BACS HW8

106070020

2021年4月17日

Question 1

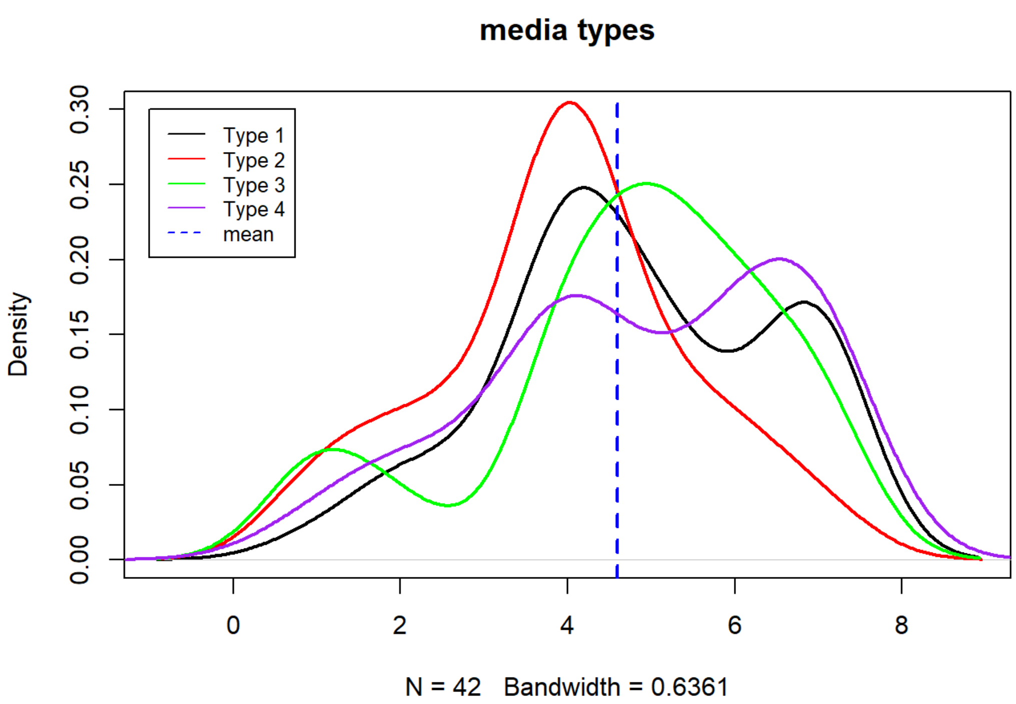
(a)

```
#Q1(a)

data1<-read.csv("C:/Users/eva/Desktop/作業 上課資料(清大)/大四下/BACS/HW8 BACS/pls-media/pls-media1.csv")
data2<-read.csv("C:/Users/eva/Desktop/作業 上課資料(清大)/大四下/BACS/HW8 BACS/pls-media/pls-media2.csv")
data3<-read.csv("C:/Users/eva/Desktop/作業 上課資料(清大)/大四下/BACS/HW8 BACS/pls-media/pls-media3.csv")
data4<-read.csv("C:/Users/eva/Desktop/作業 上課資料(清大)/大四下/BACS/HW8 BACS/pls-media/pls-media4.csv")
means<-c(mean(data1$INTEND.0),mean(data2$INTEND.0),mean(data3$INTEND.0),mean(data4$INTEND.0))
means1<-as.data.frame(means)
rownames(means1)<-c('type1','type2','type3','type4')
means1
```

```
##          means
## type1 4.809524
## type2 3.947368
## type3 4.725000
## type4 4.891304
```

(b)

```
#(b)
{plot(density(data1$INTEND.0),main='media types',lwd=2, ylim=c(0,0.3))
abline(v=mean(means),lty=2,col="blue",lwd=2)
lines(density(data2$INTEND.0),lwd=2,col='red')
lines(density(data3$INTEND.0),lwd=2,col='green')
lines(density(data4$INTEND.0),lwd=2,col='purple')
legend(-1,0.3, legend=c("Type 1", "Type 2", "Type 3", "Type 4", "mean"),
       col=c("black", "red", 'green', 'purple', 'blue'), lty=c(1,1,1,1,2), cex=0.8)
}
```



(c)

> For the visualization, the means of the four media types are similar, though they have different distribution. Thus, it is hard to distinguish that the media type makes a difference on intention to share.

Question 2

(a)

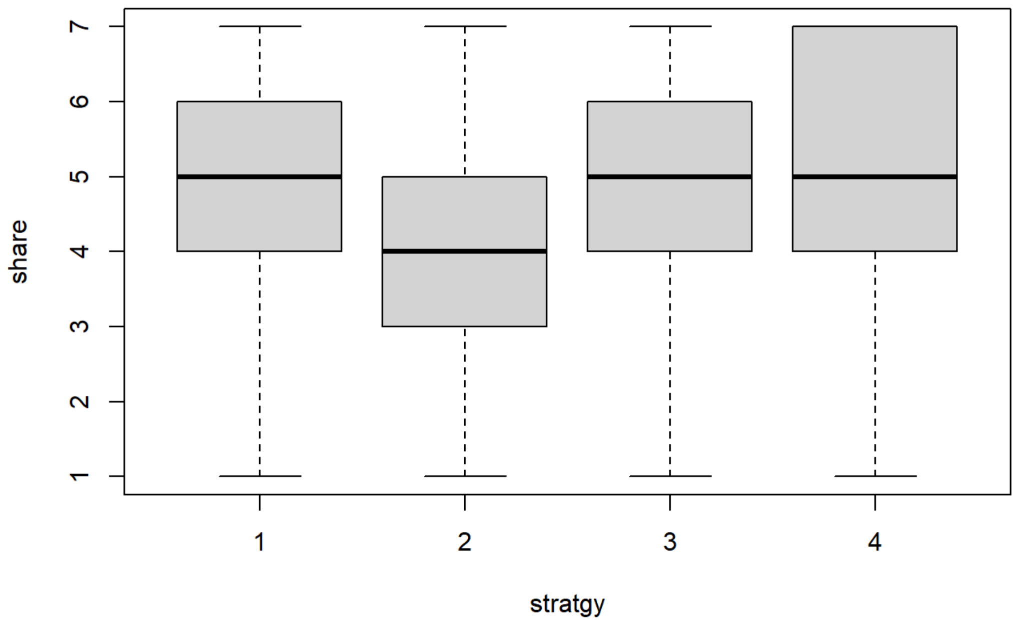
```
#Q2(a)

md1<-data.frame(stratgy=rep(1,length(data1$INTEND.0)),share=data1$INTEND.0)
md2<-data.frame(stratgy=rep(2,length(data2$INTEND.0)),share=data2$INTEND.0)
md3<-data.frame(stratgy=rep(3,length(data3$INTEND.0)),share=data3$INTEND.0)
md4<-data.frame(stratgy=rep(4,length(data4$INTEND.0)),share=data4$INTEND.0)
mds<-rbind(md1,md2,md3,md4)

oneway.test(mds$share ~ mds$stratgy, var.equal = T)
```

```
##
##  One-way analysis of means
##
## data:  mds$share and mds$stratgy
## F = 2.6167, num df = 3, denom df = 162, p-value = 0.05289
```

```
summary(aov(mds$share ~ factor(mds$stratgy)))
```

```
##                      Df Sum Sq Mean Sq F value Pr(>F)
## factor(mds$stratgy)   3   22.5   7.508   2.617 0.0529 .
## Residuals           162  464.8   2.869
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
boxplot(share~stratgy,data=mds)
```



H0: The means are equal, H1: The means are not equal

(b)

```
#(b)
m<-mean(means)

sstr1=length(data1$INTEND.0)*((mean(data1$INTEND.0)-m)^2)
sstr2=length(data2$INTEND.0)*((mean(data2$INTEND.0)-m)^2)
sstr3=length(data3$INTEND.0)*((mean(data3$INTEND.0)-m)^2)
sstr4=length(data4$INTEND.0)*((mean(data4$INTEND.0)-m)^2)
sstr=sum(sstr1,sstr2,sstr3,sstr4)
df_mstr=4-1
mstr=sstr/df_mstr

sse1<-sum((length(data1$INTEND.0)-1)*var(data1$INTEND.0))
sse2<-sum((length(data2$INTEND.0)-1)*var(data2$INTEND.0))
sse3<-sum((length(data3$INTEND.0)-1)*var(data3$INTEND.0))
sse4<-sum((length(data4$INTEND.0)-1)*var(data4$INTEND.0))
sse<-sum(sse1,sse2,sse3,sse4)
mse<-sse/162

F_score<-mstr/mse
F_score
```

```
## [1] 2.625303
```

(c)

```
#(c)cutoff
cutoff<-c(qf(p=0.95,df1=3, df2=162), qf(p=0.99,df1=3, df2=162))
cutoff
```

```
## [1] 2.660406 3.904807
```

(d)

```
#(d)
oneway.test(mds$share ~ mds$stratgy, var.equal = T)
```

```
##
##  One-way analysis of means
##
## data:  mds$share and mds$stratgy
## F = 2.6167, num df = 3, denom df = 162, p-value = 0.05289
```

According to the p-value from ANOVA test, the p-value is 0.05289, which is larger than 0.05 and 0.01. Therefore, we should accept H0 under 95% and 99% confidence interval.

(e)

```
#(e)
md11<-data.frame(stratgy=rep("M1",length(data1$INTEND.0)),share=data1$INTEND.0)
md22<-data.frame(stratgy=rep("M2",length(data2$INTEND.0)),share=data2$INTEND.0)
md33<-data.frame(stratgy=rep("M3",length(data3$INTEND.0)),share=data3$INTEND.0)
md44<-data.frame(stratgy=rep("M4",length(data4$INTEND.0)),share=data4$INTEND.0)
mds1<-rbind(md11,md22,md33,md44)

library(car)
```

```
## Warning: package 'car' was built under R version 4.0.5
```

```
leveneTest(share~stratgy,data=mds1)
```

```
## Warning in leveneTest.default(y = y, group = group, ...): group coerced to
## factor.
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value Pr(>F)
## group   3  1.5403 0.2061
##       162
```

> The p-value of Levene Test is larger than 0.05, so we can say that the variance of 4 data are same. The assumption of ANOVA is reasonable.

Question 3

(a)

```
#Q3(a)
boot_anova<-function(t1, t2, t3, t4, treat_nums) {
  null_grp1 =sample(t1-mean(t1), replace=TRUE)
  null_grp2 =sample(t2-mean(t2), replace=TRUE)
  null_grp3 =sample(t3-mean(t3), replace=TRUE)
  null_grp4 =sample(t4-mean(t4), replace=TRUE)
  null_values= c(null_grp1, null_grp2, null_grp3,null_grp4)
  alt_grp1 = sample(t1, replace=TRUE)
  alt_grp2 = sample(t2, replace=TRUE)
  alt_grp3 = sample(t3, replace=TRUE)
  alt_grp4 = sample(t4, replace=TRUE)
  alt_values= c(alt_grp1, alt_grp2, alt_grp3,alt_grp4)
  c(oneway.test(null_values~treat_nums,var.equal=TRUE)$statistic,
    oneway.test(alt_values~treat_nums,var.equal=TRUE)$statistic)
}

f_val<-replicate(5000,boot_anova(data1$INTEND.0, data2$INTEND.0, data3$INTEND.0, data4$INTEND.0, mds$stratgy))
f_nulls<-f_val[1,]
f_alts<-f_val[2,]

f_nulls
```

```
##  [1] 1.43919945 0.05960848 0.44200768 0.62404123 1.12379584 0.42291467
##  [7] 0.08408164 1.12507002 0.85922666 0.19647067
##  [ reached getOption("max.print") -- omitted 4990 entries ]
```

```
f_alts
```

```
##  [1] 4.965267 4.028072 3.632107 1.465042 3.768576 2.923776 2.382294 1.701736
##  [9] 1.087080 6.683959
##  [ reached getOption("max.print") -- omitted 4990 entries ]
```

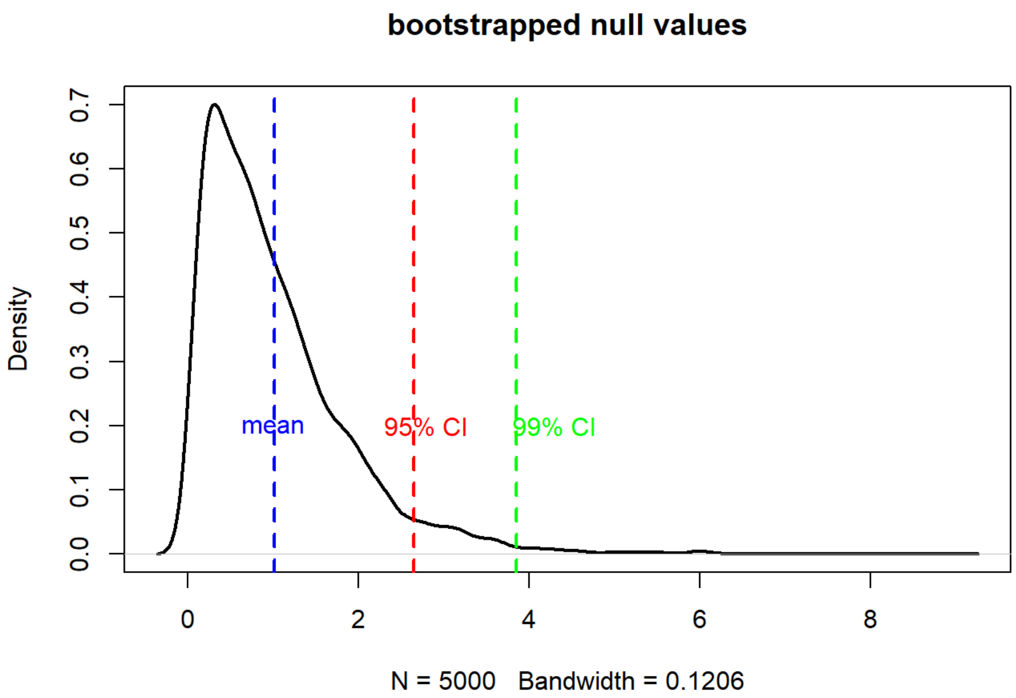
(b)

```
#(b)
cutoffs<-c(quantile(f_nulls,0.95), quantile(f_nulls,0.99))
cutoffs
```

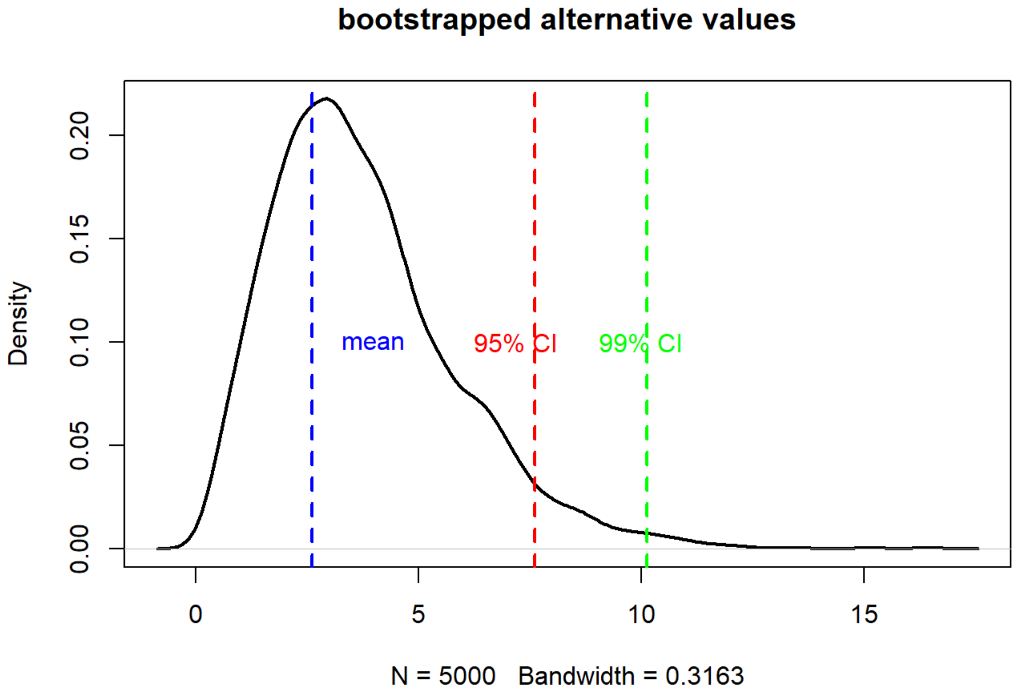
```
##      95%      99%
## 2.644979 3.850201
```

(c)

```
#(c)

{plot(density(f_nulls),main='bootstrapped null values',lwd=2)
  abline(v=mean(f_nulls),lty=2,col="blue",lwd=2)
  abline(v=quantile(f_nulls,0.95),lty=2,col="red",lwd=2)
  abline(v=quantile(f_nulls,0.99),lty=2,col="green",lwd=2)
  text(x=1,y=0.2, 'mean',col='blue')
  text(x=2.8,y=0.2, '95% CI',col='red')
  text(x=4.3,y=0.2, '99% CI',col='green')
}
```




```
anova<-oneway.test(mds$share ~ mds$stratgy, var.equal = T)
{plot(density(f_alts),main='bootstrapped alternative values',lwd=2)
  # abline(v=mean(f_alts),lty=2,col="blue",lwd=2)
  abline(v=quantile(f_alts,0.95),lty=2,col="red",lwd=2)
  abline(v=quantile(f_alts,0.99),lty=2,col="green",lwd=2)
  abline(v=anova$statistic,lty=2,col="blue",lwd=2)
  text(x=4,y=0.1, 'mean',col='blue')
  text(x=7.2,y=0.1, '95% CI',col='red')
  text(x=10,y=0.1, '99% CI',col='green')
}
```



(d)

> According to the bootstrap outcome, the mean of the F-value is less than the 95% and 99% cutoffs, so we should accept H0 under 95% CI and 99% CI.