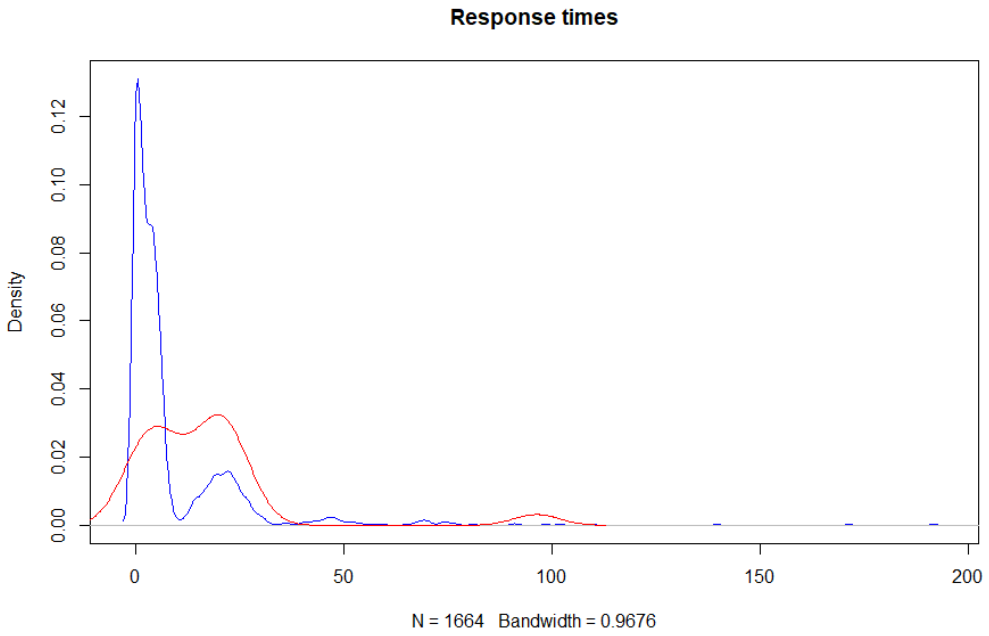
106073401

BACS HW (Week 5)

1. a) Divide customers into 2 groups and visualizing the ILEC & CLEC customers.

```
> time<-read.csv("verizon.csv",header = TRUE)
> ILEC<-time$Time[time$Group=="ILEC"]
> CLEC<-time$Time[time$Group=="CLEC"]
> plot(density(ILEC),col="blue",main="Response times")
> lines(density(CLEC),col="red")
```



b) Test the *difference between the mean of ILEC and mean of CLEC* by t.test function

```
> t.test(CLEC,ILEC,"greater",conf.level=0.99)

        Welch Two Sample t-test

data:  CLEC and ILEC
t = 1.9834, df = 22.346, p-value = 0.02987
alternative hypothesis: true difference in means is greater than 0
99 percent confidence interval:
 -2.130858       Inf
sample estimates:
mean of x mean of y
16.509130  8.411611
```

(i) In this case, H_0: The mean of CLEC <=ILEC

H_1:The mean of CLEC > ILEC

(ii) Since p>0.01, cannot reject H_0.

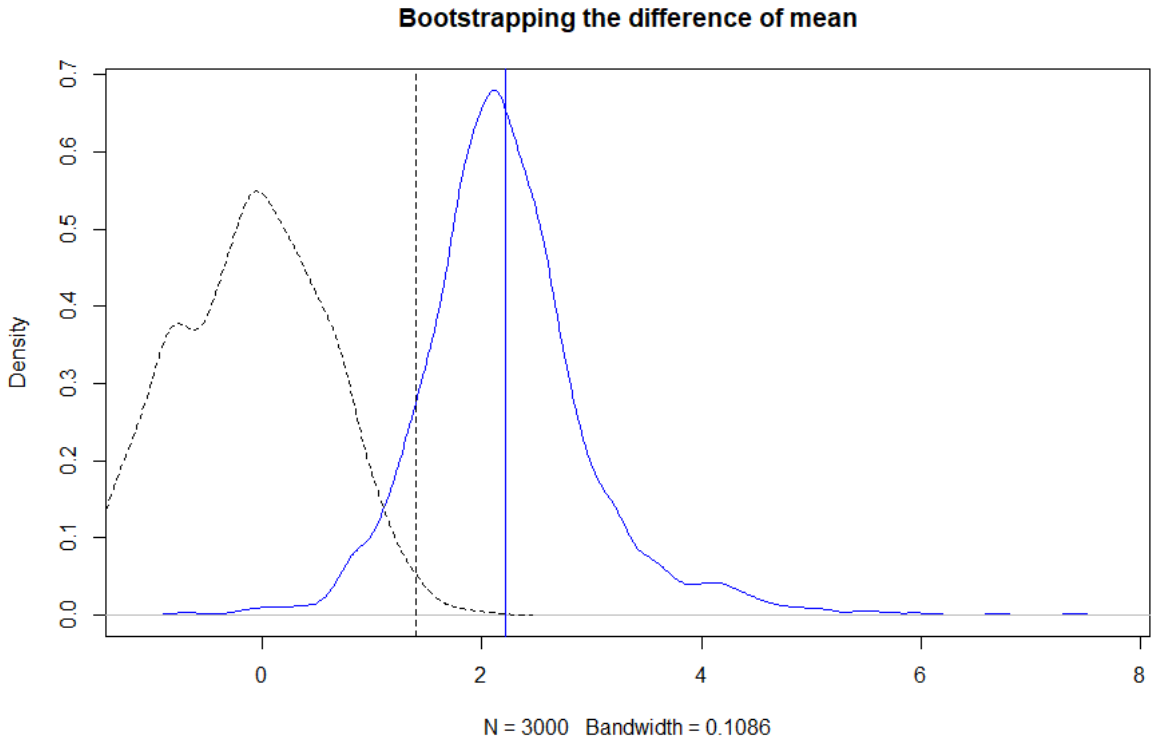
106073401

c) Test the *alternative and null values of t* by bootstrapping compare and compare with t.test function.

```
> boot_mean_diff<-function(CLEC,ILEC){
 resample_CLEC<-sample(CLEC,length(CLEC),replace=TRUE)
 resample_ILEC<-sample(ILEC,length(ILEC),replace=TRUE)
 t_test_alt<- t.test(resample_CLEC,resample_ILEC, "greater",conf.level = 0.99)
 t_test_null<-t.test(resample_CLEC,CLEC)
 return(c(t_test_alt$statistic,t_test_null$statistic))
  }
> t_test<-replicate(3000,boot_mean_diff(CLEC,ILEC))
> t_alt<-t_test[1,]
> t_null<-t_test[2,]
```

(i) Plot the null hypothesis and alternative hypothesis, show the 1% rejection zone.

```
> plot(density(t_alt),col="blue",main="Bootstrapping the difference of mean")
> lines(density(t_null),lty="dashed")
> abline(v=quantile(t_null,c(0.99)),lty="dashed")
> abline(v=mean(t_alt),col="blue")
```



(ii) As the alternative t-value largely outside the 1% rejection zone, we could reject the null hypothesis.

106073401

2. Test whether the variance of ILEC & CLEC response times are different.

a) The null & alternative hypothesis

H_0: There is no difference between the variance of ILEC & CLEC

H_1: The variance between ILEC & CLEC are not the same.

b)

```
#Identify which variance is higher
> var(CLEC)>var(ILEC)
[1] TRUE
```

(i) F-statistic of the ratio of variance=1.762717

```
> var.test(CLEC,ILEC,alternative="greater")

        F test to compare two variances

data:  CLEC and ILEC
F = 1.7627, num df = 22, denom df = 1663, p-value = 0.01582
alternative hypothesis: true ratio of variances is greater than 1
95 percent confidence interval:
 1.138356      Inf
sample estimates:
ratio of variances
          1.762717
```

(ii) The cut-off value of F=1.548476

```
> qf(p=0.95,df1=length(CLEC)-1,df2=length(ILEC)-1)
[1] 1.548476
```

(iii) We can reject the null hypothesis since F-value is greater than the cut-off value.

c) Bootstrapping

```
> boot_f_stat<-function(CLEC,ILEC){
    resample_CLEC<-sample(CLEC,length(CLEC),replace=TRUE)
    resample_ILEC<-sample(ILEC,length(ILEC),replace=TRUE)
    f_alt<-var(resample_CLEC)/var(resample_ILEC)
    f_null<-var(resample_CLEC)/var(CLEC)
    c(f_alt,f_null)
  }
```

(i) H_0: The f-value <=1

H_1:The f-value>1

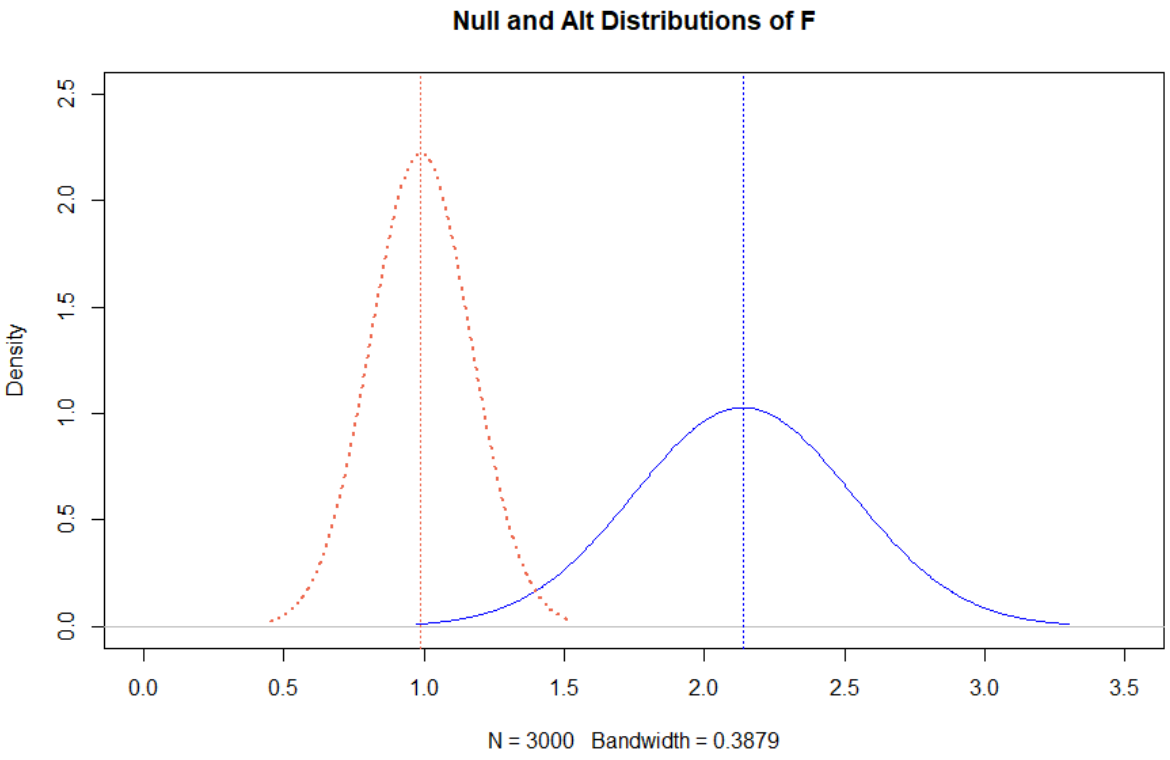
106073401

(ii) The 95% cutoff value is 0.9890.

```
> f_stats <- replicate(3000, boot_f_stat(CLEC,ILEC))
> f_alts <- f_stats[1,]
> f_nulls <- f_stats[2,]
> quantile(f_nulls, probs=0.95)
      95%
0.9890281
```

(iii) Visualize H_0, H_1 and the cut off value.

```
> plot(density(f_alts), col="blue",ylim=c(0,2.5), xlim=c(0,3.5),
       main="Null and Alt Distributions of F")
> lines(density(f_nulls),col="coral2", lwd=2, lty="dotted")
> abline(v=quantile(f_nulls, probs=0.95),lty="dotted", col="coral2")
> abline(v=median(f_alts),lty="dotted", col="blue")
```



(iv) We can reject the null hypothesis since F-value is greater than the cut-off value.