

STA32HW5

Rickey Huang

3/27/2022

Question 1

Part c

```
library(faraway)
teengamb <- faraway::teengamb

# Fit both models
lm0 <- lm(gamble~sex+income, data = teengamb)
lmo <- lm(gamble~income,data = teengamb)
# Compute the SSE of the both models
SSEO <- (t(lm0$residuals)%*%lm0$residuals)[1]
SSEo <- (t(lmo$residuals)%*%lmo$residuals)[1]
# Define n, p, and q
n <- dim(model.matrix(lmo))[1]
p <- dim(model.matrix(lm0))[2]
q <- dim(model.matrix(lmo))[2]
# Compute the F statistic
Fstat <- ((SSEo-SSEO)/(p-q))/(SSEO/(n-p))
Fstat

## [1] 10.09598
```

Part d

```
# Compute the p value
pval <- 1 - pf(Fstat,p-q,n-p)
pval

## [1] 0.00271732
```

Part e

```
# Compute the Fstar
Fstar <- qf(0.95,p-q,n-p)
Fstar

## [1] 4.061706
```

Question 2

```
# Using the same data set and models, conduct a permutation test for the same hypothe-  
# ses. How do your p-values compare?  
# Compute the original F statistic  
forg <- summary(lm0)$fstat  
# Initialize the p-value  
pval2=0  
# Create the for loop  
for (i in 1:4000){  
  # Fit the model with the permuting  
  lmnew <- lm(gamble~sample(sex)+income, data = teengamb)  
  # Find out whether the F stistic is bigger  
  if(summary(lmnew)$fstat > forg){  
    # if bigger, add it to the p value  
    pval=pval+1/4000  
  }  
}
```

```
## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and  
## only the first element will be used
```

```
## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and  
## only the first element will be used
```

```
## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and  
## only the first element will be used
```

```
## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and  
## only the first element will be used
```

```
## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and  
## only the first element will be used
```

```
## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and  
## only the first element will be used
```

```
## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and  
## only the first element will be used
```

```
## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and  
## only the first element will be used
```

```
## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and  
## only the first element will be used
```

```
## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and  
## only the first element will be used
```

```
## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and  
## only the first element will be used
```

```
## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and  
## only the first element will be used
```


[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]


```
## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and
## only the first element will be used

## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and
## only the first element will be used

## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and
## only the first element will be used

## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and
## only the first element will be used

## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and
## only the first element will be used

## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and
## only the first element will be used

## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and
## only the first element will be used

## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and
## only the first element will be used

## Warning in if (summary(lmnew)$fstat > forg) {: the condition has length > 1 and
## only the first element will be used

# return the p-value
pval

## [1] 0.00496732
```

Question 3

Part c

```
library(faraway)
teengamb <- faraway::teengamb

# Fit both models
lm02 <- lm(gamble~verbal + status +income, data = teengamb)
lmo2 <- lm(gamble~income,data = teengamb)
# Compute the SSE of the both models
SSE02 <- (t(lm02$residuals)%*%lm02$residuals)[1]
SSEo2 <- (t(lmo2$residuals)%*%lmo2$residuals)[1]
# Define n, p, and q
n2 <- dim(model.matrix(lmo2))[1]
p2 <- dim(model.matrix(lm02))[2]
q2 <- dim(model.matrix(lmo2))[2]
# Compute the F statistic
```



```
Fstat2 <- ((SSEo2-SSE02)/(p2-q2))/(SSE02/(n2-p2))
Fstat2
```

```
## [1] 2.245865
```

Part d

```
# Compute the p value
pval3 <- 1 - pf(Fstat2,p2-q2,n2-p2)
pval3
```

```
## [1] 0.1181095
```

Part e

```
# Compute the Fstar
Fstar2 <- qf(0.95,p2-q2,n2-p2)
Fstar2
```

```
## [1] 3.21448
```

Question 5

Part a

```
# Fit the model
lm1 <- lm(gamble~sex+income, data = teengamb)
intsex <- confint(lm1,2,level=0.95)
intsex
```

```
##           2.5 %      97.5 %
## sex -35.35662 -7.912162
```

Part b

```
# Construct a 95% confidence level boot-strapped interval for the coefficient of sex.
# Find the residuals and fitted value of the original model
resids <- lm1$residuals
fit <- lm1$fitted
# Construct a vector with 4000 entries
betasexs <- numeric(4000)
# Construct a for loop to find the boot-strapped interval
for (i in 1:4000){
  # Randomize the residuals to create 4000 random response variables with errors
  fitboot <- fit + sample(resids,rep = TRUE)
  # Fit new models with 4000 error terms
  lmboot <- lm(fitboot ~ sex + income, data = teengamb)
  # Store the betas for the variable sex
  betasexs[i] <- lmboot$coeff[2]
}
# Find the 0.025 and 0.975 quantile of the betas to construct a boot-strapped confidence interval
cbind(quantile(betasexs,0.025),quantile(betasexs,0.975))
```

```
##           [,1]      [,2]
```

2.5% -34.87431 -8.514709