      0.615,0.668,0.601,0.576,0.670,0.606,0.611,0.553,0.933)
z=log(r)
plot(density(r))
```




```r
plot(density(z))
```




```
boxplot(r,z,horizontal = TRUE,main="Box Plots")
```




```
s=IQR(r)/sqrt(var(r))
s
```

```
## [1] 0.7620725

t=IQR(z)/sqrt(var(z))
t

## [1] 0.8544223

# 1)Both the ratios and the log of the ratios are very similar when tested
# for normality but log of ratios behave like normal distribution. This can
# be easily seen by looking at the density plot or boxplot. In all the
# cases the log of the ratios is slightly better with respect to normal
# distribution i.e. more shifted to right, In General Density plot of
# ratios has two bumps where as Density plot of log of ratios has only one
# and after looking at IQR to Stdev ratio we can say that log of ratios
# is closer to what we expect to be normal deviation.
#______________________________________________________________________
# 2) I would use the log of the ratios for which an assumption of
# normality seems more plausible. Therefore the mean we would like to test
# now is
log(0.618034)

## [1] -0.4812118

# For the hypothesis testing, from the point of view of the anthropologist
# would be:
# H0 : μ =-0.4812. vs. H1 : μ!=0.4812
# One could argue that the anthropologist wants to minimize Type I error,
# i.e., that the Shoshoni civilization actually used golden rectangles
# but the test shows otherwise. This is why in the test H0 represent
# the golden ratio.
# TO Calculate the Student's 1-sample t-test ,we need mean
m=mean(z)
m

## [1] -0.4230678

st= sqrt(var(z))
st

## [1] 0.1287264

tn= (m+0.4812)/(st/sqrt(20))
tn

## [1] 2.019596

# p = 2 * pt(-2.02, df = 19) = 0.05771 > 0.05 = α fail to reject H0
y<- sort(r)
# 3) To build a confidence interval with confidence 0.90, the following
# needs to hold: 1- α = 0.90 =α/2 = 0.05
k=qbinom(0.05, 20, 0.5)
# By Experimentation
1- pbinom(k,20,0.5)
```

```
## [1] 0.9423409
```

```
# We can construct a confidence interval of 94% which is very close to
# 95% and any other choice would be way off the value.
# The form of interval is  (sorting the values) :
# (x(k+1), x(n-k))= (x7, x14)
y[7]
```

```
## [1] 0.609
```

```
y[14]
```

```
## [1] 0.67
```

Note:- Question 2 Discussed with Krish Mahajan.