hm2stat631

Jessica Grover

2/16/2022

#Exercise 3.1 The problem specifies that rats were given four different types of diet. Therefore, treatments are 4. According to the table given, responses are 29.

Inserting data

group1<-c(3.52,3.36,3.57,4.19,3.88,3.76,3.94)  
group2<-c(3.47,3.73,3.38,3.87,3.69,3.51,3.35,3.64)  
group3<-c(3.54,3.52,3.61,3.76,3.65,3.51)  
group4<-c(3.74,3.83,3.87,4.08,4.31,3.98,3.86,3.71)  
  
diets<-c(group1,group2,group3,group4)

Computing overall mean

mean(diets)

## [1] 3.718276

Computing treatment effects

trt1<-mean(group1)-mean(diets)  
trt2<-mean(group2)-mean(diets)  
trt3<-mean(group3)-mean(diets)  
trt4<-mean(group4)-mean(diets)  
trts<-c(trt1,trt2,trt3,trt4)  
trts

## [1] 0.02743842 -0.13827586 -0.11994253 0.20422414

trt<-c(rep(1,7),rep(2,8),rep(3,6),rep(4,8))  
df<-data.frame(diets,trt)

Let us set up the hypothesis test.

H0: All the four group’s means are equal. H1: Atleast one of the mean is different.

Let us now factor “df”. It will be converted into factor variable. A factor variable is treated as qualitative material by R.

df$ftrt<-as.factor(df$trt)  
df

## diets trt ftrt  
## 1 3.52 1 1  
## 2 3.36 1 1  
## 3 3.57 1 1  
## 4 4.19 1 1  
## 5 3.88 1 1  
## 6 3.76 1 1  
## 7 3.94 1 1  
## 8 3.47 2 2  
## 9 3.73 2 2  
## 10 3.38 2 2  
## 11 3.87 2 2  
## 12 3.69 2 2  
## 13 3.51 2 2  
## 14 3.35 2 2  
## 15 3.64 2 2  
## 16 3.54 3 3  
## 17 3.52 3 3  
## 18 3.61 3 3  
## 19 3.76 3 3  
## 20 3.65 3 3  
## 21 3.51 3 3  
## 22 3.74 4 4  
## 23 3.83 4 4  
## 24 3.87 4 4  
## 25 4.08 4 4  
## 26 4.31 4 4  
## 27 3.98 4 4  
## 28 3.86 4 4  
## 29 3.71 4 4

Using ANOVA to do the analysis of variance,

model<-lm(diets~ftrt,data=df)  
anova(model)

## Analysis of Variance Table  
##   
## Response: diets  
## Df Sum Sq Mean Sq F value Pr(>F)   
## ftrt 3 0.57821 0.192736 4.6581 0.01016 \*  
## Residuals 25 1.03440 0.041376   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The overall mean is 3.718276

Fvalue is the division of two mean squares, i.e, 0.192736/0.041376

If H0 is true, it means that Fvalue should be around 1 but it is 4.6581 and p-value is 0.01016. It is clear that there are differences in the means of the diet and we can reject the null hypothesis.

#Exercise 3.3 The table data has moisture content of the silage. Let us first insert the data.

response<- c(80.5, 79.3, 79, 89.1, 75.7, 81.2, 77.8, 79.5, 77, 76.7, 77.2, 78.6)  
trt<- rep(1:4, each=3)  
df<- data.frame(response,trt)

Factor variable

df$ftrt<-as.factor(df$trt)  
df

## response trt ftrt  
## 1 80.5 1 1  
## 2 79.3 1 1  
## 3 79.0 1 1  
## 4 89.1 2 2  
## 5 75.7 2 2  
## 6 81.2 2 2  
## 7 77.8 3 3  
## 8 79.5 3 3  
## 9 77.0 3 3  
## 10 76.7 4 4  
## 11 77.2 4 4  
## 12 78.6 4 4

Using ANOVA to find variance,

model<-lm(response~ftrt,data=df)  
anova(model)

## Analysis of Variance Table  
##   
## Response: response  
## Df Sum Sq Mean Sq F value Pr(>F)  
## ftrt 3 36.18 12.06 0.9926 0.444  
## Residuals 8 97.20 12.15

Here, we can see that p-value is 0.444 which is greater than 0.05. Therefore, null hypothesis cannot be rejected. There is no sufficient evidence to conclude that all four treatment yield the same average moisture content.

#Exercise 3.4 The group means are given. Let us first insert the data.

means<-c(6.5,4.5,5.7,5.7,5.1)  
grand<-means  
x<-mean(grand)  
x<-rep(x,5)  
sst<-3\*sum((means-x)^2)  
mse<-0.75  
sse<-mse\*10  
mst<-sst/4  
fstat<-mst/mse  
pvalue<-pf(fstat,4,10,lower.tail = FALSE)  
  
response<-c("treatment","error","total")  
df<-c(4, 10, 14)  
ss<-c(sst,sse,NA)  
ms<-c(mst,mse,NA)  
f<-c(fstat,NA,NA)  
pr<-c(pvalue,NA,NA)  
table<-data.frame(response,df,ss,ms,f,pr)  
head(table)

## response df ss ms f pr  
## 1 treatment 4 6.72 1.68 2.24 0.1372519  
## 2 error 10 7.50 0.75 NA NA  
## 3 total 14 NA NA NA NA