

R Forward MTCARS

Yosef Guevara Salamanca

26/11/2020

```
data("mtcars")
attach(mtcars)
```

Paso 1. Modelo nulo

```
#SLR modelo nulo
```

```
mt.fit0 <- lm(mpg ~ 1)
anova(mt.fit0)
```

```
## Analysis of Variance Table
##
## Response: mpg
##           Df Sum Sq Mean Sq F value Pr(>F)
## Residuals 31  1126   36.324
```

```
SCE.0=anova(mt.fit0)[1,2]
cat("Suma de los cuadrados del error Modelo nulo: ", SCE.0)
```

```
## Suma de los cuadrados del error Modelo nulo:  1126.047
```

```
y_hat <- mean(mpg);
cat("y_hat:", y_hat )
```

```
## y_hat: 20.09062
```

```
# Calculo de la media
```

```
media <- y_hat
cat("La media es: ", media, "\n")
```

Calculo conceptual del SSRnulo

```
## La media es:  20.09062
```

```
# Caculo del las Sumas de cuadrados de la regresion

SCR.nulo=sum((mpg-media)^2)
cat("Las suma de cuadrados de la regresion es: ", SCR.nulo, "\n")
```

```
## Las suma de cuadrados de la regresion es: 1126.047
```

```
# F critico modelo nulo para el 10%

f.critico1<-qf(0.90,1,5)
cat("F_Critico modelo nulo: ", f.critico1 )
```

“Partial f Test”

```
## F_Critico modelo nulo: 4.06042
```

```
# Regresiones lineales necesarias

mt.fit11<-lm(mpg~ wt)
mt.fit12<-lm(mpg~ disp)
mt.fit13<-lm(mpg~ hp)

print("mpg ~ wt ")
```

TRES REGRESIONES LINEALES SIMPLES

```
## [1] "mpg ~ wt "
```

```
anova(mt.fit11)
```

```
## Analysis of Variance Table
##
## Response: mpg
##          Df Sum Sq Mean Sq F value    Pr(>F)
## wt         1  847.73   847.73  91.375 1.294e-10 ***
## Residuals 30  278.32     9.28
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print("mpg~ disp")
```

```
## [1] "mpg~ disp"
```

```
anova(mt.fit12)
```

```
## Analysis of Variance Table
##
## Response: mpg
##           Df Sum Sq Mean Sq F value    Pr(>F)
## disp       1  808.89   808.89  76.513 9.38e-10 ***
## Residuals  30  317.16    10.57
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print("mpg~ hp")
```

```
## [1] "mpg~ hp"
```

```
anova(mt.fit13)
```

```
## Analysis of Variance Table
##
## Response: mpg
##           Df Sum Sq Mean Sq F value    Pr(>F)
## hp         1  678.37   678.37  45.46 1.788e-07 ***
## Residuals  30  447.67    14.92
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
SCE.0<-anova(mt.fit0)[1,2]
SCE.11<-anova(mt.fit11)[2,2]
SCE.12<-anova(mt.fit12)[2,2]
SCE.13<-anova(mt.fit13)[2,2]
CME.11<-anova(mt.fit11)[2,3]
CME.12<-anova(mt.fit12)[2,3]
CME.13<-anova(mt.fit13)[2,3]
cbind(SCE.11,SCE.12,SCE.13,CME.11,CME.12,CME.13)
```

fparcial SHEATHER

```
##           SCE.11  SCE.12  SCE.13  CME.11  CME.12  CME.13
## [1,] 278.3219 317.1587 447.6743 9.277398 10.57196 14.92248
```

```
fparcial.11=(SCE.0-SCE.11)/CME.11
fparcial.12=(SCE.0-SCE.12)/CME.12
fparcial.13=(SCE.0-SCE.13)/CME.13
cbind(f.critico1,fparcial.11, fparcial.12, fparcial.13)
```

Cálculos de fparciales

```
##           f.critico1 fparcial.11 fparcial.12 fparcial.13
## [1,]      4.06042      91.37533      76.51266      45.4598
```

PASO 2 DE FORWARD CON fparcial. GUARDAR “SCE.11” “wt” SE EVALÚAN DOS MODELOS CON 2 COVARIABLES, CADA UNO

```
f.critico2<- qf(0.9,2,4)
f.critico2
```

```
## [1] 4.324555
```

MODELOS 21 Y 22

```
mt.fit21<-lm(mpg ~ wt + disp)
anova(mt.fit21)
```

con disp

```
## Analysis of Variance Table
##
## Response: mpg
##           Df Sum Sq Mean Sq F value    Pr(>F)
## wt         1  847.73   847.73  99.6586 6.861e-11 ***
## disp       1   31.64    31.64   3.7195  0.06362 .
## Residuals 29  246.68     8.51
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
SCE.21<-anova(mt.fit21)[3,2]
CME.21<-anova(mt.fit21)[3,3]
```

```
mt.fit22<-lm(mpg~ wt + hp)
SCE.22<-anova(mt.fit22)[3,2]
CME.22<-anova(mt.fit22)[3,3]
anova(mt.fit22)
```

con hp

```
## Analysis of Variance Table
##
## Response: mpg
##           Df Sum Sq Mean Sq F value    Pr(>F)
## wt         1  847.73   847.73 126.041 4.488e-12 ***
## hp         1   83.27    83.27  12.381  0.001451 **
## Residuals 29  195.05     6.73
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cbind(SCE.21,CME.21,SCE.22,CME.22)
```

```
##          SCE.21    CME.21    SCE.22    CME.22
## [1,] 246.6825  8.506293 195.0478  6.725785
```

```
fparcial.21<-(SCE.12-SCE.21)/CME.21
fparcial.22<-(SCE.12-SCE.22)/CME.22
cbind(f.critico2, fparcial.21, fparcial.22)
```

fparciales 21 y 22

```
##          f.critico2 fparcial.21 fparcial.22
## [1,]    4.324555    8.285182    18.15564
```

El modelo que cuyo r parcial es mayor al F critico es el modelo 22. Este modelo es mejor que el modelo que solo tiene a wt

conclusión: tanto el modelo con 21 como el modelo 22 son mejores que el modelo 11 pero de esos 22 es el mejor

PASO 2: MODELO CON Covariabes “disp” y “wt”

```
mt.fit22<-lm(mpg ~ wt + hp)
```

PASO 3 DE FORWARD ... (GUARDAR SCE.22) Sólo queda chequear si el modelo con las 3 covariables es mejor que el modelo sólo con estas dos PREVIAMENTE ESCOGIDAS.

```
fcrítico3<-qf(0.9, 3, 3)
fcrítico3
```

```
## [1] 5.390773
```

```
mt.fit3<-lm(mpg ~ wt + hp + disp)
anova(mt.fit3)
```

ajustar modelo con las 3 covariables

```
## Analysis of Variance Table
##
## Response: mpg
##          Df Sum Sq Mean Sq  F value    Pr(>F)
## wt         1  847.73   847.73  121.7305 1.052e-11 ***
## hp         1   83.27    83.27   11.9579  0.001758 **
## disp        1    0.06     0.06    0.0082  0.928507
## Residuals 28 194.99     6.96
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
SCE.3<-anova(mt.fit3)[4,2]  
CME.3<-anova(mt.fit3)[4,3]
```

obtener SCE.3 y CME.3

```
fparcial.3<-(SCE.22-SCE.3)/CME.3  
cbind(fcrítico3,fparcial.3)
```

CALCULAR fparcial.3

```
##      fcrítico3  fparcial.3  
## [1,]  5.390773 0.008196498
```

Como el fcrítico es mayor que el fparcial3 el modelo de 3 covariables no es mejor que el modelo con 2 variables dado por las variables wt y hp