Stat500(Section002): Homework #3

Due on Oct. 6, 2021 at 11:59pm

Instructor:Naisyin Wang

Tiejin Chen **tiejin@umich.edu**

Problem 1

Part a
We get $H0 : \beta_{elevation} \leq 0, H1 : \beta_{elevation} > 0$. We can calculate the t-value of $\beta_{elevation}$ is 7.725. We use the following code to get the p-value:

R_code.Rmd
```
library(faraway)
data(gala)
linear = lm(Endemics~.-Species,data=gala)
1-pt(7.725,df=24)
```

We can the result p-value is 2.91431e-08 which is less than 0.01.Hence we reject H0, and accept H1, which means an island with a large highest elevation level tends to have more endemic species.
Part b

R_code.Rmd
```
confint(linear,level=0.99)
```

Using the code, we can get the 99% confifence interval for $\beta_{elevation}$ is (0.05328709,0.113773003),for$\beta Nearest$ is (-0.56891008,0.619257004).
We use the following code to calculate the CI of $\beta_{elevation}$ by ourselves.

R_code.Rmd
```
beta_E = linear$coefficients['Elevation']
print(beta_E-qt(0.995,df = 24)*0.010813)
print(beta_E+qt(0.995,df = 24)*0.010813)
```

And we get the same result as confint function.
Part c
To this test, we should use f-test. And, we use the following code:

R_code.Rmd
```
H0_model = lm(Endemics~Area+Scruz+Adjacent, data = gala)
anova(linear,H0_model)
```

And we get the f-value is 31.789, the p-value is 1.794e-07 which is less than 0.05. Hence, we reject H0,and accept H1:One of $\beta_{Nearest}$ and β_{Scruz} is not zero. That is to say at least one of the predictors have an effect on the response.
If (0,0) is inside the confidence region, then we should accpet H0 instead of H1. Hence (0,0) is not in the region.

Problem 2

Part a
We get $H0 : \beta_{ratio} \geq 0, \beta_{ratio} < 0$. And we use t-test and following code:

R_code.Rmd
```
data(sat)
```

```
tmp = lm(total~takers+ratio+salary,sat)
t = -4.6394/2.1215
pt(t,df =46)
```

get the t-value is -2.187, p-value is 0.0169382, which is larger than 0.01. Hence we will accept H0. That is, a higher value of average pupil/teacher ratio (ratio) does not tend to lead to a lower sat score.

Part b

R_code.Rmd

```
summary(tmp)
```

```
Call:
lm(formula = total ~ takers + ratio + salary, data = sat)

Residuals:
    Min      1Q  Median      3Q     Max 
-89.244 -21.485  -0.798  17.685  68.262 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1057.8982    44.3287  23.865   <2e-16 ***
takers        -2.9134     0.2282 -12.764   <2e-16 ***
ratio         -4.6394     2.1215  -2.187   0.0339 *  
salary         2.5525     1.0045   2.541   0.0145 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 32.41 on 46 degrees of freedom
Multiple R-squared:  0.8239,	Adjusted R-squared:  0.8124 
F-statistic: 71.72 on 3 and 46 DF,  p-value: < 2.2e-16
```

We can get the R^2 is 0.8239. And σ^2 is $32.41 \times 32.41 = 1050.4$.

Problem 3

R_code.Rmd

```
set.seed(42)
nrep = 5000
tnull=rep(NA,nrep)
for (ii in 1:nrep){
lmod=lm(total~takers+sample(ratio)+salary,sat);
tnull[ii]=coef(summary(lmod))[3,3]
mean(abs(tnull)> abs(coef(summary(tmp))[3,3]))
```

With the code, we could get the p-value of permutation test is 0.0332 which is less than 0.05. Hence, we reject H0 and accpet H1 that $\beta_{ratio} \neq 0$.

Problem 3 continued on next page. . .

R_code.Rmd

```
mean(tnull< coef(summary(tmp))[3,3])
```

Using this code, we can get the p-value of $H0 : \beta_{ratio} \geq 0, \beta_{ratio} < 0$ is 0.016 which is larger than 0.01. Hence we accept H0, which is the same as t-test result in Problem2(a).