

Act 6 ANOVA

Ana Lucía Cárdenas Pérez A01284090

2023-09-02

Problema

En un instituto se han matriculado 36 estudiantes. Se desea explicar el rendimiento de ciencias naturales en función de dos variables: género y metodología de enseñanza. La metodología de enseñanza se analiza en tres niveles: explicación oral y realización del experimento (1er nivel) explicación oral e imágenes (2º nivel) y explicación oral (tercer nivel). En los alumnos matriculados había el mismo número de chicos que de chicas, por lo que formamos dos grupos de 18 sujetos; en cada uno de ellos, el mismo profesor aplicará a grupos aleatorios de 6 estudiantes las 3 metodologías de estudio. A fin de curso los alumnos son sometidos a la misma prueba de rendimiento. Los resultados son los siguientes:

Factor 1: Metodo de Enseñanza

Factor 2: El Sexo

```
Rendimiento <-  
c(10,7,9,9,9,10,5,7,6,6,8,4,2,6,3,5,5,3,9,7,8,8,10,6,8,3,5,6,7,7,2,6,2,1,4,3)  
  
Metodo <- c(rep("M1",6), rep("M2",6), rep("M3",6), rep("M1",6), rep("M2",6),  
rep("M3",6))  
  
Sexo <- c(rep("H",18), rep("M",18))  
  
Metado = factor(Metodo)  
Sexo = factor(Sexo)
```

Hipotesis

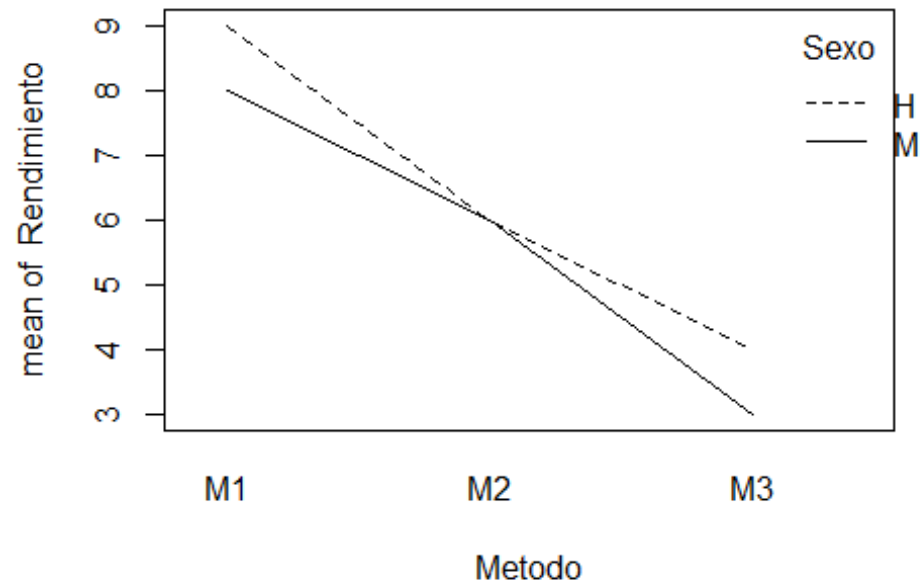
Primera: - H₀ : - H₁ : Segunda: - H₀ : - H₁ : Tercera: - H₀ : - H₁ : #ANOVA (con interacción)

```
A <- aov(Rendimiento ~ Metodo * Sexo)  
summary(A)
```

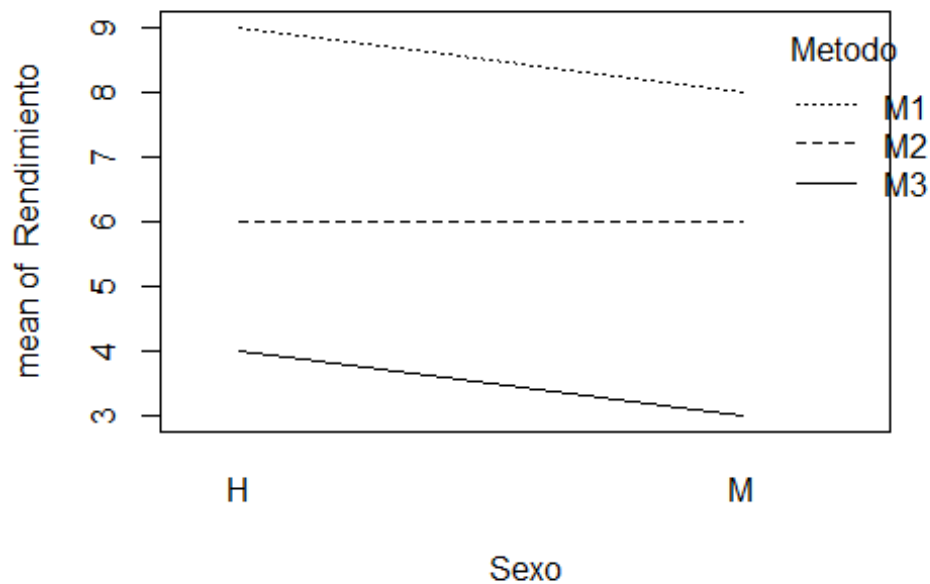
##		Df	Sum Sq	Mean Sq	F value	Pr(>F)
##	Metodo	2	150	75.00	32.143	3.47e-08 ***
##	Sexo	1	4	4.00	1.714	0.200
##	Metodo:Sexo	2	2	1.00	0.429	0.655
##	Residuals	30	70	2.33		

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

interaction.plot(Metodo, Sexo, Rendimiento)
```



```
interaction.plot(Sexo, Metodo, Rendimiento)
```



El M1 es mejor que el M2 y M3 #ANOVA (sin interacción) En el modelo, se consideran sólo los efectos principales. Ya no se usa *, se usa mas

```
B <- aov(Rendimiento ~ Metodo + Sexo)
summary(B)
```

	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
```

Para observar mejor los efectos de los factores principales, se calcula la media por nivel y se grafica por nivel. También se calcula la media general.

```
tapply(Rendimiento, Sexo, mean)
```

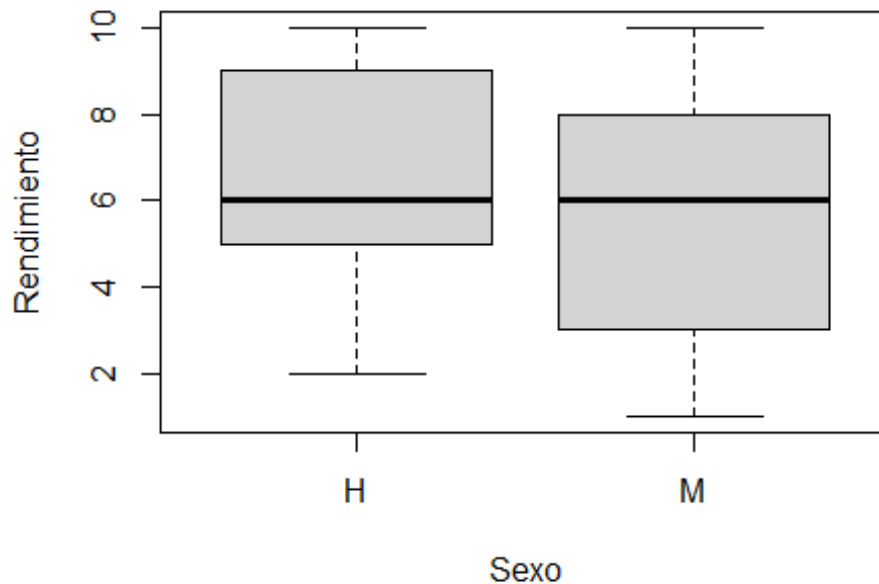
	H	M
mean	6.333333	5.666667

```
tapply(Rendimiento, Metodo, mean)
```

	M1	M2	M3
mean	8.5	6.0	3.5

```
M = mean(Rendimiento)
M
```

```
## [1] 6
boxplot(Rendimiento ~ Sexo)
```



Podemos ver que el rendimiento medio de los hombres es mayor que el de las mujeres, y este también está por encima del rendimiento medio de los dos.

ANOVA con un solo factor (el significativo)

En el modelo, se consideran solo el efecto significativo.

```
C <- aov(Rendimiento ~ Metodo)
summary(C)

##           Df Sum Sq Mean Sq F value    Pr(>F)
## Metodo      2    150    75.00  32.57 1.55e-08 ***
## Residuals  33     76     2.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

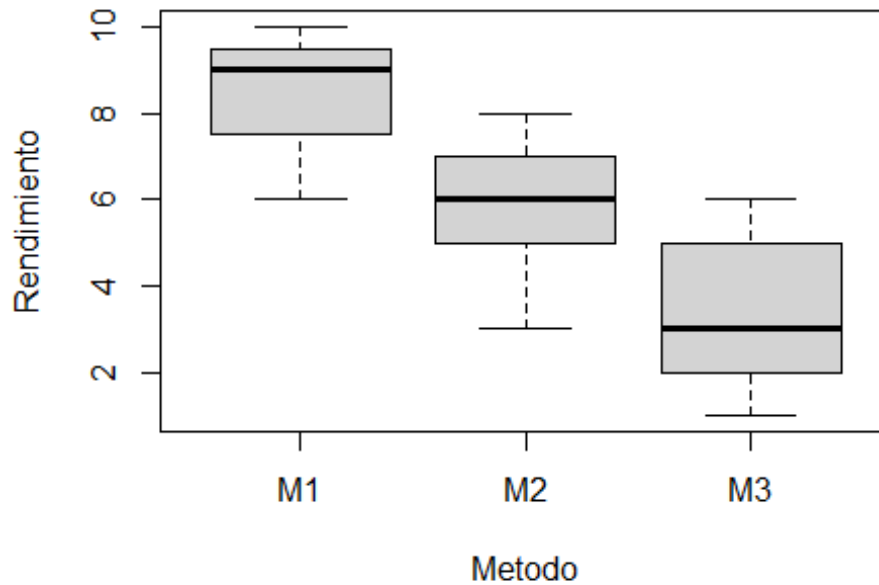
tapply(Rendimiento, Metodo, mean)

##  M1  M2  M3
## 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
## M2-M1 -2.5 -4.020241 -0.9797592 0.0008674
## M3-M1 -5.0 -6.520241 -3.4797592 0.0000000
## M3-M2 -2.5 -4.020241 -0.9797592 0.0008674

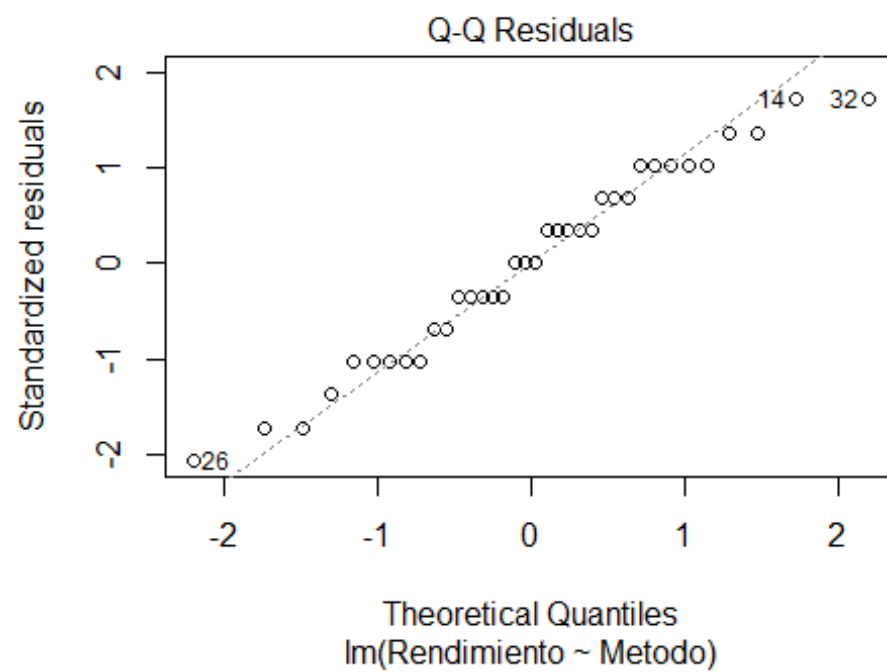
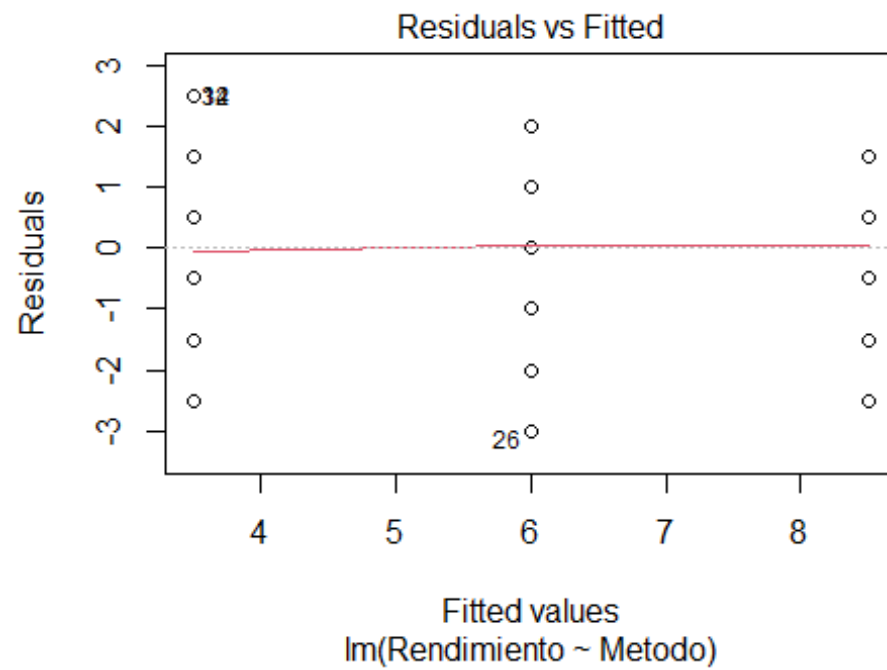
plot(I)
```

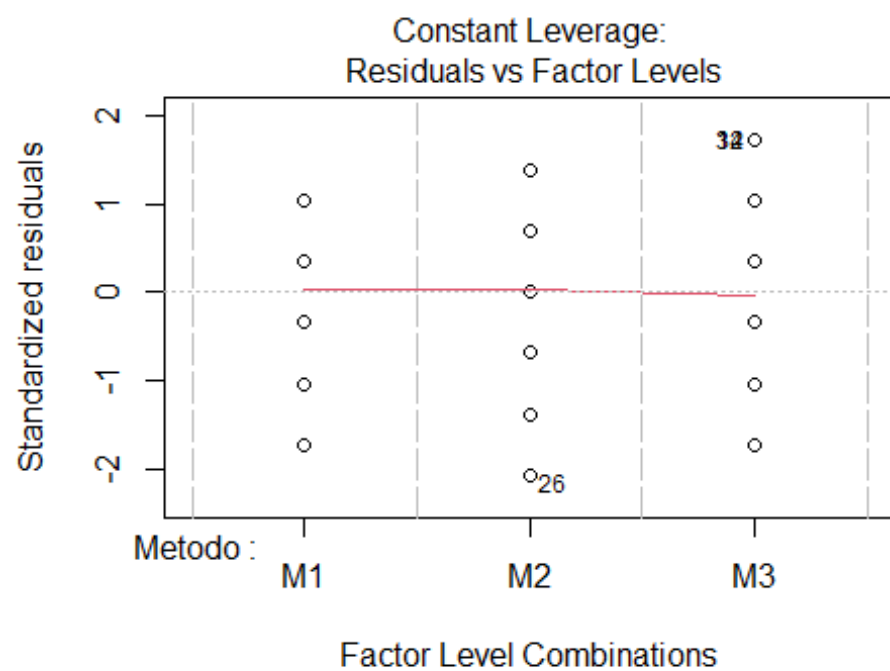
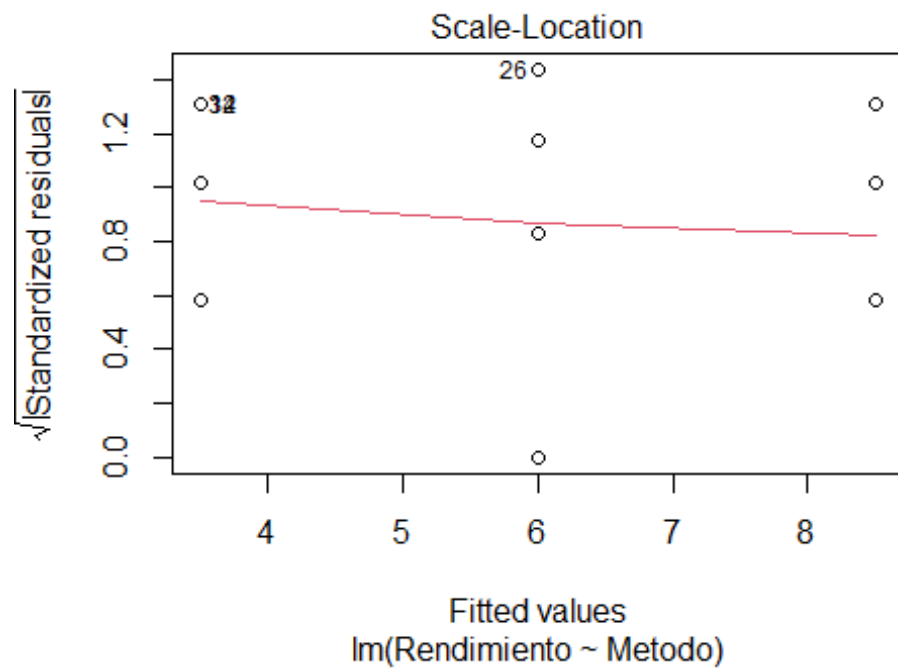


Analisis del modelo

Se verifica la validez del modelo por medio de las gráficas de residuos y la gráfica de normalidad. También se pueden calcular los coeficientes de determinación del modelo para conocer la variación explicada por el modelo.

```
plot(lm(Rendimiento ~ Metodo))
```





CD = $150 / (150 + 76)$

CD

[1] 0.6637168