Act9 ANOVA

Luis Maximiliano López Ramírez

2024-08-27

 # El rendimiento

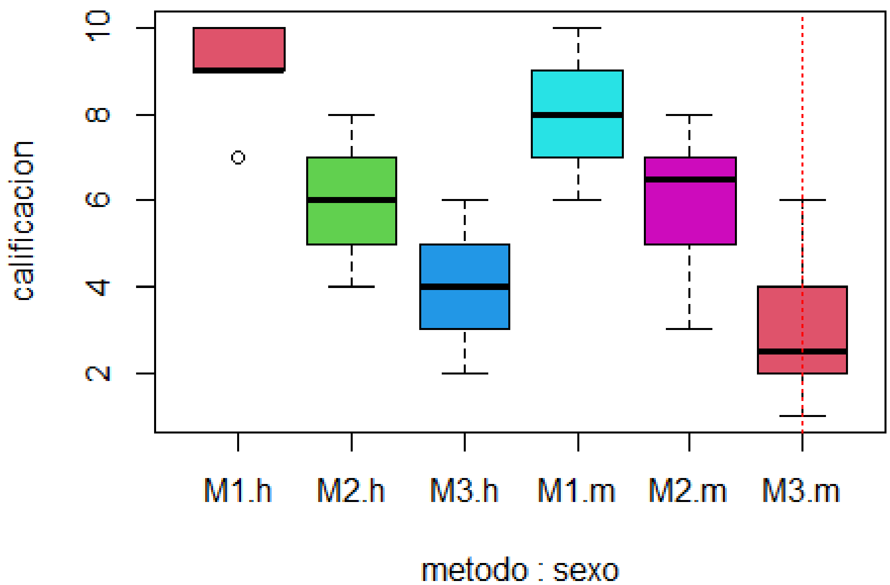
```
calificacion=c(10,7,9,9,9,10,5,7,6,6,8,4,2,6,3,5
,5,3,9,7,8,8,10,6,8,3,5,6,7,7,2,6,2,1,4,3)
metodo=c(rep("M1",6),rep("M2",6),rep("M3",6),rep("M1",6),rep("M2",6),rep(
"M3",6))
sexo = c(rep("h", 18), rep("m",18))
metodo = factor(metodo)
sexo = factor(sexo)

datos <- data.frame(calificacion, metodo, sexo)
```

1. Análisis exploratorio. Calcula la media para el rendimiento por método de enseñanza.

```
m = tapply(calificacion, metodo, mean)
m

##  M1  M2  M3
## 8.5 6.0 3.5
```

Haz el boxplot de la evaluación de los estudiantes por método de enseñanza y sexo.

```
boxplot(calificacion ~ metodo:sexo, col = 2:6)
abline(v = mean(calificacion), lty = 3, col = "red")
```



Interpreta el resultado desde la perspectiva estadística y en el contexto del problema.

Parece que hay diferencias entre las medias de cada método pero no por sexo.

M1.h (Método 1 para hombres): La mediana es alta, cerca de 10, con una pequeña dispersión de los datos, lo que indica que la mayoría de las calificaciones son muy altas y consistentes.

M2.h (Método 2 para hombres): La mediana es más baja que en

M1.h, alrededor de 6, con una mayor dispersión de los datos, lo que sugiere una mayor variabilidad en las calificaciones.

M3.h (Método 3 para hombres): Similar a M2.h, pero con una dispersión un poco menor.

M1.m (Método 1 para mujeres): La mediana es comparable a la de M2.h, con una dispersión similar.

M2.m (Método 2 para mujeres): La mediana es similar a la de

M3.h, con una dispersión un poco mayor.

M3.m (Método 3 para mujeres): La mediana es baja, alrededor de 4, con una dispersión moderada.

Escribe tus conclusiones parciales

Desempeño por Método y Género:

Los hombres que usan el Método 1 (M1.h) tienden a tener calificaciones significativamente más altas y consistentes en comparación con otros métodos y con las mujeres. Las mujeres tienen calificaciones más bajas en el Método 3 (M3.m), lo que sugiere que este método podría ser menos efectivo o adecuado para ellas.

Variabilidad y Consistencia:

Hay más consistencia en las calificaciones de los hombres que usan el Método 1, mientras que otros grupos muestran más variabilidad. Es posible que los métodos no sean igualmente efectivos para hombres y mujeres, o que haya otros factores que expliquen las diferencias en las calificaciones.

2. Las hipótesis. Establece las hipótesis estadísticas (tienen que ser 3).

Método de enseñanza (tres niveles: M1, M2, M3) \

Sexo (dos niveles: hombre, mujer)

$$Y_{ijk} = \mu + \tau_i + \alpha_j + \tau_i\alpha_j + \epsilon_{ijk}$$

Bajo este modelo:

$$\sum_{i=1}^{n_\tau} \tau_i = 0 \quad \text{Factor 1}$$

$$\sum_{j=1}^{n_\alpha} \alpha_j = 0 \quad \text{Factor 2}$$

$$\sum_{i=1}^{n_\tau} \sum_{j=1}^{n_\alpha} \tau_i \, \alpha_j = 0 \quad \text{Interacción}$$

$H_0 \quad : \tau_i = 0$
$H_1 \quad$:algún τ_i es distinto de cero

$H_0 \quad : \alpha_j = 0$
$H_1 \quad$:algún α_j es distinto de cero

$H_0 \quad : \tau_i\alpha_j = 0$
$H_1 \quad : \tau_i\alpha_j$ es distinto de cero

3. Realiza el ANOVA para dos niveles con interacción.

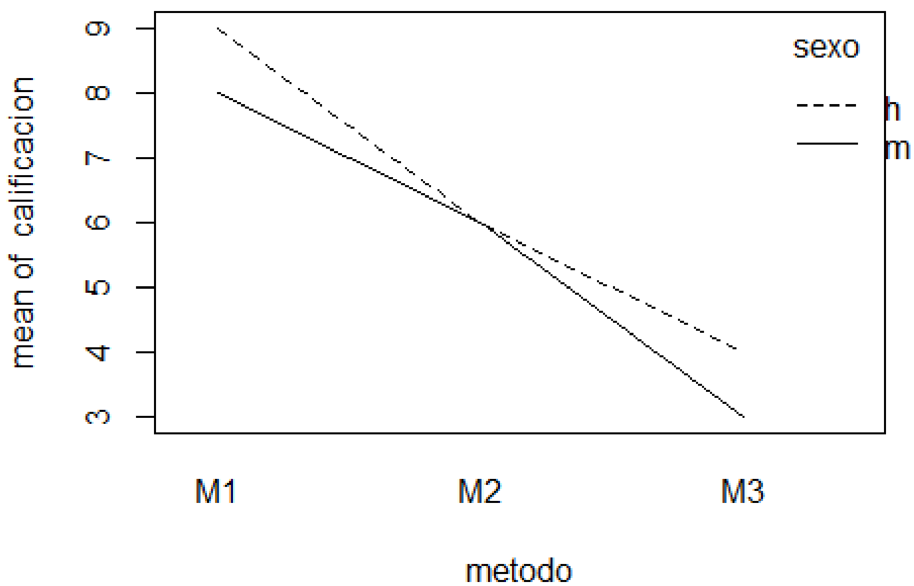
Consulta el código en R en los apoyos de clase de "ANOVA":

```
A = aov(calificacion ~ metodo*sexo, datos)
summary(A)
```

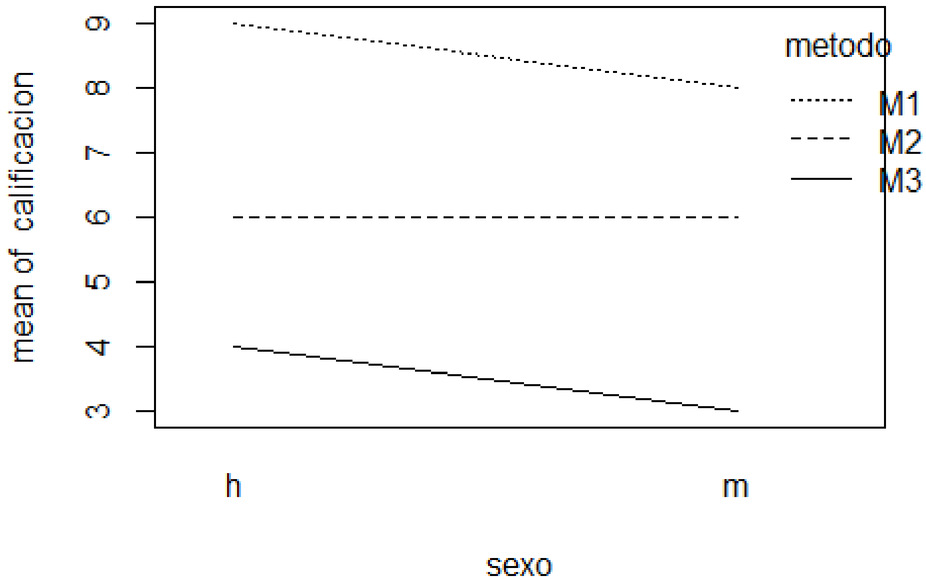
```
##              Df Sum Sq Mean Sq F value   Pr(>F)    
## metodo        2    150   75.00  32.143 3.47e-08 ***
## sexo          1      4    4.00   1.714    0.200    
## metodo:sexo   2      2    1.00   0.429    0.655    
## Residuals    30     70    2.33                     
## ---
## Signif. codes: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

Haz la gráfica de interacción de dos factores en ANOVA

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

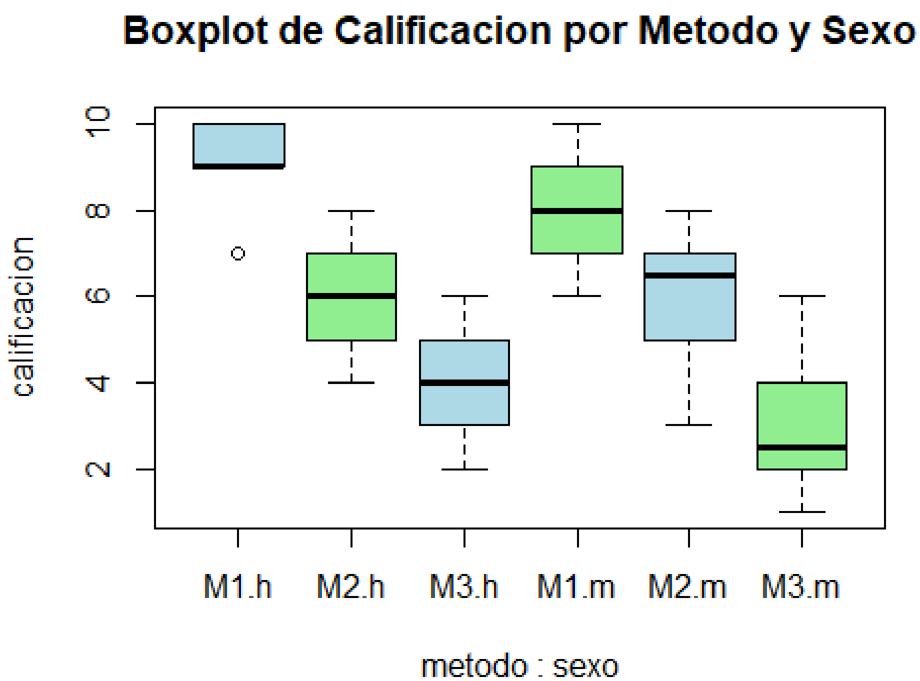



```
interaction.plot(sexo, metodo, calificacion)
```



Haz el boxplot para visualizar la interacción de los factores, por ejemplo, peso por dieta interacción ejercicio:

```
boxplot(calificacion ~ metodo * sexo, data = datos, col = c("lightblue",
"lightgreen"), main = "Boxplot de Calificacion por Metodo y Sexo")
```



Interpreta el resultado desde la perspectiva estadística y en el contexto del problema

El análisis ANOVA revela que el método de enseñanza (metodo) tiene un efecto estadísticamente significativo en las calificaciones, con un valor p muy inferior a 0.05, indicando diferencias notables entre los métodos M1, M2 y M3. Por otro lado, ni el sexo (sexo) ni la interacción entre método y sexo (metodo:sexo) mostraron efectos significativos sobre las calificaciones, con valores p superiores a 0.05, lo que sugiere que estos factores no influyen de manera considerable en los resultados.

Escribe tus conclusiones parciales

Las calificaciones de los estudiantes están significativamente influenciadas por el método de enseñanza utilizado, mientras que el sexo de los participantes y la interacción entre método y sexo no tienen un impacto relevante en las calificaciones. Esto implica que la elección del método de enseñanza es clave para mejorar el rendimiento académico, independientemente del sexo de los estudiantes.

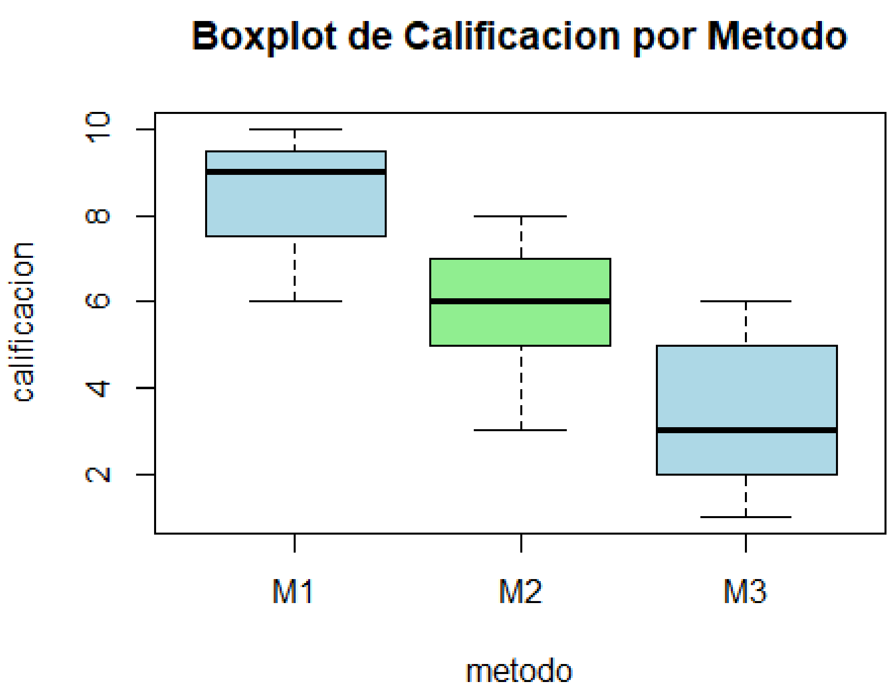
4. Realiza el ANOVA para dos niveles sin interacción.

```r
A = aov(calificacion ~ metodo + sexo, datos)
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
```

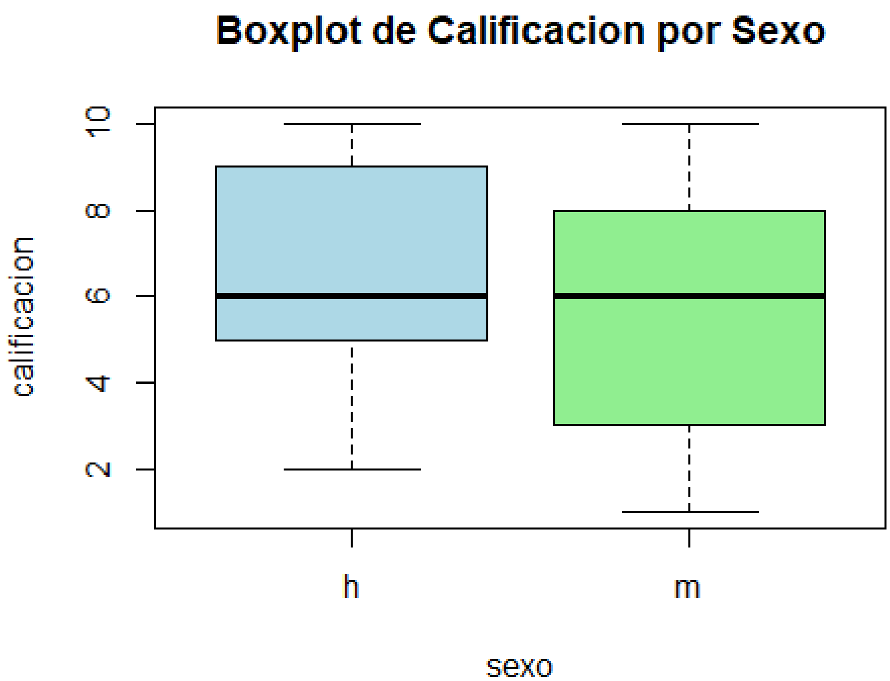
```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Haz el boxplot de rendimiento por sexo. Calcula la media para el rendimiento por sexo y método.

```
m = tapply(calificacion, sexo, mean)
m

##        h        m
## 6.333333 5.666667

m = tapply(calificacion, metodo, mean)
m

##  M1  M2  M3
## 8.5 6.0 3.5

boxplot(calificacion ~ metodo, data = datos, col = c("lightblue",
"lightgreen"), main = "Boxplot de Calificacion por Metodo")
```



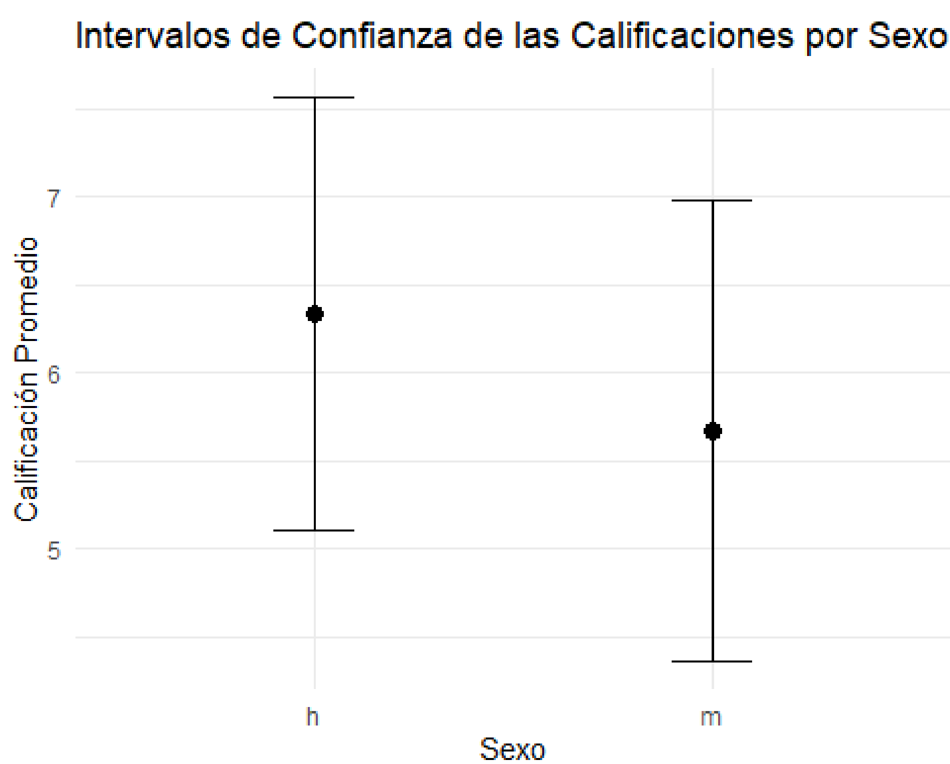

```
boxplot(calificacion ~ sexo, data = datos, col = c("lightblue",
"lightgreen"), main = "Boxplot de Calificacion por Sexo")
```



Haz los intervalos de confianza de rendimiento por sexo. Grafícalos

```
# Calcular la media y el intervalo de confianza por sexo
library(dplyr)

## Warning: package 'dplyr' was built under R version 4.3.2

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.3.2

# Agrupamos por sexo y calculamos la media y el intervalo de confianza
datos_summary <- datos %>%
  group_by(sexo) %>%
  summarise(
    mean_calificacion = mean(calificacion),
    n = n(),
```

```
    se = sd(calificacion) / sqrt(n),
    lower_ci = mean_calificacion - qt(1 - 0.05 / 2, n - 1) * se,
    upper_ci = mean_calificacion + qt(1 - 0.05 / 2, n - 1) * se
  )

print(datos_summary)

## # A tibble: 2 × 6
##   sexo  mean_calificacion     n    se lower_ci upper_ci
##   <fct>             <dbl> <int> <dbl>    <dbl>    <dbl>
## 1 h                  6.33    18 0.583     5.10     7.56
## 2 m                  5.67    18 0.621     4.36     6.98

# Graficar los intervalos de confianza
ggplot(datos_summary, aes(x = sexo, y = mean_calificacion)) +
  geom_point(size = 3) +
  geom_errorbar(aes(ymin = lower_ci, ymax = upper_ci), width = 0.2) +
  labs(title = "Intervalos de Confianza de las Calificaciones por Sexo",
       x = "Sexo", y = "Calificación Promedio") +
  theme_minimal()
```



Interpreta el resultado desde la perspectiva estadística y en el contexto del problema.

El gráfico muestra los intervalos de confianza del 95% para las calificaciones promedio según el sexo. La media de las calificaciones para los hombres se encuentra alrededor de 6.2, mientras que para las mujeres está alrededor de 6.8. Sin embargo, los intervalos de confianza son amplios y se superponen, lo que indica que no hay una diferencia estadísticamente significativa en las calificaciones promedio entre hombres y mujeres.

Escribe tus conclusiones parciales

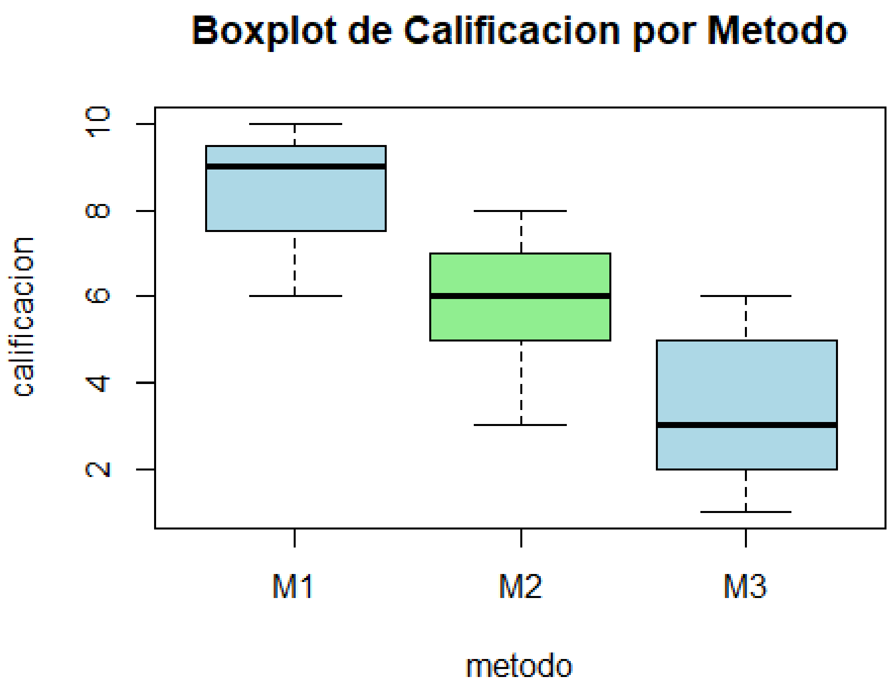
Dado que los intervalos de confianza se superponen, no podemos concluir con certeza que exista una diferencia en el rendimiento académico entre hombres y mujeres en este estudio. Es probable que otros factores, como el método de enseñanza, tengan un impacto mayor en las calificaciones. Sería recomendable investigar más para determinar si existen diferencias en otros contextos o con tamaños de muestra más grandes.

5. Realiza el ANOVA para un efecto principal

```
A = aov(calificacion ~ metodo, datos)
summary(A)

##             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
```

Haz el boxplot de rendimiento por método de enseñanza. Calcula la media.

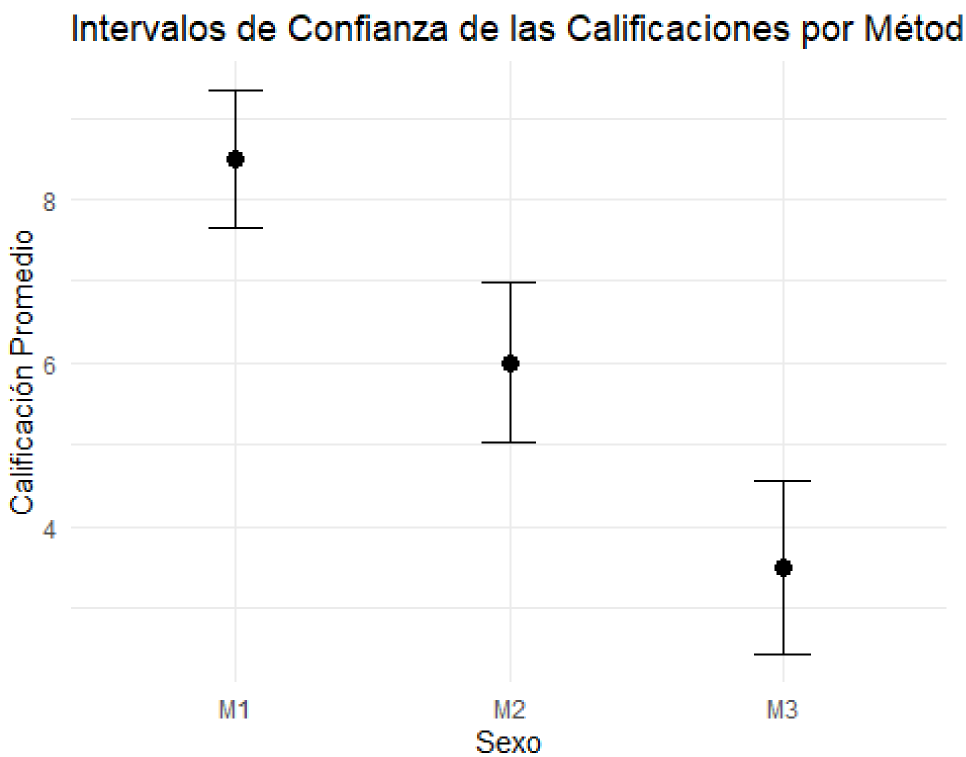
```
m = tapply(calificacion, metodo, mean)
m

##  M1  M2  M3
## 8.5 6.0 3.5

boxplot(calificacion ~ metodo, data = datos, col = c("lightblue",
"lightgreen"), main = "Boxplot de Calificacion por Metodo")
```



Haz los intervalos de confianza de rendimiento por método. Grafícalos

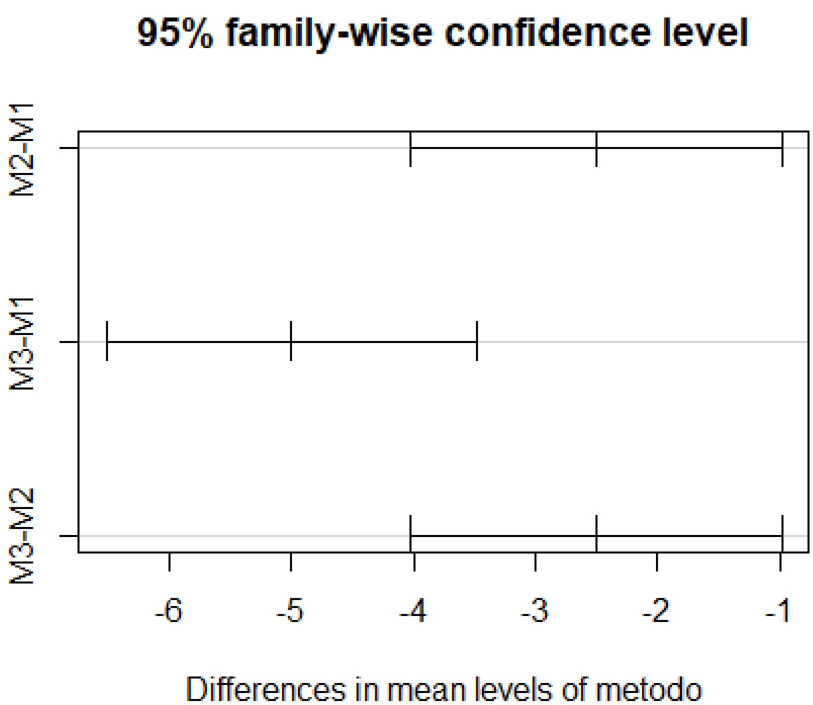
```
# Calcular la media y el intervalo de confianza por sexo
library(dplyr)
library(ggplot2)

# Agrupamos por sexo y calculamos la media y el intervalo de confianza
datos_summary <- datos %>%
  group_by(metodo) %>%
  summarise(
    mean_calificacion = mean(calificacion),
    n = n(),
    se = sd(calificacion) / sqrt(n),
    lower_ci = mean_calificacion - qt(1 - 0.05 / 2, n - 1) * se,
    upper_ci = mean_calificacion + qt(1 - 0.05 / 2, n - 1) * se
  )

print(datos_summary)

## # A tibble: 3 × 6
##   metodo mean_calificacion     n    se lower_ci upper_ci
##   <fct>              <dbl> <int> <dbl>    <dbl>    <dbl>
## 1 M1                   8.5    12 0.379     7.66     9.34
## 2 M2                   6      12 0.444     5.02     6.98
## 3 M3                   3.5    12 0.485     2.43     4.57
```

```r
# Graficar los intervalos de confianza
ggplot(datos_summary, aes(x = metodo, y = mean_calificacion)) +
  geom_point(size = 3) +
  geom_errorbar(aes(ymin = lower_ci, ymax = upper_ci), width = 0.2) +
  labs(title = "Intervalos de Confianza de las Calificaciones por 
Método",
       x = "Sexo", y = "Calificación Promedio") +
  theme_minimal()
```



Realiza la prueba de comparaciones múltiples de Tukey. Grafica los intervalos de confianza de Tukey.

```r
Tu=TukeyHSD(A)
Tu

##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = calificacion ~ metodo, data = datos)
##
## $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(TukeyHSD(A))
```



Interpreta el resultado desde la perspectiva estadística y en el contexto del problema.

Gráfico de intervalos de confianza de diferencias de medias: En esta gráfica se presentan las diferencias en los niveles medios entre diferentes métodos (M2-M1 y M3-M2) con un intervalo de confianza del 95%. Si un intervalo incluye el valor cero, esto sugiere que no hay una diferencia estadísticamente significativa entre los métodos correspondientes. En este caso, parece que las diferencias entre M2 y M1, y entre M3 y M2, no incluyen cero, lo que indica que probablemente hay una diferencia significativa entre estos métodos.

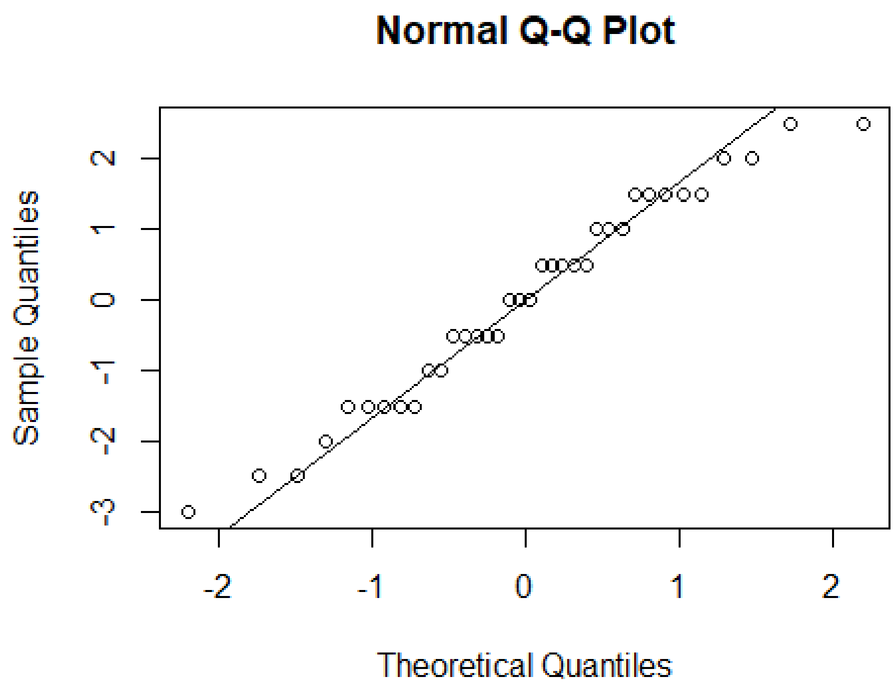
Escribe tus conclusiones parciales

Los métodos M1, M2 y M3 parecen diferir significativamente en sus medias, lo que sugiere que el cambio de un método a otro podría llevar a un cambio notable en el resultado o la medición que se esté evaluando. El tamaño de los intervalos de confianza también puede indicar la precisión de las estimaciones de diferencia media. Un intervalo más estrecho sugiere una estimación más precisa de la diferencia entre los métodos.

6. Comprueba la validez del modelo. Comprueba

Normalidad

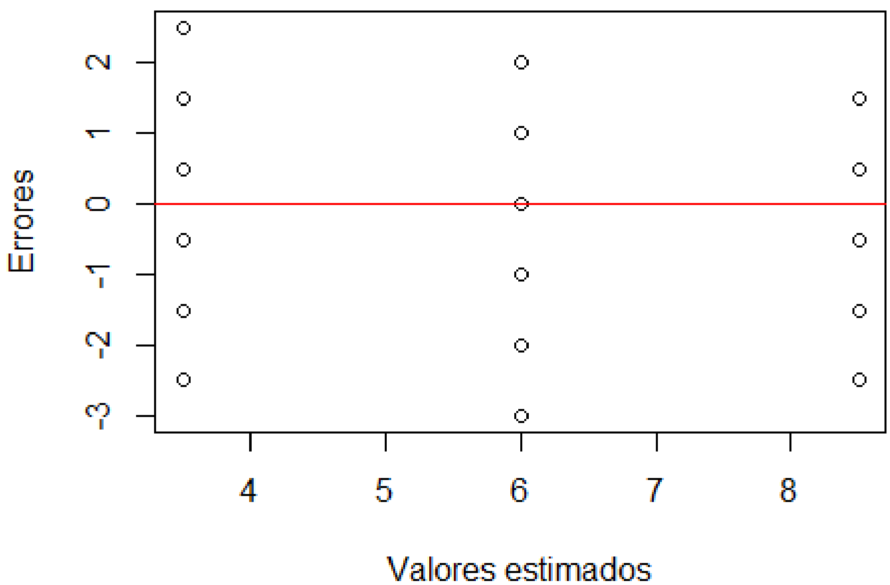
```
residuos=A$residuals
qqnorm(residuos)
qqline(residuos)
```




```r
library(stats)

# Realizar la prueba de Shapiro-Wilk
shapiro_test = shapiro.test(residuos)

# Imprimir el resultado de la prueba
print(shapiro_test)

##
##  Shapiro-Wilk normality test
##
## data:  residuos
## W = 0.96734, p-value = 0.3573

# Establecer regla de decisión basada en el p-valor
if (shapiro_test$p.value < 0.05) {
  cat("Rechazar la hipótesis nula: Los datos no siguen una distribución
normal.\n")
} else {
  cat("No rechazar la hipótesis nula: Los datos siguen una distribución
normal.\n")
}

## No rechazar la hipótesis nula: Los datos siguen una distribución
normal.
```

Homocedasticidad

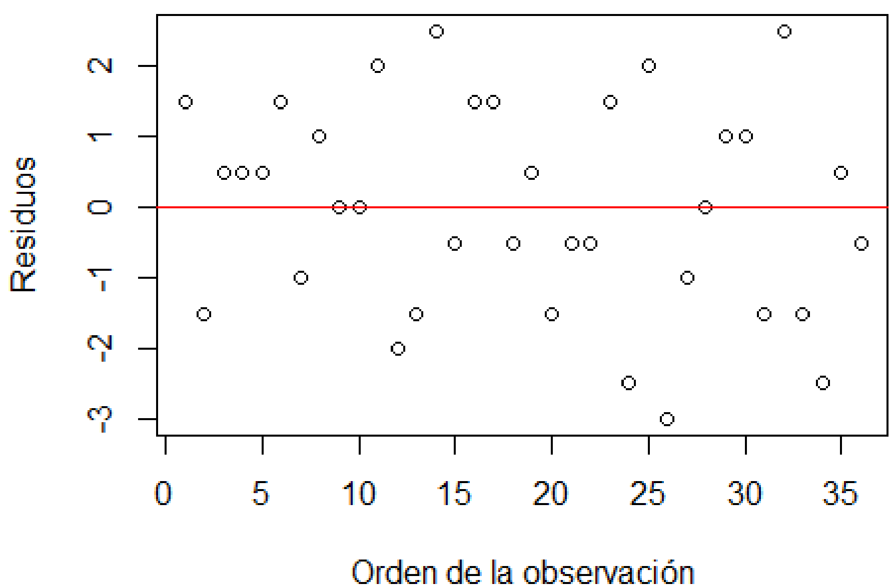
```r
plot(A$fitted.values,A$residuals,ylab="Errores",xlab="Valores estimados")
abline(h=0,col="red")
```




```r
# Cargar la librería necesaria para la prueba de Bartlett
library(stats)

# Realizar la prueba de Bartlett
bartlett_test = bartlett.test(calificacion ~ metodo, data = datos)

# Imprimir el resultado de la prueba
print(bartlett_test)

##
##  Bartlett test of homogeneity of variances
##
## data:  calificacion by metodo
## Bartlett's K-squared = 0.63268, df = 2, p-value = 0.7288

# Establecer regla de decisión basada en el p-valor
if (bartlett_test$p.value < 0.05) {
  cat("Rechazar la hipótesis nula: Las varianzas no son homogéneas.\n")
} else {
  cat("No rechazar la hipótesis nula: Las varianzas son homogéneas.\n")
}

## No rechazar la hipótesis nula: Las varianzas son homogéneas.
```

Independencia

```r
library(lmtest)

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

n = tapply(calificacion, metodo, length)
plot(c(1:sum(n)),A$residuals,xlab="Orden de la
observación",ylab="Residuos")
abline(h=0,col="red")
```




```r
# Realizar la prueba de Durbin-Watson
dw_test = dwtest(A, alternative = "two.sided")

# Imprimir el resultado de la prueba
print(dw_test)

##
##  Durbin-Watson test
##
## data:  A
## DW = 2.6974, p-value = 0.05887
## alternative hypothesis: true autocorrelation is not 0
```

```r
# Establecer regla de decisión basada en el p-valor
if (dw_test$p.value < 0.05) {
  cat("Rechazar la hipótesis nula: Existe evidencia de autocorrelación en los residuos.\n")
} else {
  cat("No rechazar la hipótesis nula: No hay evidencia suficiente de autocorrelación en los residuos.\n")
}

## No rechazar la hipótesis nula: No hay evidencia suficiente de
autocorrelación en los residuos.
```

Relación lineal entre las variables (coeficiente de determinación).

```r
# Ajustar el modelo ANOVA
anova_metodo <- aov(calificacion ~ metodo, data = datos)

# Calcular el R^2
suma_cuadrados_total <- sum((datos$calificacion -
mean(datos$calificacion))^2)
suma_cuadrados_modelo <- sum((fitted(anova_metodo) -
mean(datos$calificacion))^2)

R2 <- suma_cuadrados_modelo / suma_cuadrados_total

# Mostrar el R^2
print(paste("Coeficiente de Determinación(R^2):",R2))

## [1] "Coeficiente de Determinación(R^2): 0.663716814159292"
```

7. Concluye en el contexto del problema.

Tanto la prueba de ANOVA, intervalos de confianza y la prueba de Tukey muestran que la variabilidad entre los grupos de 'metodos' son significativas y no debido al azar. Después se aplicaron los supuestos del ANOVA que son normalidad, homocedasticidad, independencia y relación lineal, en cada prueba se ´puede ver que cuumple con cada uno de los supestos por lo que el modelo el ANOVA es válido y podemos decir que el 'metodo' influye en el rendimiento de los estudiantes.

 # Vibración de motores

```r
# Materiales de la carcasa (Factor A)
Material <- c(rep("Acero", 10), rep("Aluminio", 10), rep("Plastico", 10))

# Proveedores de cojinetes (Factor B)
Proveedor <- c("Proveedor1", "Proveedor2", "Proveedor3", "Proveedor4",
"Proveedor5",
               "Proveedor1", "Proveedor2", "Proveedor3", "Proveedor4",
"Proveedor5",
               "Proveedor1", "Proveedor2", "Proveedor3", "Proveedor4",
"Proveedor5",
```

```r
                "Proveedor1", "Proveedor2", "Proveedor3", "Proveedor4",
"Proveedor5",
                "Proveedor1", "Proveedor2", "Proveedor3", "Proveedor4",
"Proveedor5",
                "Proveedor1", "Proveedor2", "Proveedor3", "Proveedor4",
"Proveedor5")


# Datos de vibración en micrones
Vibracion <- c(13.1, 16.3, 13.7, 15.7, 13.5,   # Acero
               13.2, 15.8, 14.3, 15.8, 12.5,   # Acero
               15.0, 15.7, 13.9, 13.7, 13.4,   # Aluminio
               14.8, 16.4, 14.3, 14.2, 13.8,   # Aluminio
               14.0, 17.2, 12.4, 14.4, 13.2,   # Plástico
               14.3, 16.7, 12.3, 13.9, 13.1)


# Convertir las variables 'material' y 'proveedor' en factores
Material <- factor(Material)
Proveedor <- factor(Proveedor)

# Crear un data frame con los datos
datos <- data.frame(Vibracion = Vibracion, Material = Material, Proveedor
= Proveedor)
datos

##    Vibracion Material  Proveedor
## 1       13.1    Acero Proveedor1
## 2       16.3    Acero Proveedor2
## 3       13.7    Acero Proveedor3
## 4       15.7    Acero Proveedor4
## 5       13.5    Acero Proveedor5
## 6       13.2    Acero Proveedor1
## 7       15.8    Acero Proveedor2
## 8       14.3    Acero Proveedor3
## 9       15.8    Acero Proveedor4
## 10      12.5    Acero Proveedor5
## 11      15.0 Aluminio Proveedor1
## 12      15.7 Aluminio Proveedor2
## 13      13.9 Aluminio Proveedor3
## 14      13.7 Aluminio Proveedor4
## 15      13.4 Aluminio Proveedor5
## 16      14.8 Aluminio Proveedor1
## 17      16.4 Aluminio Proveedor2
## 18      14.3 Aluminio Proveedor3
## 19      14.2 Aluminio Proveedor4
## 20      13.8 Aluminio Proveedor5
## 21      14.0 Plastico Proveedor1
## 22      17.2 Plastico Proveedor2
## 23      12.4 Plastico Proveedor3
```

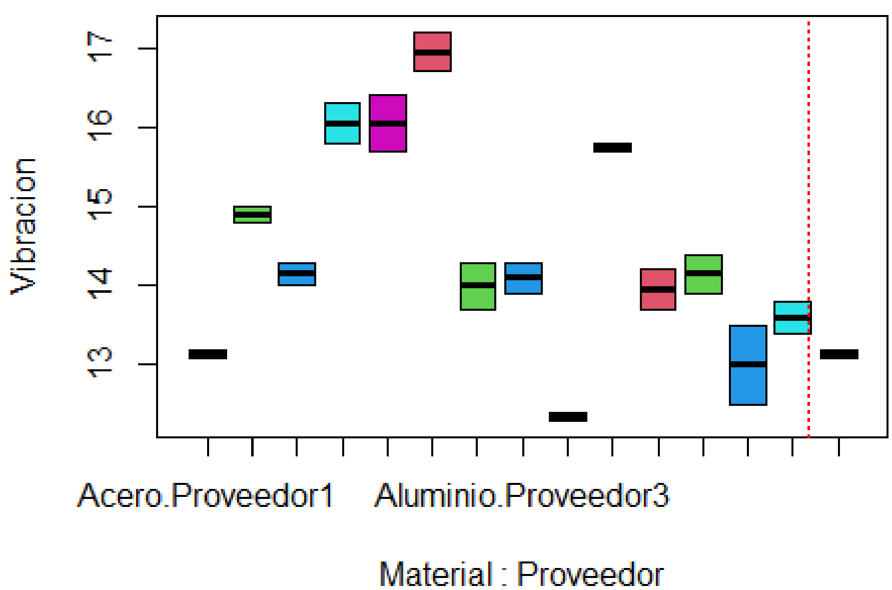
```
## 24      14.4 Plastico Proveedor4
## 25      13.2 Plastico Proveedor5
## 26      14.3 Plastico Proveedor1
## 27      16.7 Plastico Proveedor2
## 28      12.3 Plastico Proveedor3
## 29      13.9 Plastico Proveedor4
## 30      13.1 Plastico Proveedor5
```

1. Análisis exploratorio. Calcula la media para la vibración por Material y Proveedor.

```
m = tapply(Vibracion, Material, mean)
m

##    Acero Aluminio Plastico
##    14.39    14.52    14.15

m = tapply(Vibracion, Proveedor, mean)
m

## Proveedor1 Proveedor2 Proveedor3 Proveedor4 Proveedor5
##   14.06667   16.35000   13.48333   14.61667   13.25000
```

Haz el boxplot de la evaluación de la vibración por material y proveedor.

```
boxplot(Vibracion ~ Material:Proveedor, col = 2:6)
abline(v = mean(Vibracion), lty = 3, col = "red")
```



Interpreta el resultado desde la perspectiva estadística y en el contexto del problema.

Viendo las medias del material y proveedor así como el boxplot, se puede inferir que el proveedor puede que tenga alguna influencia en la vibración mientras que el material no parece ser tan visibile dicho efecto.

Escribe tus conclusiones parciales

El proveedor 2 es el que tiene en promedio mayor vibración mietras que el proveddor 5 es el que tiene menos. En cuanto los materiales hay diferencias pero no está claro sin son significativas como para mencionarlas.

2. Las hipótesis. Establece las hipótesis estadísticas (tienen que ser 3).

Proveedor (5 niveles: P1, P2, P3, P4, P5) \

Material (dos niveles: Acero, Plástico y Aluminio)

$$Y_{ijk} = \mu + \tau_i + \alpha_j + \tau_i\alpha_j + \epsilon_{ijk}$$

Bajo este modelo:

$$\sum_{i=1}^{n_\tau} \tau_i = 0 \quad \text{Factor 1}$$

$$\sum_{j=1}^{n_\alpha} \alpha_j = 0 \quad \text{Factor 2}$$

$$\sum_{i=1}^{n_\tau} \sum_{j=1}^{n_\alpha} \tau_i \, \alpha_j = 0 \quad \text{Interacción}$$

$$\begin{aligned} H_0 &\quad : \tau_i = 0 \\ H_1 &\quad :\text{algún } \tau_i \text{ es distinto de cero} \end{aligned}$$

$$\begin{aligned} H_0 &\quad : \alpha_j = 0 \\ H_1 &\quad :\text{algún } \alpha_j \text{ es distinto de cero} \end{aligned}$$

$$\begin{aligned} H_0 &\quad : \tau_i\alpha_j = 0 \\ H_1 &\quad : \tau_i\alpha_j \text{ es distinto de cero} \end{aligned}$$

3. Realiza el ANOVA para dos niveles con interacción.

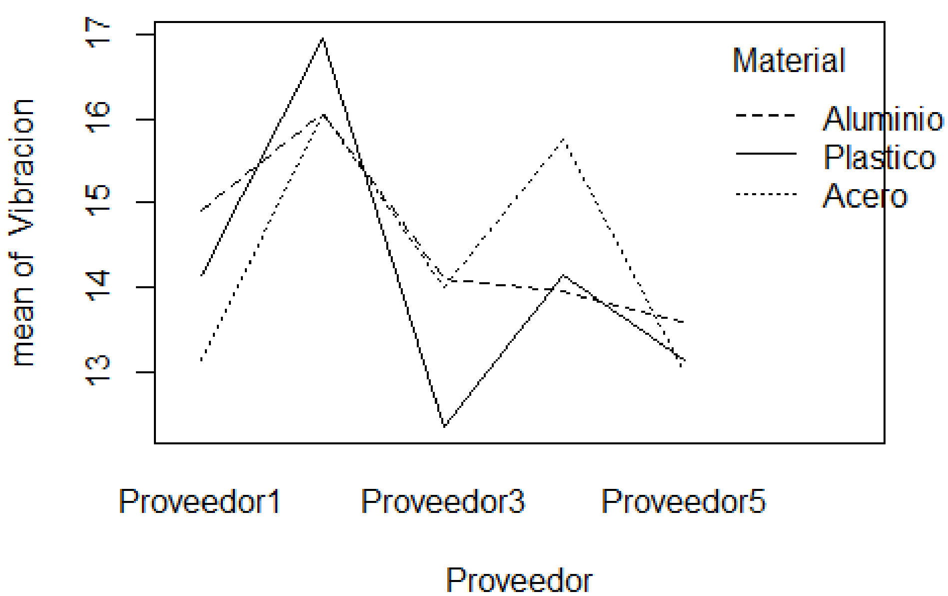
Consulta el código en R en los apoyos de clase de "ANOVA":

```r
A = aov(Vibracion ~ Proveedor*Material, datos)
summary(A)
```

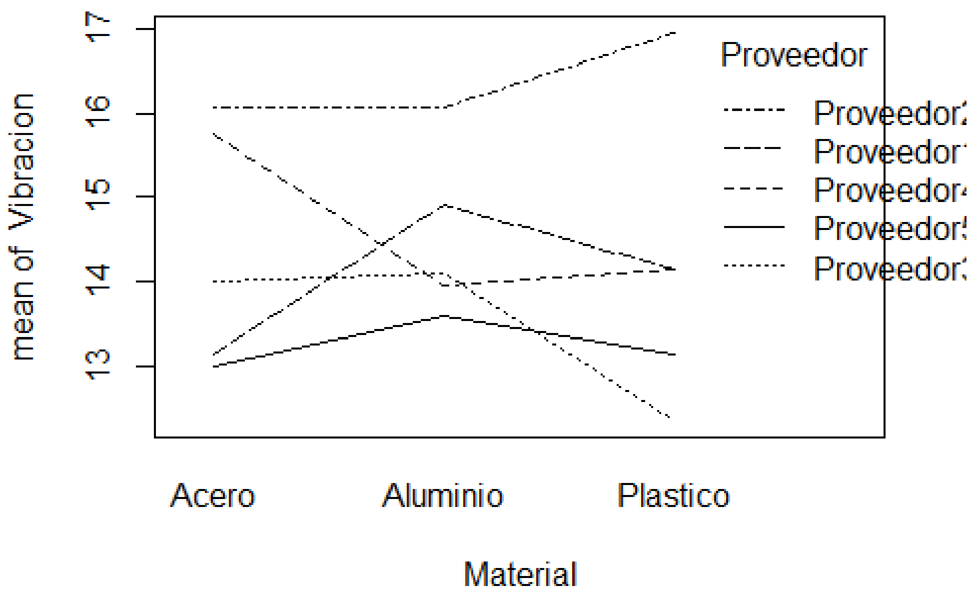
```
##                    Df Sum Sq Mean Sq F value   Pr(>F)
## Proveedor           4  36.67   9.169  82.353 5.07e-10 ***
## Material            2   0.70   0.352   3.165   0.0713 .
## Proveedor:Material  8  11.61   1.451  13.030 1.76e-05 ***
## Residuals          15   1.67   0.111
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Haz la gráfica de interacción de dos factores en ANOVA

```r
interaction.plot(Proveedor, Material, Vibracion)
```

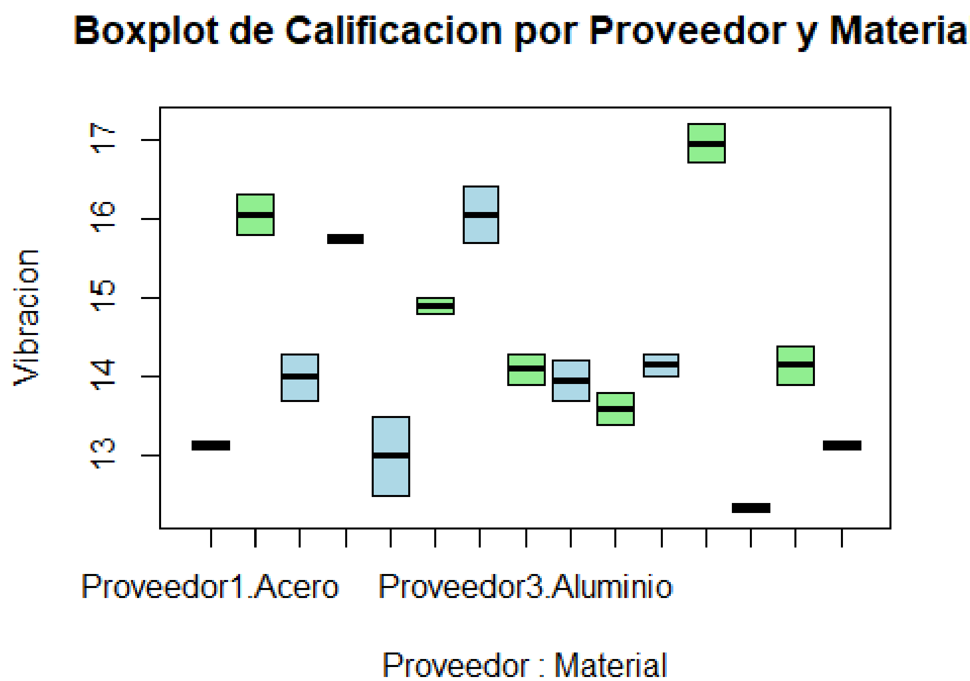



```r
interaction.plot(Material, Proveedor, Vibracion)
```



Haz el boxplot para visualizar la interacción de los factores, por ejemplo, peso por dieta interacción ejercicio:

```
boxplot(Vibracion ~ Proveedor * Material, data = datos, col =
c("lightblue", "lightgreen"), main = "Boxplot de Calificacion por
Proveedor y Material")
```



Interpreta el resultado desde la perspectiva estadística y en el contexto del problema

El análisis ANOVA revela que el proveedor tiene un efecto estadísticamente significativo en las calificaciones, con un valor p muy inferior a 0.05, indicando diferencias notables entre los proveedores P1, P2, P3, P4, P5 al igual que la interación entre el proveedor y el material también tiene un valor p muy bajo. Por otro lado, el material no moestró efectos significativos sobre las vibraciones, con valores p superiores a 0.05, lo que sugiere que este factor no influyen de manera considerable en los resultados.

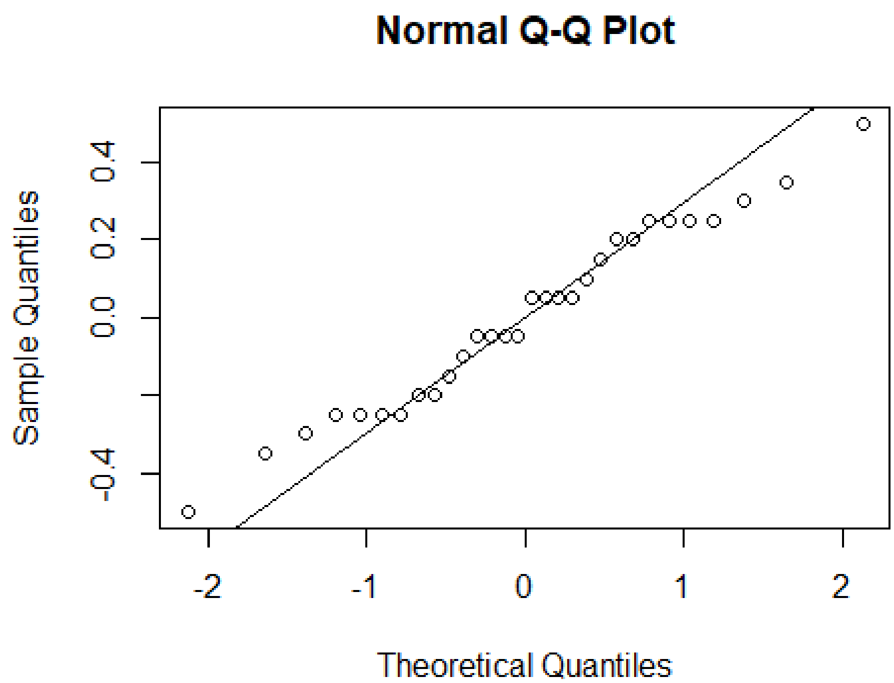
Escribe tus conclusiones parciales

Las vibraciones están significativamente influenciadas por el proveedor y la interacción entre el proveedor y el material, mientras que el material por sí solo no tiene un impacto relevante en las vibraciones. Esto implica que la elección del proveedor es clave definir las vibraciones, independientemente del material.**Sin embargo como la interacción entre el Material y el Proveedor es significativo según el anova. Comprobaremos los supuestos de dicho anova**

6. Comprueba la validez del modelo. Comprueba

Normalidad

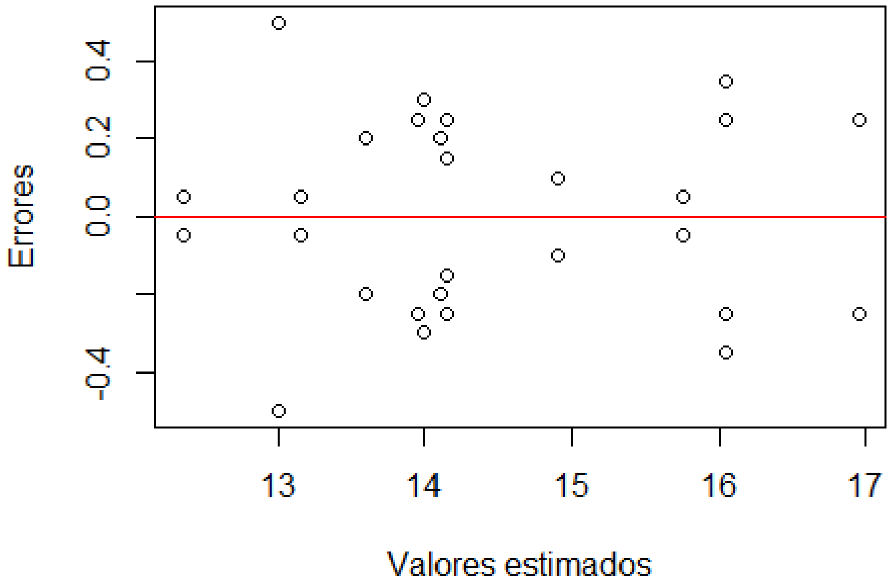
```
residuos=A$residuals
qqnorm(residuos)
qqline(residuos)
```




```
library(stats)

# Realizar la prueba de Shapiro-Wilk
shapiro_test = shapiro.test(residuos)

# Imprimir el resultado de la prueba
print(shapiro_test)

##
##  Shapiro-Wilk normality test
##
## data:  residuos
## W = 0.97627, p-value = 0.7202

# Establecer regla de decisión basada en el p-valor
if (shapiro_test$p.value < 0.05) {
  cat("Rechazar la hipótesis nula: Los datos no siguen una distribución
normal.\n")
} else {
  cat("No rechazar la hipótesis nula: Los datos siguen una distribución
normal.\n")
}

## No rechazar la hipótesis nula: Los datos siguen una distribución
normal.
```

Homocedasticidad

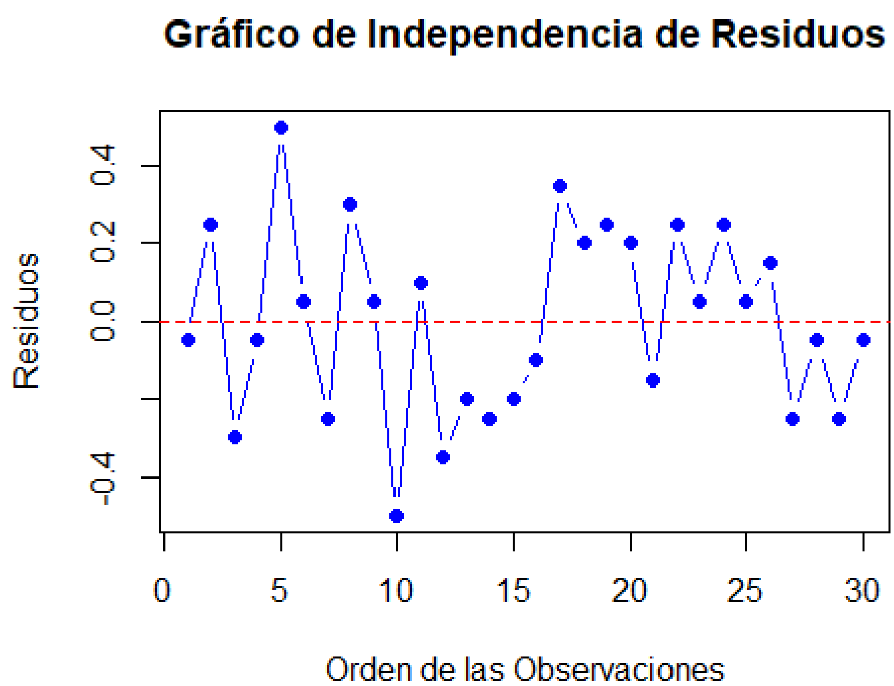
```r
plot(A$fitted.values,A$residuals,ylab="Errores",xlab="Valores estimados")
abline(h=0,col="red")
```




```r
# Cargar la librería necesaria para la prueba de Bartlett
library(stats)

# Realizar la prueba de Bartlett
bartlett_test = bartlett.test(calificacion ~ metodo, data = datos)

# Imprimir el resultado de la prueba
print(bartlett_test)

##
##  Bartlett test of homogeneity of variances
##
## data:  calificacion by metodo
## Bartlett's K-squared = 0.63268, df = 2, p-value = 0.7288

# Establecer regla de decisión basada en el p-valor
if (bartlett_test$p.value < 0.05) {
  cat("Rechazar la hipótesis nula: Las varianzas no son homogéneas.\n")
} else {
  cat("No rechazar la hipótesis nula: Las varianzas son homogéneas.\n")
}

## No rechazar la hipótesis nula: Las varianzas son homogéneas.
```

Independencia

```r
# Instalar y cargar el paquete lmtest si no está instalado
library(lmtest)

# Crear una secuencia que represente el orden de las observaciones
orden <- 1:length(residuos)

# Graficar los residuos en función del orden de las observaciones
plot(orden, residuos, type = "b", pch = 19, col = "blue",
     xlab = "Orden de las Observaciones", ylab = "Residuos",
     main = "Gráfico de Independencia de Residuos")
abline(h = 0, col ="red",lty=2)
```




```r
# Realizar la prueba de Durbin-Watson
dw_test = dwtest(A, alternative = "two.sided")

# Imprimir el resultado de la prueba
print(dw_test)

##
##  Durbin-Watson test
##
## data:  A
## DW = 1.9401, p-value = 0.9904
## alternative hypothesis: true autocorrelation is not 0

# Establecer regla de decisión basada en el p-valor
if (dw_test$p.value < 0.05) {
```

```
  cat("Rechazar la hipótesis nula: Existe evidencia de autocorrelación en los residuos.\n")
} else {
  cat("No rechazar la hipótesis nula: No hay evidencia suficiente de autocorrelación en los residuos.\n")
}

## No rechazar la hipótesis nula: No hay evidencia suficiente de
autocorrelación en los residuos.
```

Relación lineal entre las variables (coeficiente de determinación).

```
# Ajustar el modelo ANOVA
anova <- aov(Vibracion ~ Proveedor*Material, data = datos)

# Calcular el R^2
suma_cuadrados_total <- sum((datos$Vibracion - mean(datos$Vibracion))^2)
suma_cuadrados_modelo <- sum((fitted(anova) - mean(datos$Vibracion))^2)

R2 <- suma_cuadrados_modelo / suma_cuadrados_total

# Mostrar el R^2
print(paste("Coeficiente de Determinación(R^2):",R2))

## [1] "Coeficiente de Determinación(R^2): 0.967031665394435"
```

7. Concluye en el contexto del problema.

En el modelo del ANOVA con interacción dió un valor p muy bajo sugiriendo que el Proveedor y el Material tienen influencia en las Vibraciones en conjunto.

Después se aplicaron los supuestos del ANOVA que son normalidad, homocedasticidad, independencia y relación lineal, en cada prueba se ´puede ver que cuumple con cada uno de los supestos por lo que el modelo el ANOVA es válido y podemos decir que la interacción Proveedor y Material influye en las vibraciones.