

# Actividad 6 ANOVA

Elías Garza A01284041

3/9/2023

```
df = read.csv("C:\\Users\\elias\\Dropbox\\Carrera\\7mo Semestre\\Bloque 1\\Estadistica\\instituto_anova")
head(df)
```

```
##   i..sexo metodo calificacion
## 1  chico      1          10
## 2  chico      1           7
## 3  chico      1           9
## 4  chico      1           9
## 5  chico      1           9
## 6  chico      1          10
```

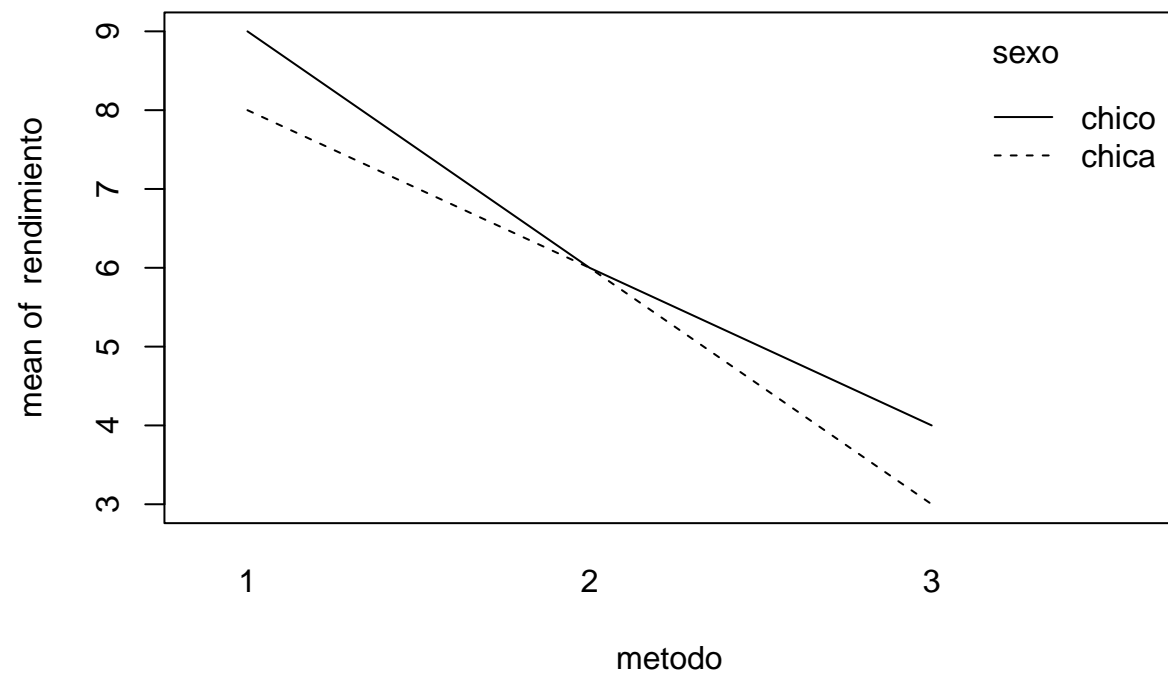
## Con interacción

```
sexo = df$i..sexo
metodo = df$metodo
rendimiento = df$calificacion

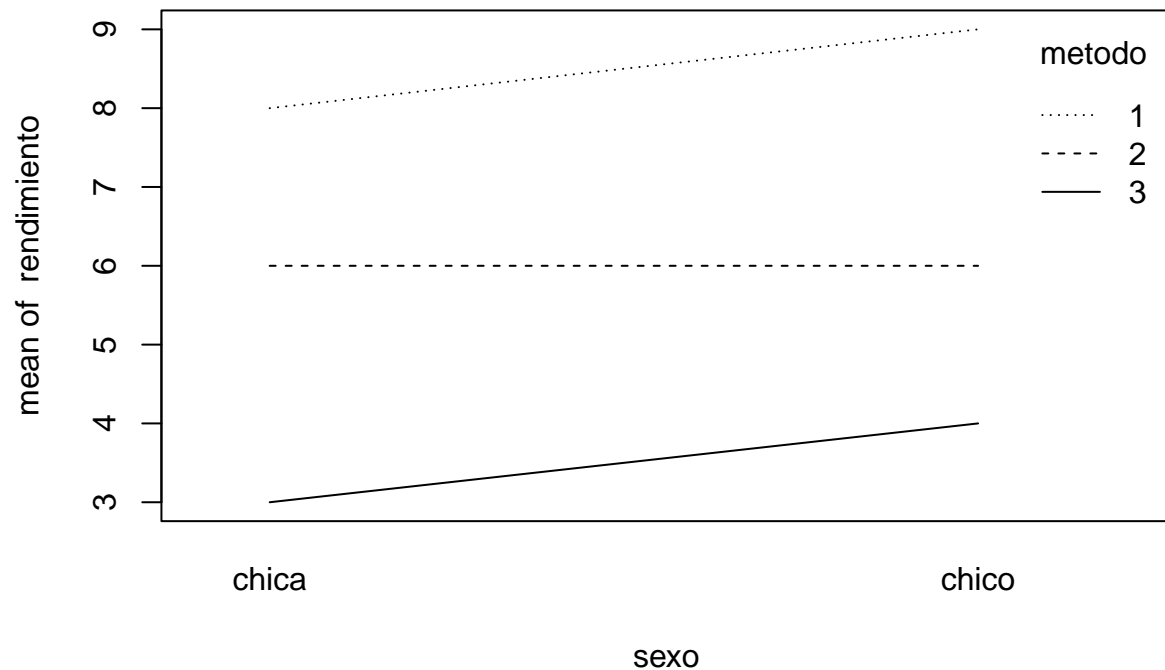
metodo = factor(metodo)
sexo = factor(sexo)
A<-aov(rendimiento~metodo*sexo)
summary(A)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## metodo      2    150   75.00  32.143 3.47e-08 ***
## sexo        1     4    4.00   1.714  0.200
## metodo:sexo  2     2    1.00   0.429  0.655
## Residuals   30     70    2.33
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
interaction.plot(metodo,sexo,rendimiento)
```



```
interaction.plot(sexo, metodo, rendimiento)
```



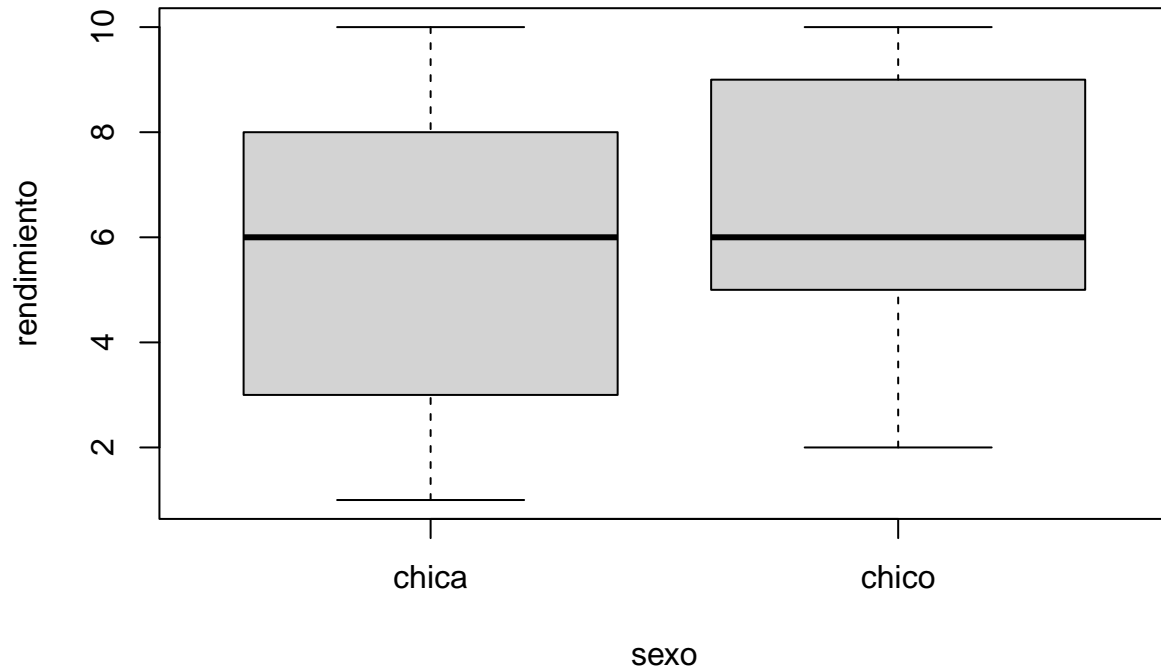
En la variable de sexo vemos que las calificaciones no se intersectan por lo que nos dice que siempre el metodo 1 es el mejor. En cambio, en la primera grafica vemos que las graficas casi lo hacen por lo que no podemos que siempre los hombres tienen mejores calificaciones con seguridad significativa.

## Sin interacción

```
A<-aov(rendimiento~metodo+sexo)
summary(A)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## metodo      2     150    75.00  33.333 1.5e-08 ***
## sexo        1         4     4.00   1.778  0.192
## Residuals   32        72     2.25
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
boxplot(rendimiento ~ sexo)
```



Vamos que el modelo sigue sin ser significativo para sexo por lo que lo descartaremos. Quedandonos solo con un modelo por metodo.

```
C<-aov(rendimiento~metodo)
summary(C)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## metodo      2    150    75.0    32.57 1.55e-08 ***
## Residuals   33     76     2.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

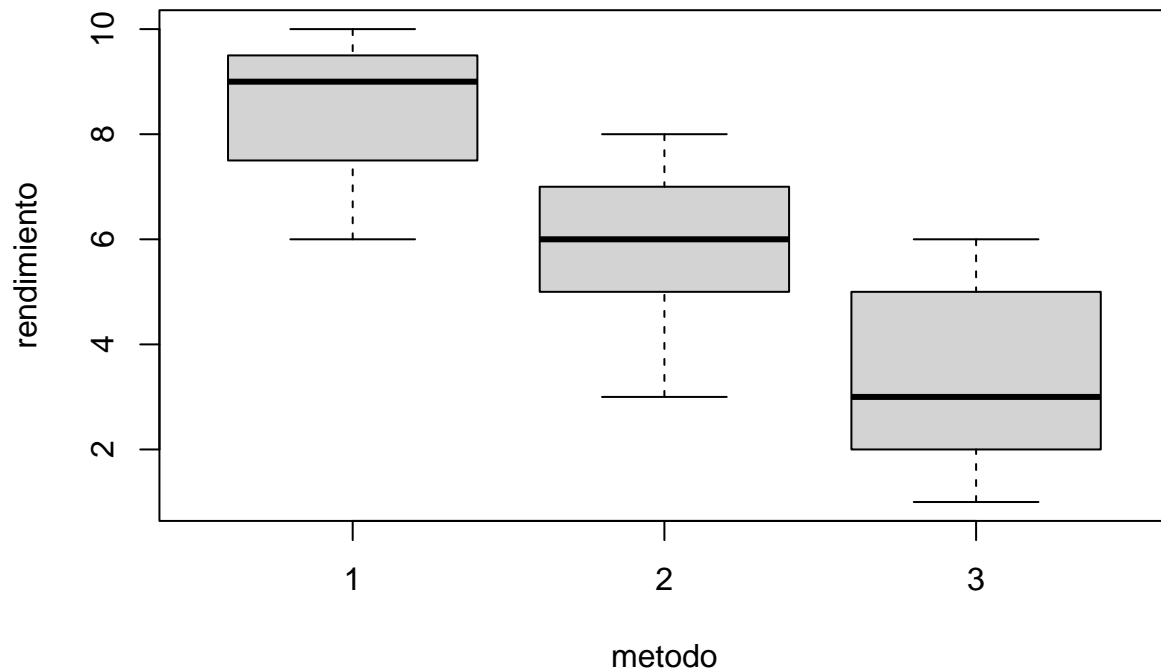
```
tapply(rendimiento,metodo,mean)
```

```
##      1      2      3
## 8.5 6.0 3.5
```

```
mean(rendimiento)
```

```
## [1] 6
```

```
boxplot(rendimiento ~ metodo)
```

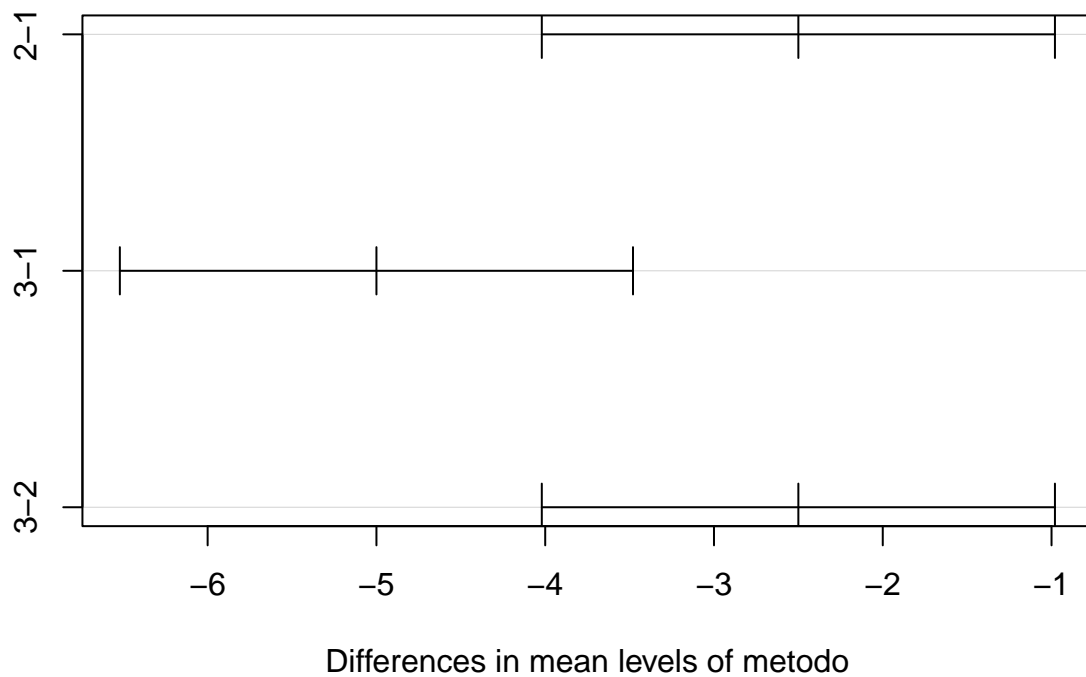


```
I = TukeyHSD(aov(rendimiento ~ metodo))
I

##    Tukey multiple comparisons of means
##      95% family-wise confidence level
##
## Fit: aov(formula = rendimiento ~ metodo)
##
## $metodo
##      diff      lwr      upr    p adj
## 2-1 -2.5 -4.020241 -0.9797592 0.0008674
## 3-1 -5.0 -6.520241 -3.4797592 0.0000000
## 3-2 -2.5 -4.020241 -0.9797592 0.0008674
```

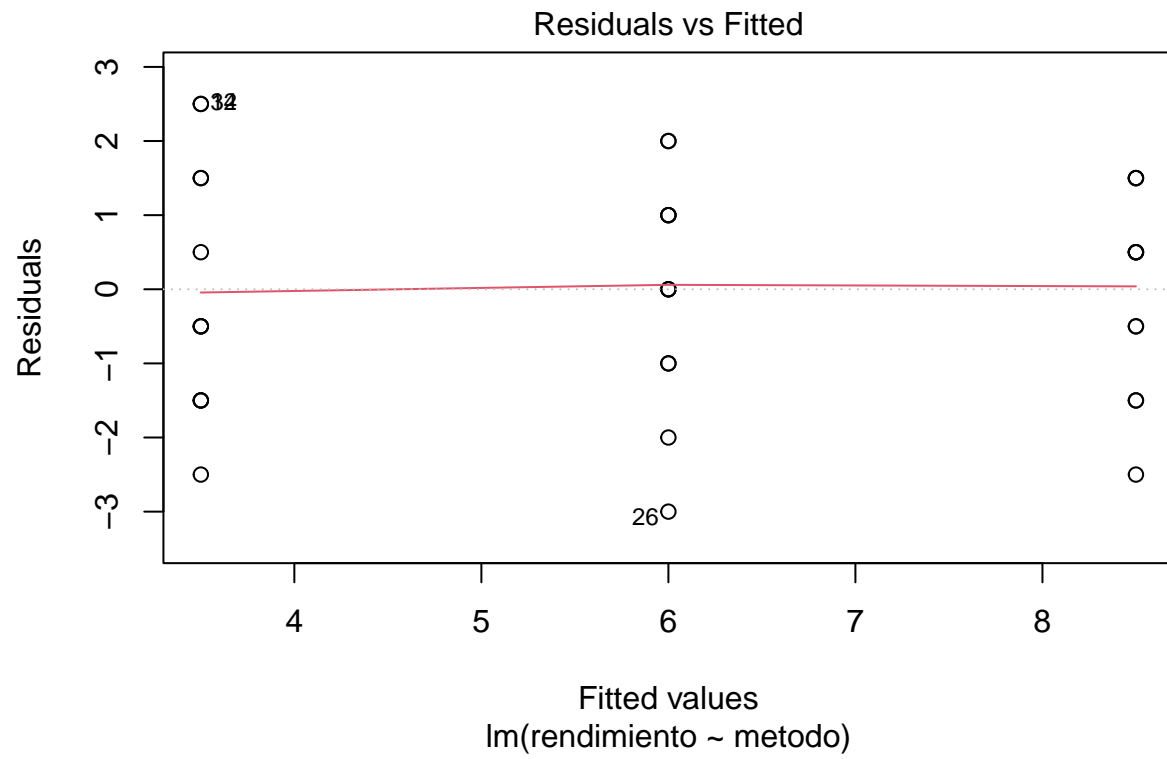
```
plot(I)
```

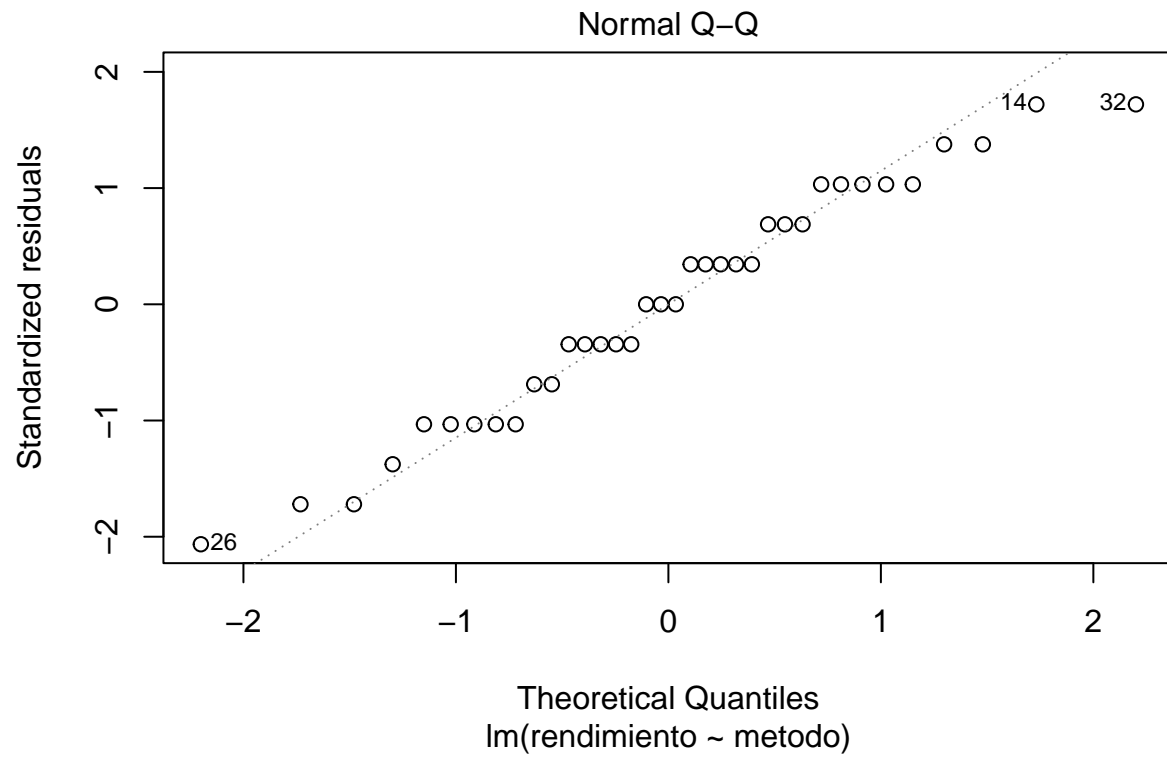
### 95% family-wise confidence level



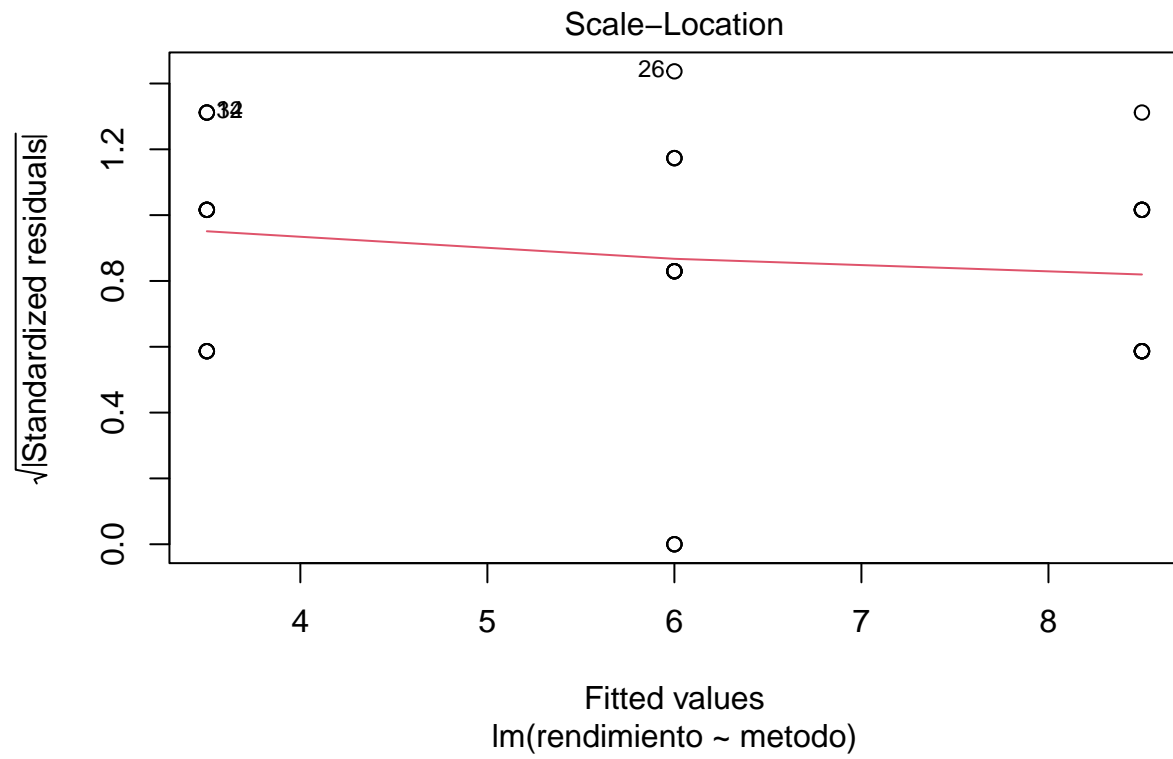
Podemos decir que el efecto del metodo es significativo ya que ninguno de los intervalos contiene al 0 lo cual nos dice que la diferencia entre las medias es distinta.

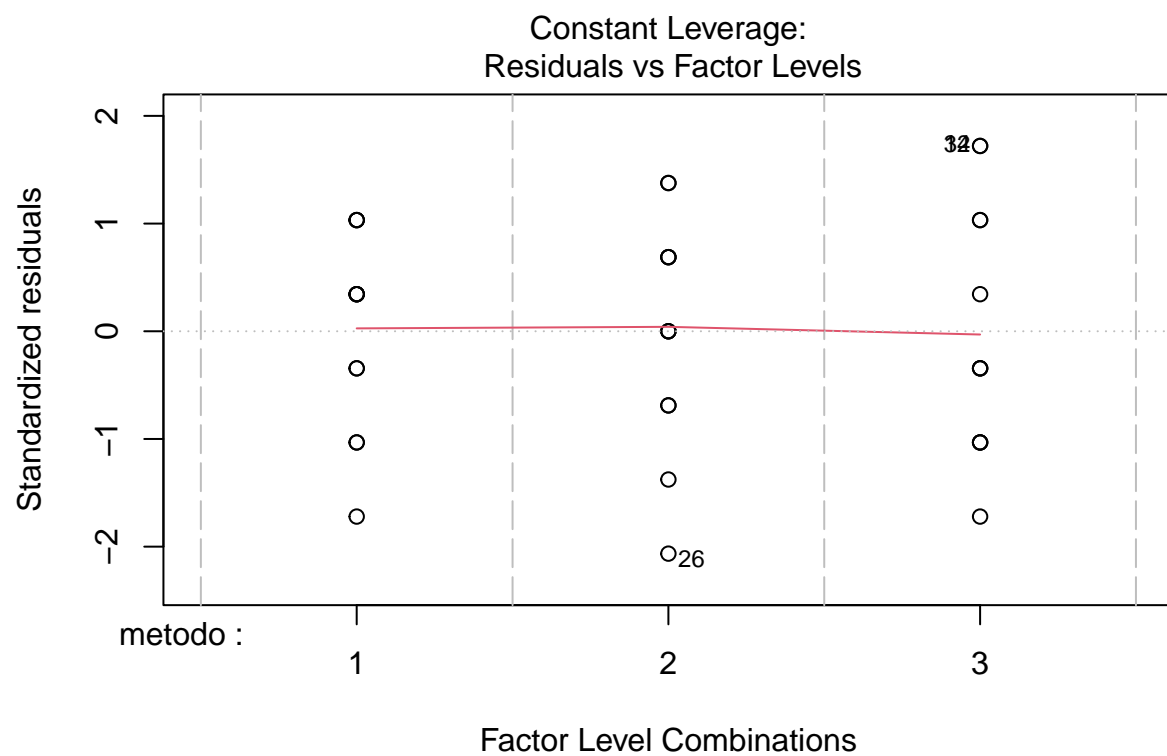
```
plot(lm(rendimiento~metodo))
```











De nuevo, concluimos que el factor de metodo es significativo y que el M1 es el mejor.