Hm3stat631

Jessica Grover

3/1/2022

#Problem 2 Reading the table

cars<-read.table("car.dat",header = TRUE)  
cars

## type response  
## 1 1 4.705398  
## 2 1 4.120209  
## 3 1 4.798508  
## 4 1 4.393436  
## 5 2 4.127314  
## 6 2 4.404476  
## 7 2 4.693979  
## 8 2 4.218640  
## 9 3 4.546443  
## 10 3 4.750001  
## 11 3 4.776935  
## 12 3 4.362813  
## 13 4 4.996235  
## 14 4 4.490389  
## 15 4 4.848673  
## 16 4 4.547607  
## 17 5 4.879283  
## 18 5 5.230027  
## 19 5 4.936292  
## 20 5 4.768095  
## 21 6 5.976583  
## 22 6 6.310510  
## 23 6 5.906525  
## 24 6 6.200180

#a.

cars$ftype<-as.factor(cars$type)  
cars

## type response ftype  
## 1 1 4.705398 1  
## 2 1 4.120209 1  
## 3 1 4.798508 1  
## 4 1 4.393436 1  
## 5 2 4.127314 2  
## 6 2 4.404476 2  
## 7 2 4.693979 2  
## 8 2 4.218640 2  
## 9 3 4.546443 3  
## 10 3 4.750001 3  
## 11 3 4.776935 3  
## 12 3 4.362813 3  
## 13 4 4.996235 4  
## 14 4 4.490389 4  
## 15 4 4.848673 4  
## 16 4 4.547607 4  
## 17 5 4.879283 5  
## 18 5 5.230027 5  
## 19 5 4.936292 5  
## 20 5 4.768095 5  
## 21 6 5.976583 6  
## 22 6 6.310510 6  
## 23 6 5.906525 6  
## 24 6 6.200180 6

model1<-lm(response~ftype,data=cars)  
anova(model1)

## Analysis of Variance Table  
##   
## Response: response  
## Df Sum Sq Mean Sq F value Pr(>F)   
## ftype 5 7.9958 1.59916 29.193 4.887e-08 \*\*\*  
## Residuals 18 0.9860 0.05478   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The null hypothesis is rejected. There is a significant differance between means oil use in given cars as p-value is less than 0.05. Hence we can conclude that data provides significant evidence that there is a difference in the mean oil use of the cars.

#b.

library(emmeans)  
lsmcars<-lsmeans(model1,~ftype)  
print(lsmcars)

## ftype lsmean SE df lower.CL upper.CL  
## 1 4.50 0.117 18 4.26 4.75  
## 2 4.36 0.117 18 4.12 4.61  
## 3 4.61 0.117 18 4.36 4.85  
## 4 4.72 0.117 18 4.47 4.97  
## 5 4.95 0.117 18 4.71 5.20  
## 6 6.10 0.117 18 5.85 6.34  
##   
## Confidence level used: 0.95

levels(cars$ftype)

## [1] "1" "2" "3" "4" "5" "6"

#b1.

contrast(lsmcars,list("ImportedWithExported"=c(-1/4,-1/4,1/2,-1/4,-1/4,1/2)))

## contrast estimate SE df t.ratio p.value  
## ImportedWithExported 0.719 0.101 18 7.093 <.0001

summary(contrast(lsmcars,list("ImportedWithExported"=c(-1/4,-1/4,1/2,-1/4,-1/4,1/2))),infer=c(T,T),level=0.95,side="two-sided")

## contrast estimate SE df lower.CL upper.CL t.ratio p.value  
## ImportedWithExported 0.719 0.101 18 0.506 0.932 7.093 <.0001  
##   
## Confidence level used: 0.95

Since we have interval of (0.506,0.932), it is clear that there is no 0. It means, the usage of oils of imported cars is more than the exported cars.

#b2.

contrast(lsmcars,list("CheapWithExpensive"=c(1/3,1/3,1/3,-1/3,-1/3,-1/3)))

## contrast estimate SE df t.ratio p.value  
## CheapWithExpensive -0.766 0.0955 18 -8.017 <.0001

summary(contrast(lsmcars,list("CheapWithExpensive"=c(1/3,1/3,1/3,-1/3,-1/3,-1/3))),infer = c(T,T),level = 0.95,side="two-sided")

## contrast estimate SE df lower.CL upper.CL t.ratio p.value  
## CheapWithExpensive -0.766 0.0955 18 -0.967 -0.565 -8.017 <.0001  
##   
## Confidence level used: 0.95

Since the interval is (-0.967,-0.565), it does not involve 0. The value is negative so we can conclude that oil usage of cheaper cars are less than the expensive cars.