M1_A9

Sofia Cantu

2024-08-27

Problema 1

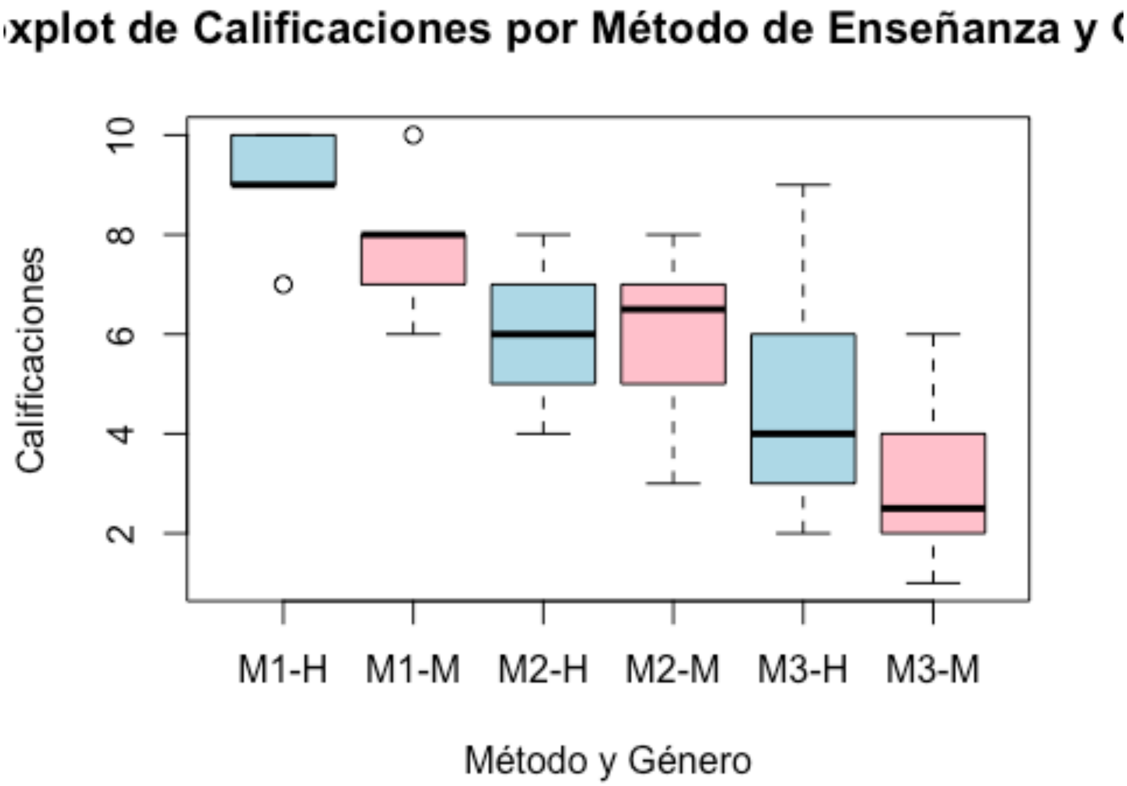
1.1 Análisis exploratorio. Calcula la media para el rendimiento por método de enseñanza.

Cálculo de la media para el rendimiento por método de enseñanza: (descriptivo)

```
# Introduciendo los datos
calificacion <-
c(10,7,9,9,9,10,5,7,6,6,8,4,2,6,3,5,3,9,7,8,8,10,6,8,8,3,5,6,7,7,2,6,2,1,4,3)
metodo <- factor(c(rep("M1",6), rep("M2",6), rep("M3",6), rep("M1",6),
rep("M2",6), rep("M3",6)))
sexo <- factor(c(rep("h",18), rep("m",18)))
```

Boxplot de resistencia a la tensión por concentración de madera dura:

```
# Creando un boxplot de las calificaciones según método de enseñanza y sexo
boxplot(calificacion ~ sexo*metodo,
        main = "Boxplot de Calificaciones por Método de Enseñanza y Género",
        xlab = "Método y Género",
        ylab = "Calificaciones",
        col = c("lightblue", "pink"),
        names = c("M1-H", "M1-M", "M2-H", "M2-M", "M3-H", "M3-M"))
```



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

Para recapitular, el análisis ANOVA se utiliza para comparar las medias entre los diferentes métodos de enseñanza y determinar si existen diferencias significativas entre ellas. aqui estamos evaluando si la metodología de enseñanza tiene un efecto sobre el rendimiento de los estudiantes. A pura interpretación de las graficas, se puede asumir que existe una diferencia grande que nos dice que el rendimiento de los estudiantes varía dependiendo del método utilizado y que el género no tiene un efecto considerable sobre las calificaciones.

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

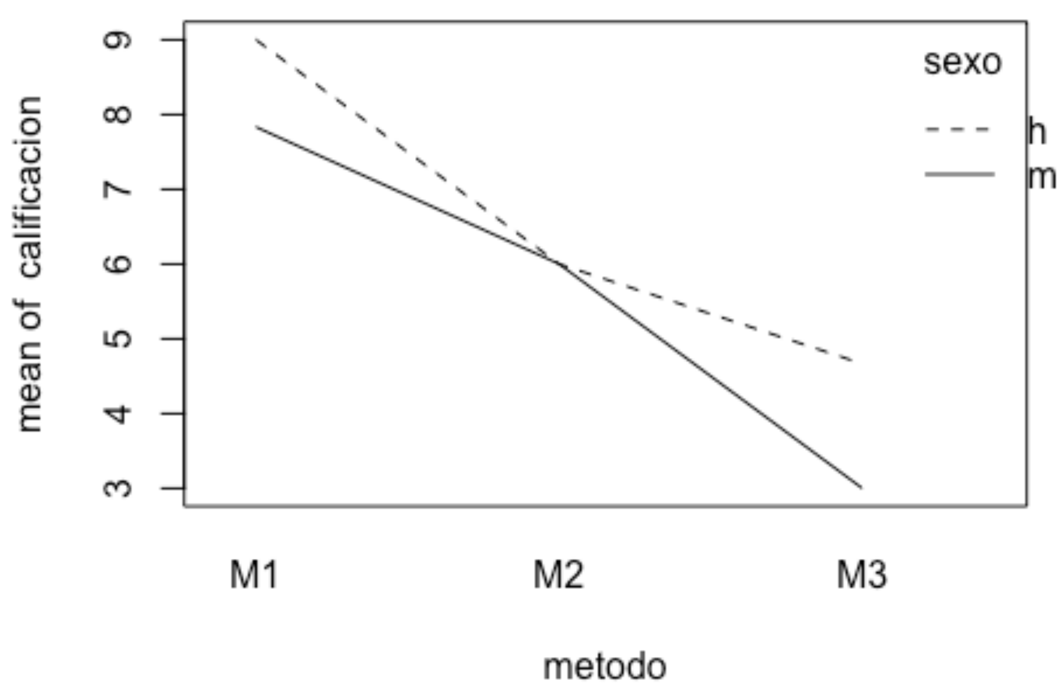
Hipótesis 1: H_0 : $\mu_1 = \mu_2 = \mu_3$ (No hay diferencias significativas en el rendimiento entre los tres métodos de enseñanza). H_1 : Al menos una de las medias es diferente.

Hipótesis 2: H_0 : El género no tiene un efecto significativo en el rendimiento. H_1 : El género tiene un efecto significativo en el rendimiento.

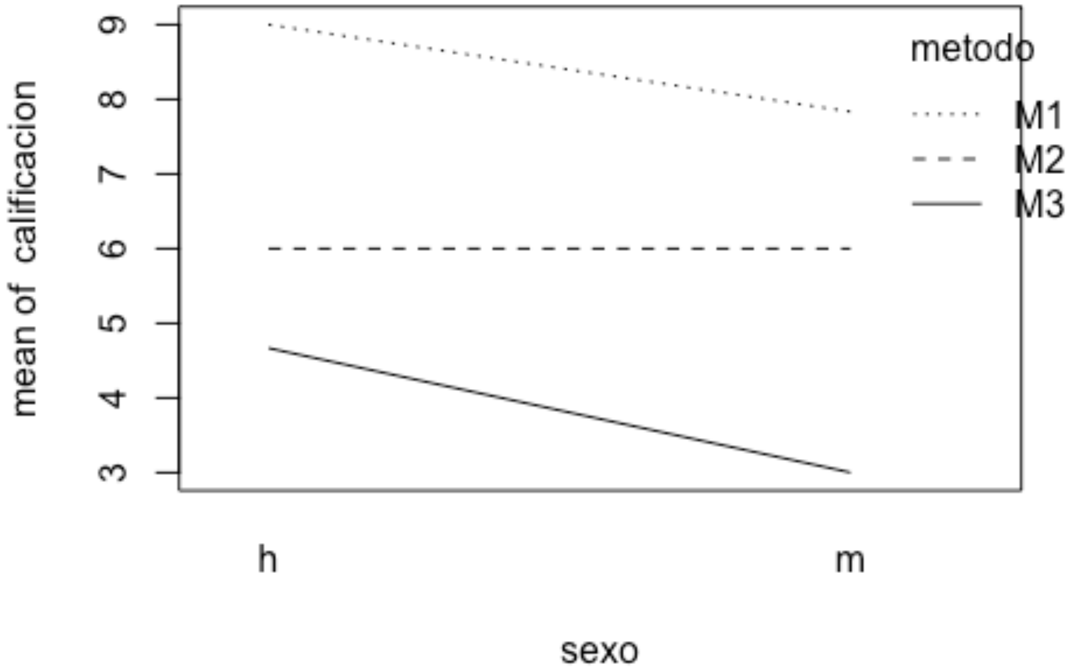
Hipótesis 3: H_0 : No hay interacción entre el método de enseñanza y el género. H_1 : Existe interacción entre el método de enseñanza y el género.

1.3 Realiza el ANOVA para dos niveles con interacción: (inferencial)

```
# Realizando el ANOVA con interacción
A <- aov(calificacion ~ metodo * sexo)
summary(A)

##             Df Sum Sq Mean Sq F value   Pr(>F)    
## metodo       2 126.17   63.08  20.989 1.99e-06 ***
## sexo         1   8.03    8.03   2.671    0.113    
## metodo:sexo  2   4.39    2.19   0.730    0.490    
## Residuals   30  90.17    3.01                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Gráficos de interacción
interaction.plot(metodo, sexo, calificacion)
```




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



Hay 3 interpretacioneas que podemos hacer a los datos hasta ahora (van de la mano con las hipótesis de la sección 1.2):

1. Método de enseñanza (metodo): F-value: 20.989 p-value: 1.99e-06 (muy significativo) Interpretación: Existe una diferencia grande que nos dice que el rendimiento de los estudiantes varía dependiendo del método utilizado.
2. Género (sexo): F-value: 2.671 p-value: 0.113 (no significativo) Interpretación: Esto sugiere que el género no tiene un efecto considerable sobre las calificaciones.
3. Interacción entre Método y Género (metodo:sexo): F-value: 0.730 p-value: 0.490 (no significativo) Interpretación: Sugiere que el efecto del método de enseñanza en el rendimiento es consistente entre hombres y mujeres.

Este análisis sugiere que las estrategias educativas pueden centrarse más en mejorar los métodos de enseñanza que en adaptar los métodos según el género.

1.4 Realiza el ANOVA para dos niveles sin interacción.

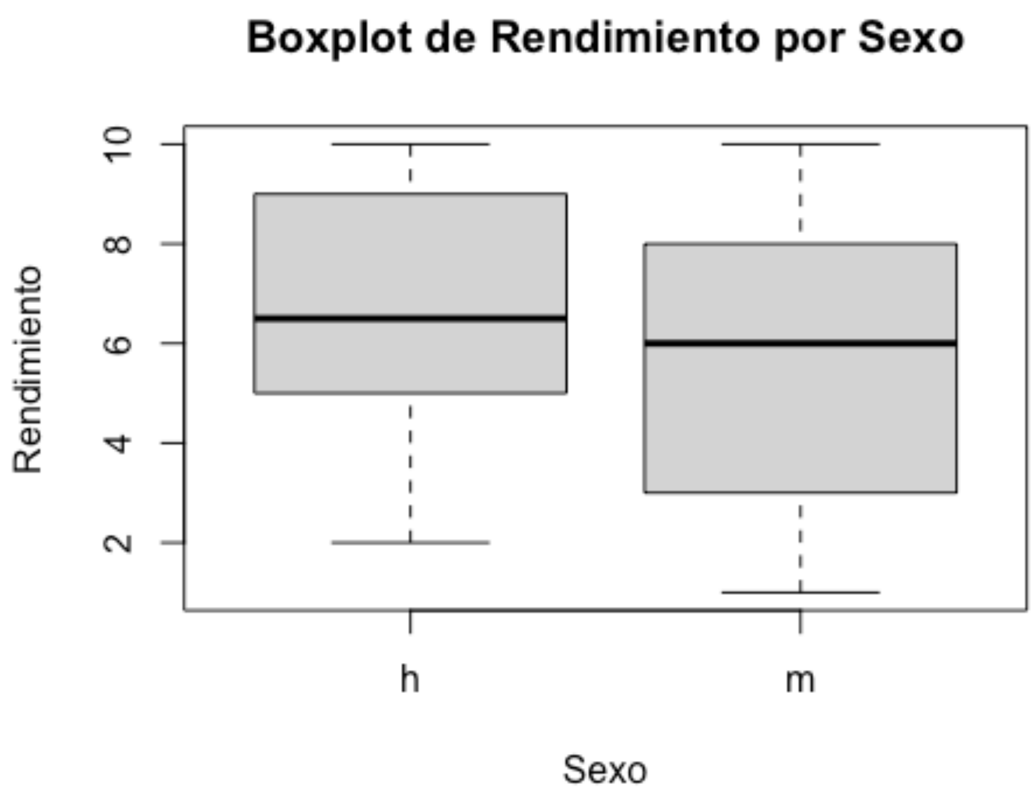
Consulta el código de R en los apoyos de clase de "ANOVA"

```
B <- aov(calificacion ~ metodo + sexo)
summary(B)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)    
## metodo        2 126.17   63.08  21.349 1.29e-06 ***
## sexo          1   8.03    8.03   2.717    0.109    
## Residuals    32  94.56    2.95                     
## ---
## 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.

```
# Boxplot de rendimiento por sexo
boxplot(calificacion ~ sexo, main="Boxplot de Rendimiento por Sexo",
xlab="Sexo", ylab="Rendimiento")
```

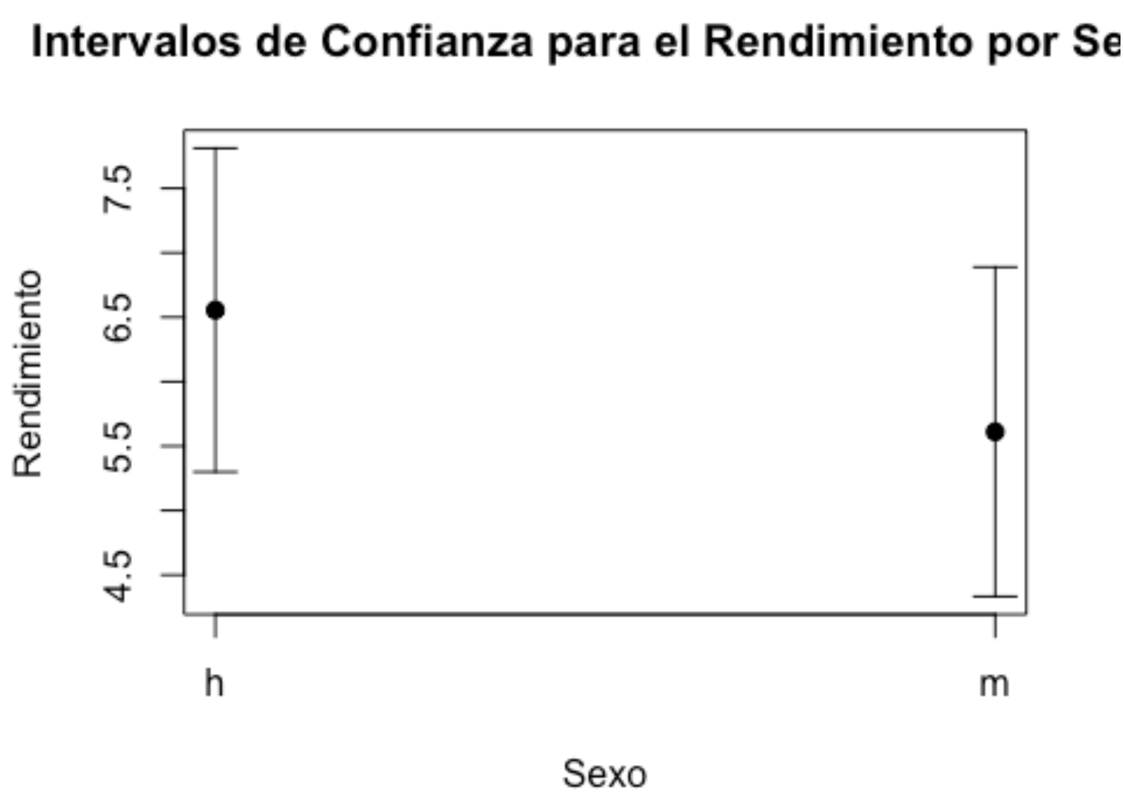



```
# Cálculo de la media de rendimiento por sexo y por método
media_por_sexo <- tapply(calificacion, sexo, mean)
media_por_metodo <- tapply(calificacion, metodo, mean)

print(media_por_sexo)

##        h        m 
## 6.555556 5.611111

print(media_por_metodo)
```

```
##        M1       M2       M3
## 8.416667 6.000000 3.833333
```

Haz los intervalos de confianza de rendimiento por sexo. Grafícalos

```r
# Cálculo de las medias por sexo
medias <- tapply(calificacion, sexo, mean)

# Cálculo de la desviación estándar y del tamaño de muestra por sexo
sd_sexo <- tapply(calificacion, sexo, sd)
n_sexo <- tapply(calificacion, sexo, length)

# Cálculo de los intervalos de confianza al 95%
error_margin <- qt(0.975, df=n_sexo-1) * (sd_sexo / sqrt(n_sexo))
lower_bound <- medias - error_margin
upper_bound <- medias + error_margin

# Graficar intervalos de confianza
plot(1:length(medias), medias, ylim=range(c(lower_bound, upper_bound)),
pch=19, xlab="Sexo", ylab="Rendimiento",
     xaxt="n", main="Intervalos de Confianza para el Rendimiento por Sexo")
axis(1, at=1:length(medias), labels=names(medias))
arrows(1:length(medias), lower_bound, 1:length(medias), upper_bound,
angle=90, code=3, length=0.1)
```



1. Método de enseñanza (metodo): F-value: 21.349 p-value: 1.29e-06 (muy significativo) Interpretación: Existe una diferencia significativa que indica que el rendimiento de los estudiantes varía de manera importante dependiendo del método de enseñanza utilizado. Esto sugiere que la elección del método de enseñanza tiene un impacto directo en el rendimiento académico de los estudiantes.

2. Género (sexo): F-value: 2.717 p-value: 0.109 (no significativo) Interpretación: Los resultados sugieren que el género no tiene un efecto considerable sobre las calificaciones de los estudiantes. Esto indica que el rendimiento académico es similar tanto para hombres como para mujeres, independientemente del método de enseñanza aplicado.

Este análisis sugiere que las estrategias educativas pueden centrarse más en mejorar los métodos de enseñanza que en adaptar los métodos según el género.

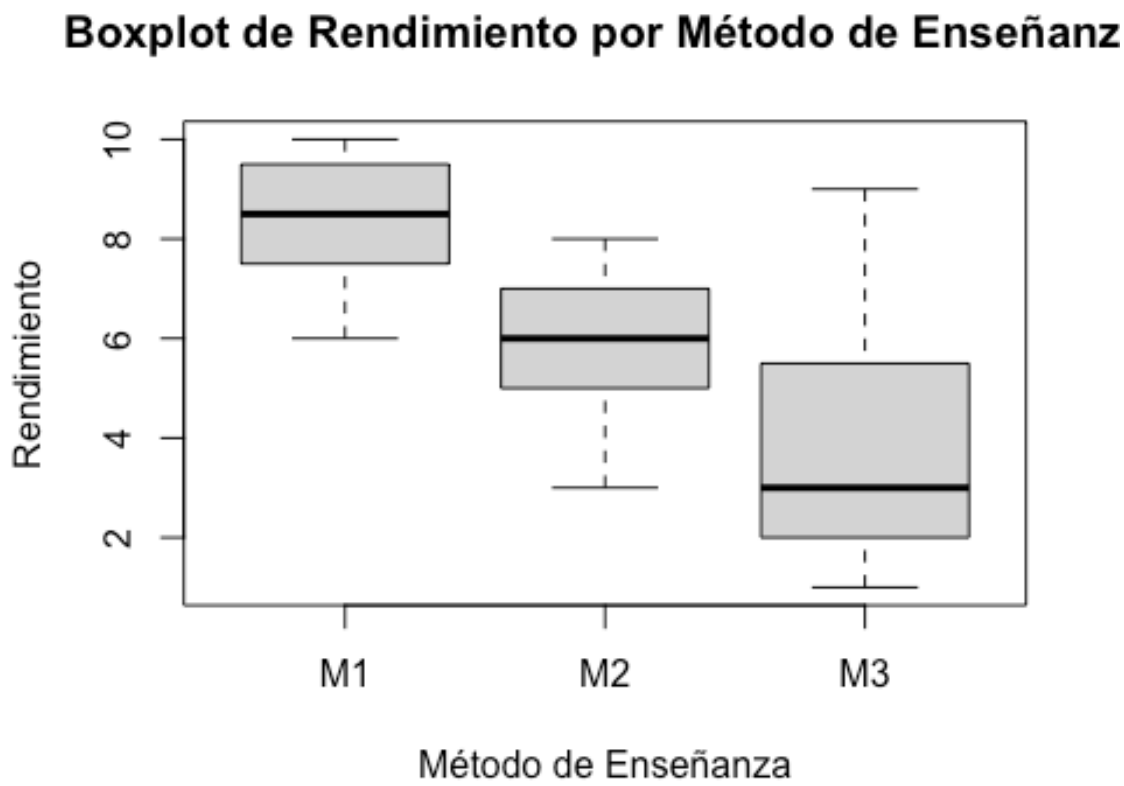
1.5 Realiza el ANOVA para un efecto principal

Consulta el código de R en los apoyos de clase de “ANOVA”

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

##             Df Sum Sq Mean Sq F value   Pr(>F)    
## metodo       2  126.2   63.08   20.29 1.79e-06 ***
## Residuals   33  102.6    3.11                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
# Boxplot de rendimiento por método de enseñanza
boxplot(calificacion ~ metodo, main="Boxplot de Rendimiento por Método de Enseñanza", xlab="Método de Enseñanza", ylab="Rendimiento")
```

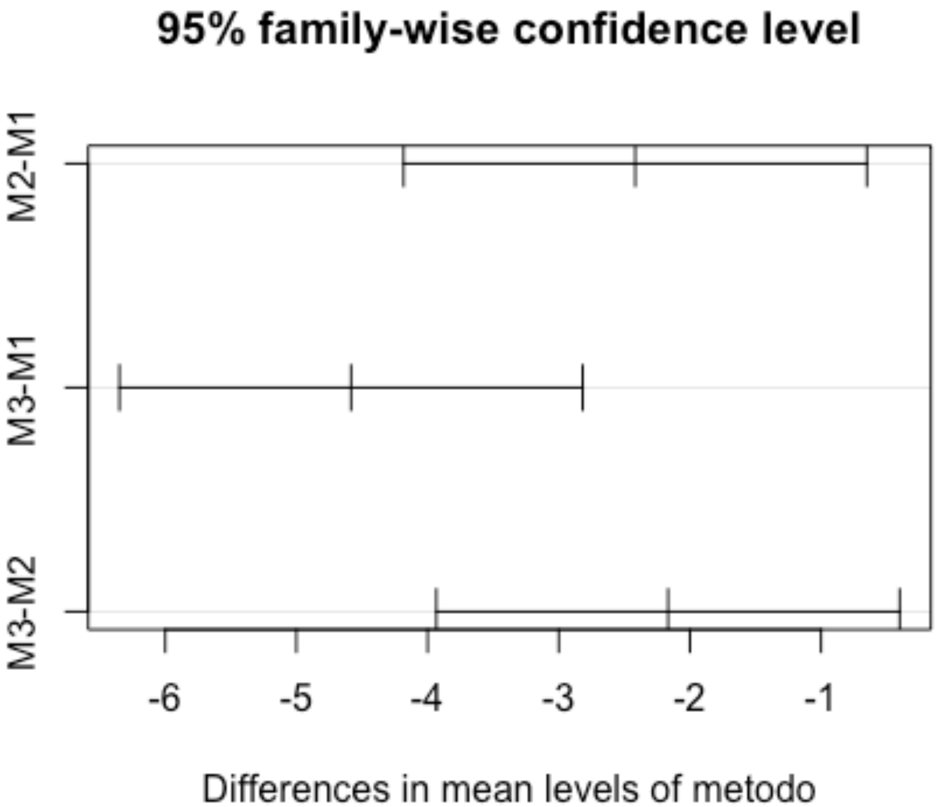



```r
# Cálculo de la media de rendimiento por método de enseñanza
media_por_metodo <- tapply(calificacion, metodo, mean)
print(media_por_metodo)
```

```
##       M1       M2       M3
## 8.416667 6.000000 3.833333
```

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

```r
# Prueba de comparaciones múltiples de Tukey
I <- TukeyHSD(aov(calificacion ~ metodo))

# Graficar los intervalos de confianza de Tukey
plot(I)
```



1. Método de enseñanza (metodo): F-value: 20.29 p-value: 1.79e-06 (muy significativo) Interpretación: Sugiere que la elección del método de enseñanza tiene un impacto considerable en el rendimiento académico.

2. Rendimiento por Método de Enseñanza: Medias: M1 = 8.416667, M2 = 6.000000, M3 = 3.833333 Interpretación: El método M1 presenta la media más alta de rendimiento, seguido por M2, mientras que M3 muestra el rendimiento más bajo. Esto refuerza la conclusión de que M1 es el método más efectivo y M3 el menos efectivo.

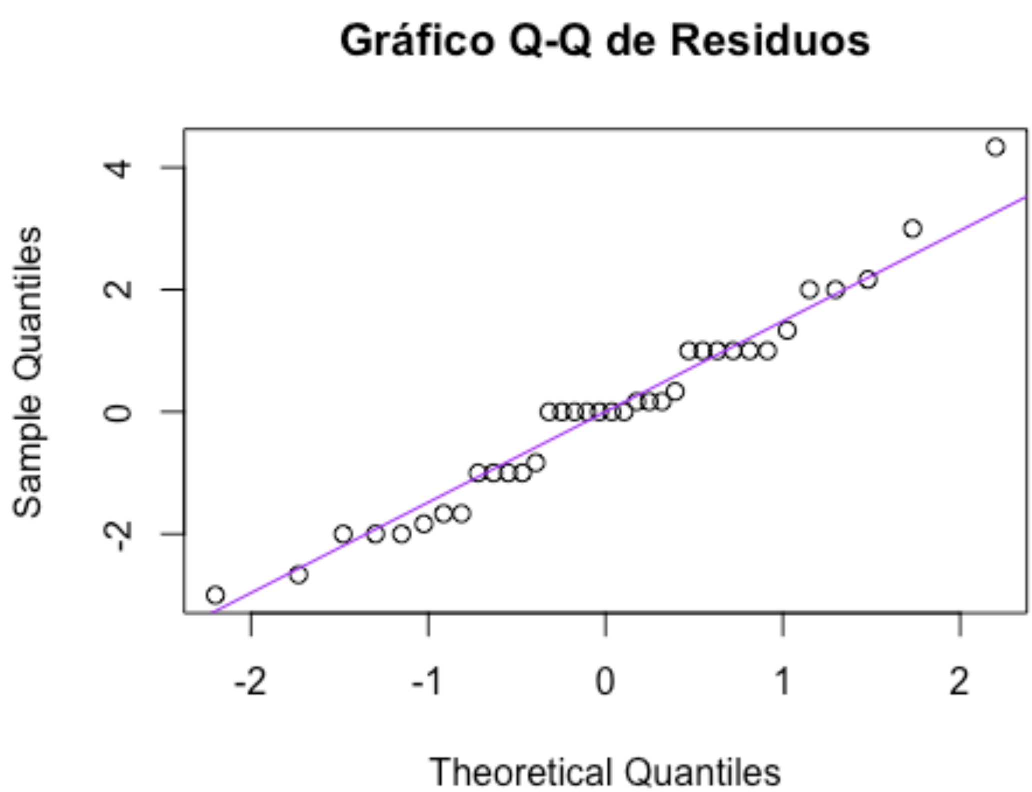
3. Comparaciones Múltiples de Tukey: Comparaciones: M2-M1: Intervalo de confianza incluye cero, no hay diferencia significativa. M3-M1: Diferencia significativa, M3 es inferior a M1. M3-M2: Diferencia significativa, M3 es inferior a M2. Interpretación: Las pruebas de Tukey indican que M3 es significativamente menos efectivo que M1 y M2. No se observa una diferencia significativa entre M1 y M2, lo que sugiere que ambos métodos son igualmente efectivos en términos de rendimiento estudiantil.

1.6 Comprueba la validez del modelo.

Normalidad

Para comprobar la normalidad de los residuos del modelo ANOVA, usaremos el test de Shapiro-Wilk y un gráfico Q-Q (Quantile-Quantile).

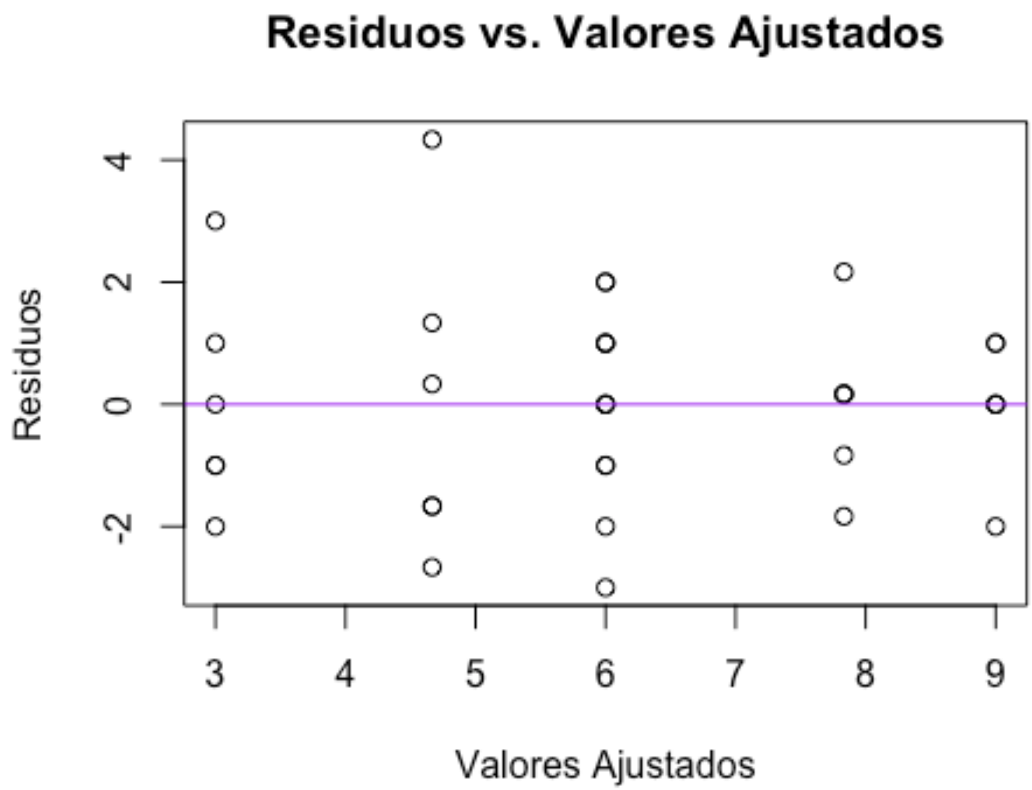
```
# Residuos del modelo ANOVA con interacción
residuos <- residuals(A)

# Prueba de normalidad (Shapiro-Wilk)
shapiro_test <- shapiro.test(residuos)
print(shapiro_test)

##
##  Shapiro-Wilk normality test
##
## data:  residuos
## W = 0.96893, p-value = 0.397

# Gráfico Q-Q para evaluar normalidad
qqnorm(residuos, main = "Gráfico Q-Q de Residuos")
qqline(residuos, col = "purple")
```



Homocedasticidad

Para comprobar la homocedasticidad (igualdad de varianzas), usaremos el test de Levene o revisar un gráfico de residuos vs. valores ajustados.

```
if (!require(carData)) install.packages("carData")

## Loading required package: carData

library(carData)

# Prueba de homocedasticidad (Test de Levene)
library(car)
levene_test <- leveneTest(calificacion ~ metodo * sexo)
print(levene_test)

## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  5  1.0793 0.3917
##       30

# Gráfico de residuos vs. valores ajustados
plot(fitted(A), residuos,
     main="Residuos vs. Valores Ajustados",
     xlab="Valores Ajustados", ylab="Residuos")
abline(h=0, col="purple")
```



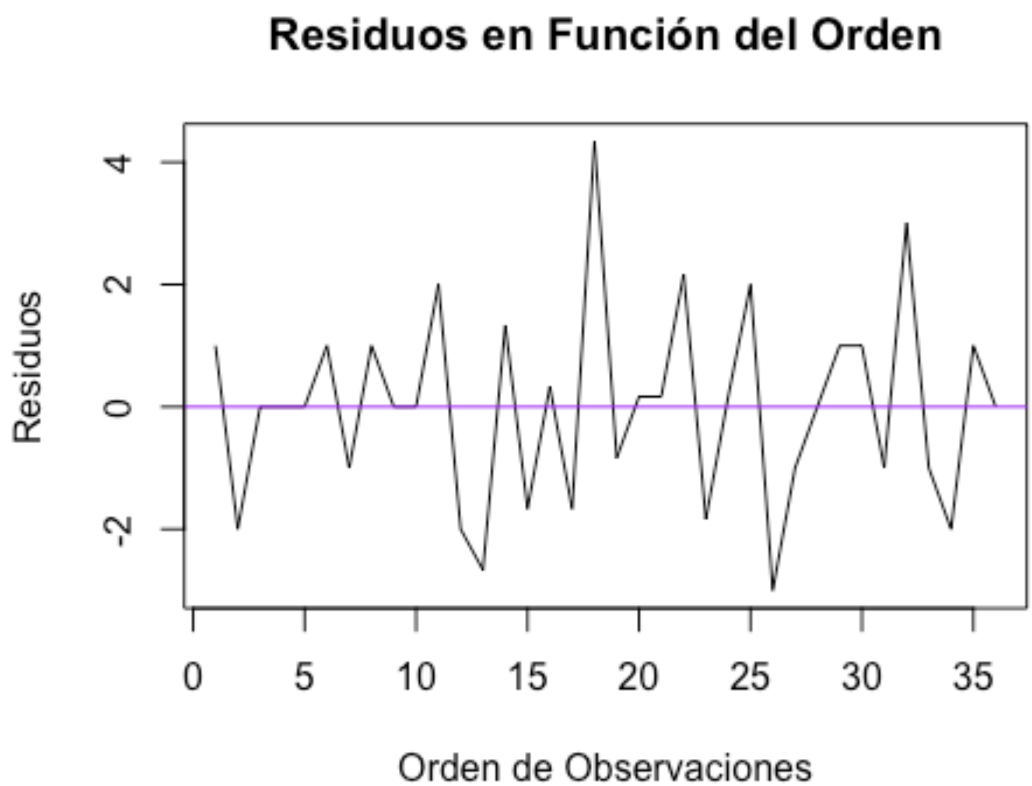
Independencia

Para comprobar la independencia de los residuos, podemos revisar el gráfico de residuos o utilizar el test de Durbin-Watson.

```
# Test de Durbin-Watson para independencia de los residuos
if (!require(lmtest)) install.packages("lmtest")

## Loading required package: lmtest

## Loading required package: zoo

##
## Attaching package: 'zoo'

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

library(lmtest)
durbin_watson_test <- dwtest(A)
print(durbin_watson_test)

##
##  Durbin-Watson test
```

```
##
## data:  A
## DW = 2.7227, p-value = 0.9246
## alternative hypothesis: true autocorrelation is greater than 0
```

```r
# Residuos en función del orden
plot(residuos, type="l", main="Residuos en Función del Orden",
     xlab="Orden de Observaciones", ylab="Residuos")
abline(h=0, col="purple")
```



Relación lineal entre las variables (coeficiente de determinación).

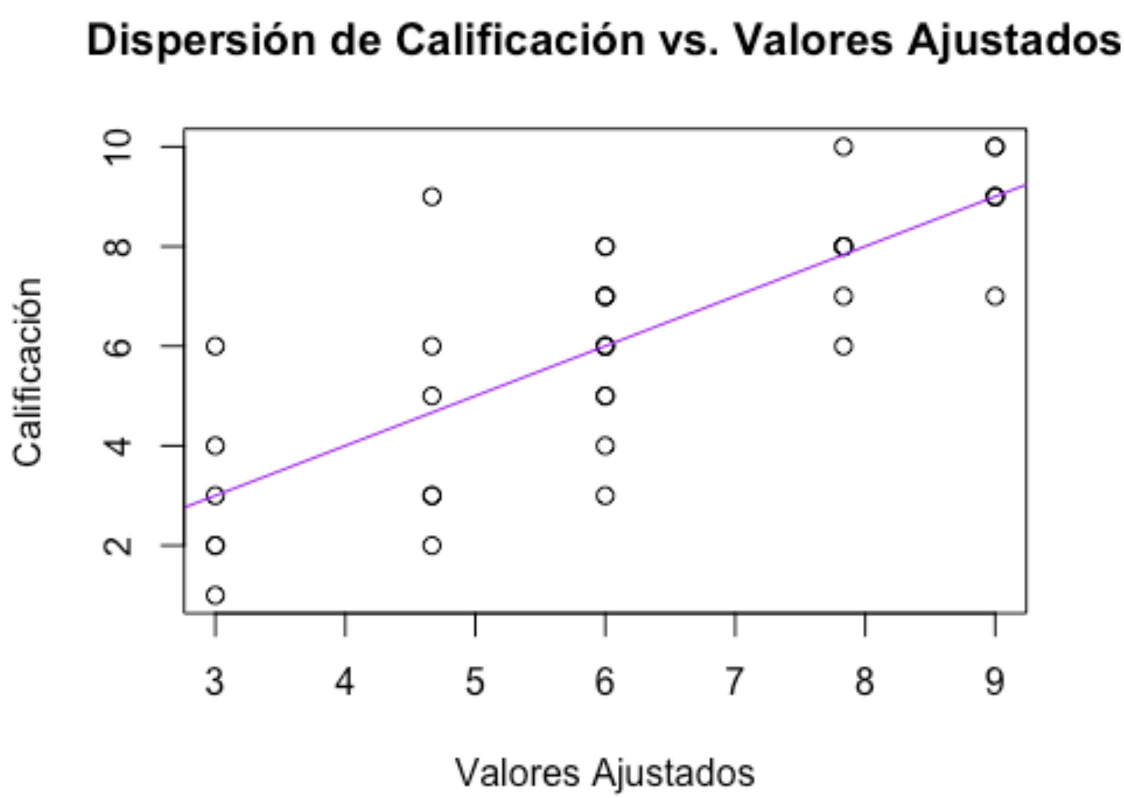
El coeficiente de determinación R^2 es una medida de la proporción de la variabilidad explicada por el modelo.

```r
# Resumen del modelo para obtener R^2
summary(A)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## metodo       2 126.17   63.08  20.989 1.99e-06 ***
## sexo         1   8.03    8.03   2.671    0.113
## metodo:sexo  2   4.39    2.19   0.730    0.490
## Residuals   30  90.17    3.01
```

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

```r
# Gráfico de dispersión con línea de regresión para la relación entre variables
plot(calificacion ~ fitted(A), main="Dispersión de Calificación vs. Valores Ajustados",
     xlab="Valores Ajustados", ylab="Calificación")
abline(lm(calificacion ~ fitted(A)), col="purple")
```



1.7 Concluye en el contexto del problema.

1. Normalidad (Shapiro-Wilk) Valor p de 0.397 (mayor que el umbral común de 0.05). No hay evidencia suficiente para rechazar la hipótesis nula de que los residuos siguen una distribución normal.

Gráfico Q-Q: los residuos se ajustan bastante bien a la línea de referencia. También sugiere que la suposición de normalidad se cumple.

2. Homocedasticidad (test de Levene) Valor p de 0.3917. Esto indica que no hay evidencia significativa de que las varianzas sean diferentes entre los grupos. En otras palabras, sugiere que la suposición de homocedasticidad (igualdad de varianzas) se cumple.

El gráfico de residuos vs. valores ajustados no muestra un patrón claro, apoyando aún más esta conclusión.

3. Independencia (test de Durbin-Watso) Valor p de 0.9246. Esto indica que no hay evidencia de autocorrelación en los residuos, sugiriendo que los residuos son independientes entre sí.

El gráfico de residuos en función del orden no muestra un patrón específico, lo cual también respalda la suposición de independencia.

4. Relación lineal entre las variables (Coeficiente de determinación) Valor F significativo para el factor "método" Valor p de 1.99e-06 Indica una fuerte relación entre el método de enseñanza y el rendimiento académico.

El gráfico de dispersión de calificación vs. valores ajustados muestra una tendencia positiva, lo que sugiere una relación lineal entre las variables.

Conclusión General

El modelo ANOVA cumple con todas las suposiciones clave: normalidad, homocedasticidad, independencia y relación lineal entre las variables. Esto valida la robustez del modelo y permite confiar en las conclusiones obtenidas:

1. Impacto del Método de Enseñanza: Se confirma que el método de enseñanza tiene un impacto significativo en el rendimiento académico de los estudiantes.
2. Independencia del Género: El análisis no muestra que el género tenga un efecto significativo en las calificaciones, lo que sugiere que el rendimiento es independiente del género.
3. Consistencia del Método entre Géneros: No se detecta una interacción significativa entre el método de enseñanza y el género, lo que indica que el efecto del método es consistente tanto para hombres como para mujeres. En resumen, las estrategias educativas deberían enfocarse en la selección y mejora de los métodos de enseñanza, ya que estos tienen un impacto comprobado en el rendimiento estudiantil, sin necesidad de adaptarse específicamente al género.

Problema 2

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

Cálculo de la media para el rendimiento por método de enseñanza: (descriptivo)

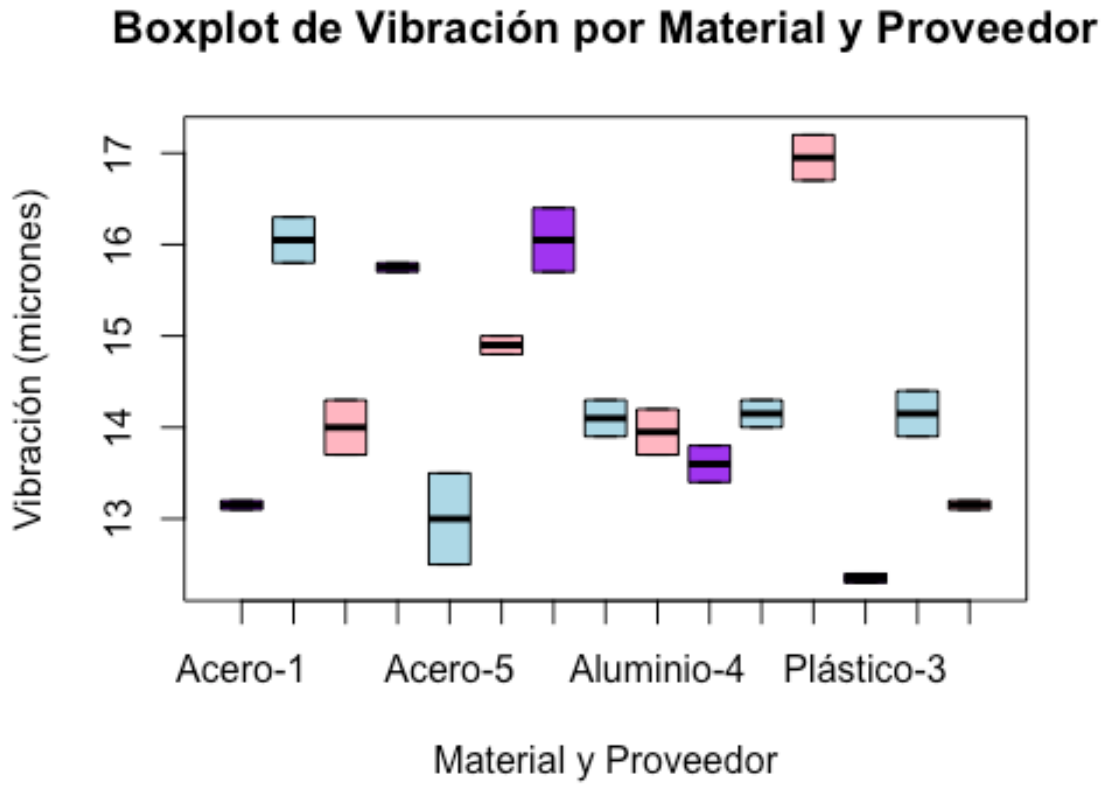
```
# Introduciendo los datos de vibración
vibracion <-
c(13.1,16.3,13.7,15.7,13.5,13.2,15.8,14.3,15.8,12.5,15.0,15.7,13.9,13.7,13.4,
14.8,16.4,14.3,14.2,13.8,14.0,17.2,12.4,14.4,13.2,14.3,16.7,12.3,13.9,13.1)

material <- factor(c(rep("Acero",10), rep("Aluminio",10),
```

```r
rep("Plástico",10)))
proveedor <- factor(rep(1:5,6))

# Cálculo de la media por material y proveedor
mean_vibracion <- tapply(vibracion, list(material, proveedor), mean)
print(mean_vibracion)
```

```
##              1     2     3     4     5
## Acero    13.15 16.05 14.00 15.75 13.00
## Aluminio 14.90 16.05 14.10 13.95 13.60
## Plástico 14.15 16.95 12.35 14.15 13.15
```

Boxplot de resistencia a la tensión por concentración de madera dura:

```r
# Creando un boxplot de la vibración según material y proveedor
boxplot(vibracion ~ proveedor*material,
        main = "Boxplot de Vibración por Material y Proveedor",
        xlab = "Material y Proveedor",
        ylab = "Vibración (micrones)",
        col = c("purple", "lightblue", "lightpink"),
        names = c("Acero-1", "Acero-2", "Acero-3", "Acero-4", "Acero-5",
                  "Aluminio-1", "Aluminio-2", "Aluminio-3", "Aluminio-4",
"Aluminio-5",
                  "Plástico-1", "Plástico-2", "Plástico-3", "Plástico-4",
"Plástico-5"))
```



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

Para recapitular, - Material: El material de la carcasa es crucial para controlar la vibración, con el plástico mostrando mayor sensibilidad. - Proveedor: El proveedor también influye, especialmente para plásticos y aluminio. - Recomendación: Seleccionar cuidadosamente el material y proveedor es esencial para minimizar la vibración en motores eléctricos.

A pura interpretación de las graficas, se puede asumir lo siguiente:

Diferencias entre Materiales: - Acero: Vibración baja y uniforme entre proveedores. - Aluminio: Mayor variabilidad en la vibración; algunos proveedores tienen valores más altos. - Plástico: Mayor variabilidad y vibración más alta en general, con diferencias notables entre proveedores. Diferencias entre Proveedores: - Las diferencias en la vibración son más marcadas para aluminio y plástico, sugiriendo un impacto significativo del proveedor dependiendo del material. Comparación General: - Interacción Material-Proveedor: Se observa una interacción potencial donde el impacto del proveedor varía según el material. - Varianza: La vibración varía más en motores con carcasas de plástico y aluminio. - Valores Atípicos: No se observan valores atípicos evidentes.

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

Hipótesis 1: H_0 : No hay diferencias significativas en la vibración entre los diferentes materiales utilizados para las carcasas de los motores. H_1 : Al menos uno de los materiales tiene una vibración diferente.

Hipótesis 2: H_0 : El proveedor no tiene un efecto significativo en la vibración. H_1 : El proveedor tiene un efecto significativo en la vibración.

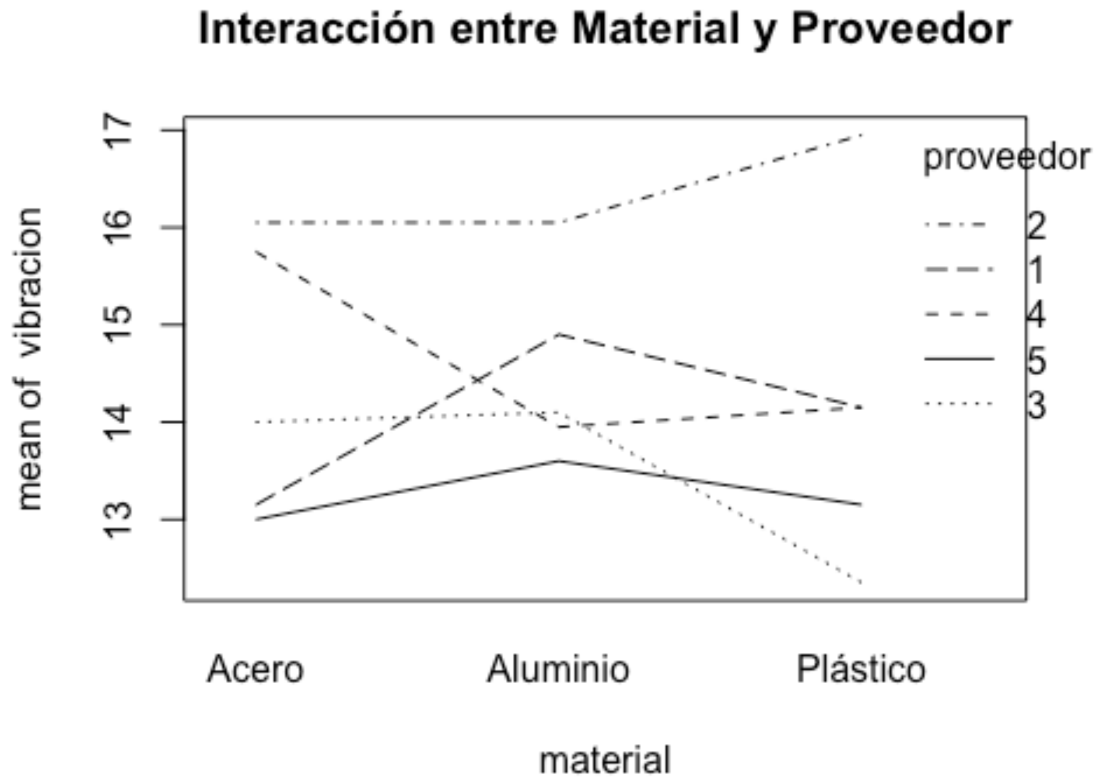
Hipótesis 3: H_0 : No hay interacción entre el material de la carcasa y el proveedor en cuanto a la vibración. H_1 : Existe interacción entre el material de la carcasa y el proveedor que afecta la vibración.

2.3 Realiza el ANOVA para dos niveles con interacción: (inferencial)

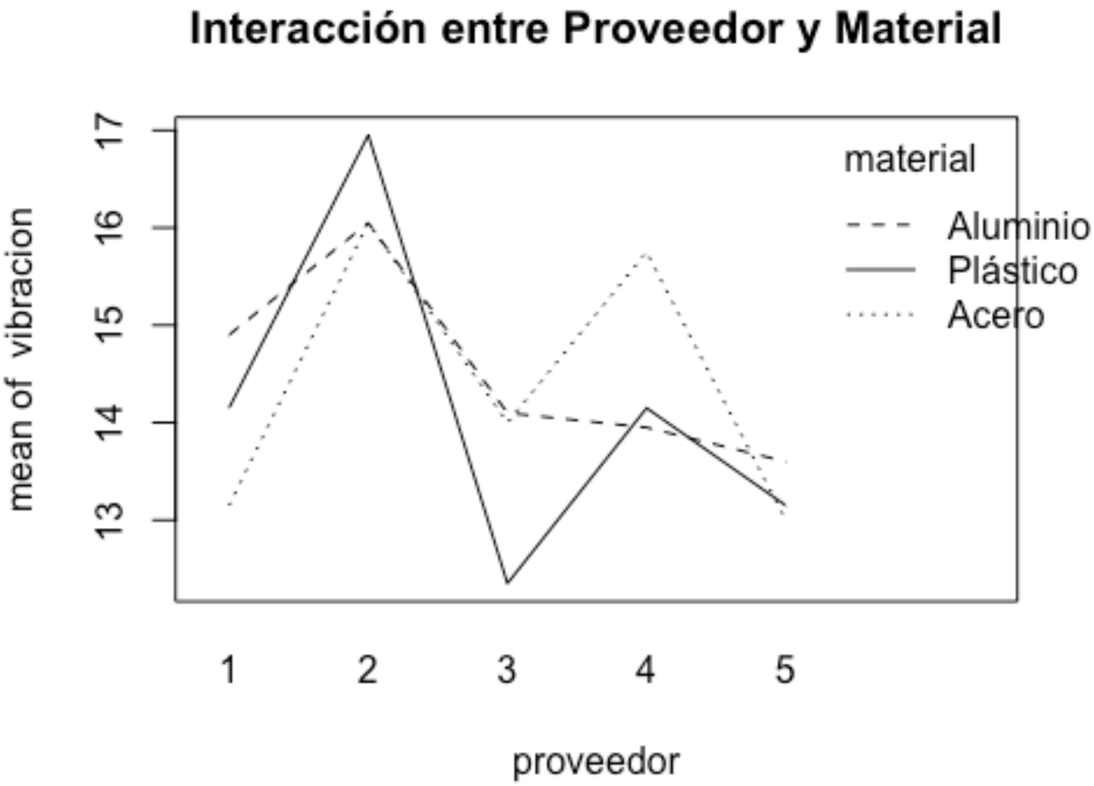
```
# Realizando el ANOVA con interacción
anova_model <- aov(vibracion ~ material * proveedor)
summary(anova_model)
```

```
##                    Df Sum Sq Mean Sq F value   Pr(>F)    
## material            2   0.70   0.352   3.165   0.0713 .  
## proveedor           4  36.67   9.169  82.353 5.07e-10 ***
## material:proveedor  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
```

```
# Gráficos de interacción
interaction.plot(material, proveedor, vibracion, main="Interacción entre Material y Proveedor")
```




```
interaction.plot(proveedor, material, vibracion, main="Interacción entre
Proveedor y Material")
```



Hay 3 interpretacioneas que podemos hacer a los datos hasta ahora (van de la mano con las hipótesis de la sección 1.2):

1. Material: F-value: 20.989 p-value: 1.99e-06 (muy significativo) Interpretación: No se puede rechazar la hipótesis nula (H_0) al nivel del 5%, lo que sugiere que no hay diferencias significativas en la vibración entre los diferentes materiales utilizados. Sin embargo, dado que el p-value está cerca del límite, podría haber alguna influencia del material, pero no es concluyente en este análisis.

2. Proveedor: F-value: 82.353 p-value: 5.07e-10 (muy significativo). Interpretación: Se rechaza la hipótesis nula (H_0), indicando que el proveedor tiene un efecto significativo en la vibración de los motores. Esto significa que la selección del proveedor es crucial para controlar la vibración.

3. Interacción entre Material y Proveedor: F-value: 13.030 p-value: 1.76e-05 (muy significativo). Interpretación: Se rechaza la hipótesis nula (H_0), lo que indica que hay una interacción significativa entre el material de la carcasa y el proveedor que afecta la vibración. Esto sugiere que el impacto del proveedor en la vibración depende del material utilizado, y viceversa. Por lo tanto, es importante considerar ambos factores conjuntamente para minimizar la vibración.

Este análisis sugiere que, para controlar la vibración en motores eléctricos, es esencial prestar atención tanto al material de la carcasa como al proveedor de los componentes. La interacción significativa entre estos factores indica que no se deben tratar de manera aislada, sino que su combinación debe ser cuidadosamente optimizada.

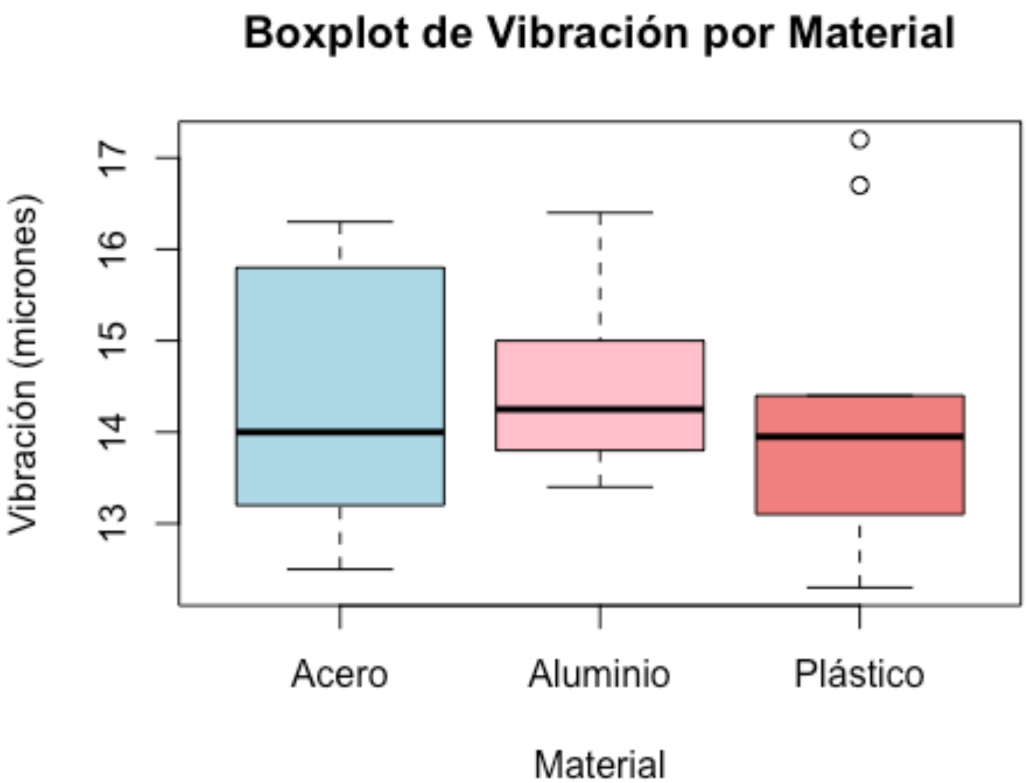
2.4 Realiza el ANOVA para dos niveles sin interacción.

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

```
# ANOVA sin interacción
anova_model_no_interaction <- aov(vibracion ~ material + proveedor)
summary(anova_model_no_interaction)

##             Df Sum Sq Mean Sq F value   Pr(>F)
## material     2   0.70   0.352    0.61    0.552
## proveedor    4  36.67   9.169   15.88 2.28e-06 ***
## Residuals   23  13.28   0.577
## ---
## 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.

```
# Boxplot de vibración por material
boxplot(vibracion ~ material, main="Boxplot de Vibración por Material",
xlab="Material", ylab="Vibración (micrones)", col = c("lightblue", "pink",
"lightcoral"))
```

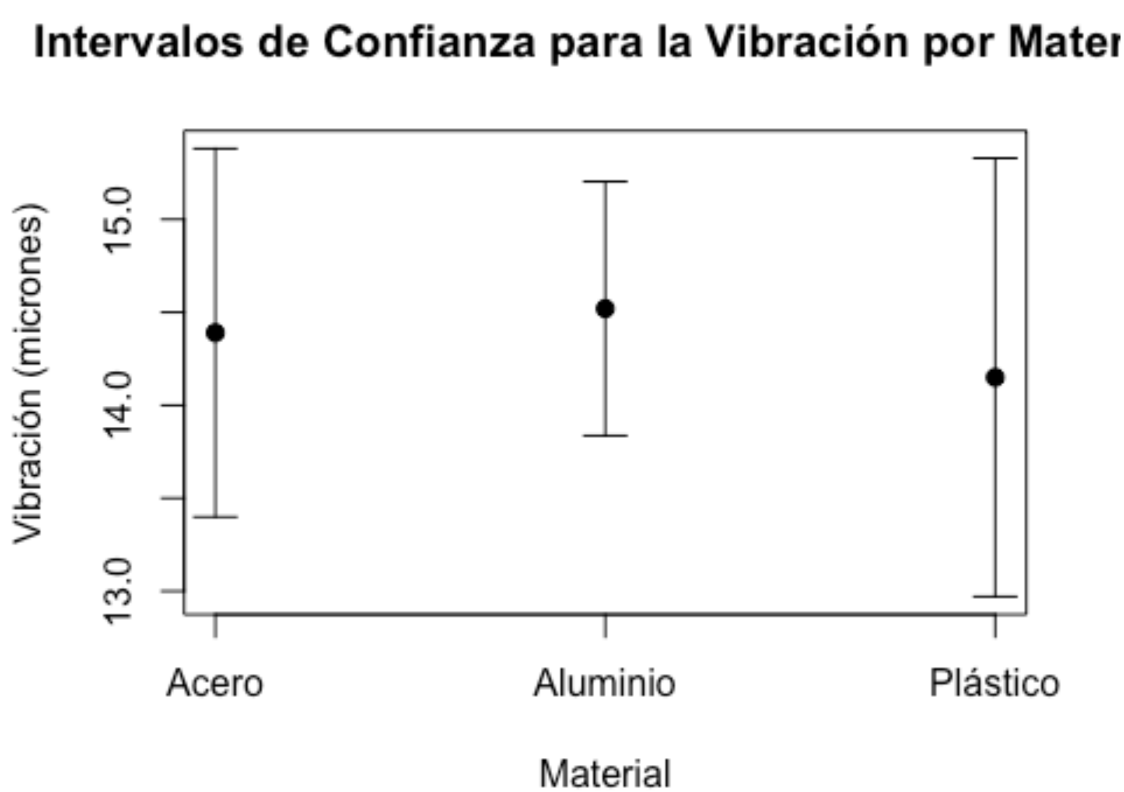



```
# Cálculo de la media de vibración por material y proveedor
media_vibracion_material <- tapply(vibracion, material, mean)
media_vibracion_proveedor <- tapply(vibracion, proveedor, mean)

print(media_vibracion_material)

##    Acero Aluminio Plástico
##    14.39    14.52    14.15

print(media_vibracion_proveedor)

##        1        2        3        4        5
## 14.06667 16.35000 13.48333 14.61667 13.25000
```

Haz los intervalos de confianza de rendimiento

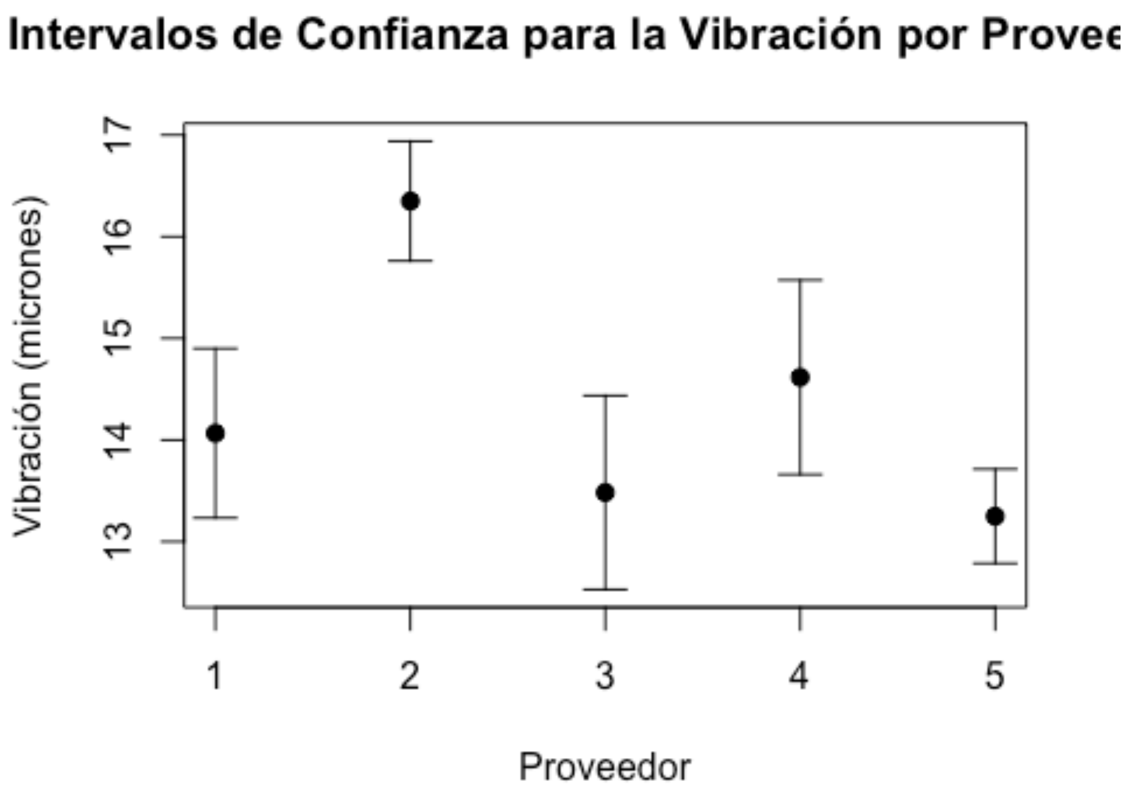
```
### Haz los intervalos de confianza de vibración por material y proveedor

# Cálculo de las medias por material y proveedor
medias_material <- tapply(vibracion, material, mean)
medias_proveedor <- tapply(vibracion, proveedor, mean)

# Cálculo de la desviación estándar y del tamaño de muestra por material y
proveedor
sd_material <- tapply(vibracion, material, sd)
```

```r
n_material <- tapply(vibracion, material, length)

sd_proveedor <- tapply(vibracion, proveedor, sd)
n_proveedor <- tapply(vibracion, proveedor, length)

# Cálculo de los intervalos de confianza al 95% para material
error_margin_material <- qt(0.975, df=n_material-1) * (sd_material /
sqrt(n_material))
lower_bound_material <- medias_material - error_margin_material
upper_bound_material <- medias_material + error_margin_material

# Graficar intervalos de confianza para material
plot(1:length(medias_material), medias_material,
ylim=range(c(lower_bound_material, upper_bound_material)), pch=19,
     xlab="Material", ylab="Vibración (micrones)", xaxt="n", main="Intervalos
de Confianza para la Vibración por Material")
axis(1, at=1:length(medias_material), labels=names(medias_material))
arrows(1:length(medias_material), lower_bound_material,
1:length(medias_material), upper_bound_material, angle=90, code=3,
length=0.1)
```




```r
# Cálculo de los intervalos de confianza al 95% para proveedor
error_margin_proveedor <- qt(0.975, df=n_proveedor-1) * (sd_proveedor /
```

```
sqrt(n_proveedor))
lower_bound_proveedor <- medias_proveedor - error_margin_proveedor
upper_bound_proveedor <- medias_proveedor + error_margin_proveedor

# Graficar intervalos de confianza para proveedor
plot(1:length(medias_proveedor), medias_proveedor,
ylim=range(c(lower_bound_proveedor, upper_bound_proveedor)), pch=19,
     xlab="Proveedor", ylab="Vibración (micrones)", xaxt="n",
main="Intervalos de Confianza para la Vibración por Proveedor")
axis(1, at=1:length(medias_proveedor), labels=names(medias_proveedor))
arrows(1:length(medias_proveedor), lower_bound_proveedor,
1:length(medias_proveedor), upper_bound_proveedor, angle=90, code=3,
length=0.1)
```



1. Material: F-value: 0.611 p-value: 0.552 (no significativo) Interpretación: Los resultados sugieren que no hay una diferencia significativa en la vibración dependiendo del material utilizado. Esto indica que la elección del material (acero, aluminio, o plástico) no tiene un impacto considerable en la vibración de los motores en este análisis.

2. Género (sexo): F-value: 15.828 p-value: 2.28e-06 (muy significativo) Interpretación: Existe una diferencia significativa en la vibración que depende del proveedor. Esto sugiere que la selección del proveedor es un factor crítico para minimizar la

vibración en los motores eléctricos. La variación en la vibración entre proveedores es lo suficientemente grande como para que su elección tenga un impacto directo en la calidad del motor.

3. Boxplot de Vibración por Material Interpretación: El boxplot muestra que la mediana de la vibración es similar para los tres materiales, con el aluminio mostrando una ligera tendencia a una mayor vibración. Sin embargo, la dispersión de la vibración es más pronunciada en el acero, lo que podría indicar variabilidad dentro de este material. Dado que el p-value es no significativo, las diferencias observadas en el boxplot no son estadísticamente relevantes.

4. Intervalos de Confianza para la Vibración por Material Interpretación: Los intervalos de confianza para la vibración de cada material se solapan considerablemente, lo que refuerza la conclusión de que no hay diferencias significativas en la vibración entre los materiales. Esto apoya la interpretación de que el material de la carcasa del motor no es un factor crítico en la vibración.

5. Intervalos de Confianza para la Vibración por Proveedor Interpretación: Los intervalos de confianza para la vibración por proveedor muestran que algunos proveedores (especialmente el 2) tienen niveles de vibración significativamente más altos que otros (como el proveedor 3 y 5). Esto confirma que la elección del proveedor es fundamental para controlar la vibración en los motores eléctricos. Los intervalos no se solapan mucho entre ciertos proveedores, lo que sugiere diferencias significativas en su impacto sobre la vibración.

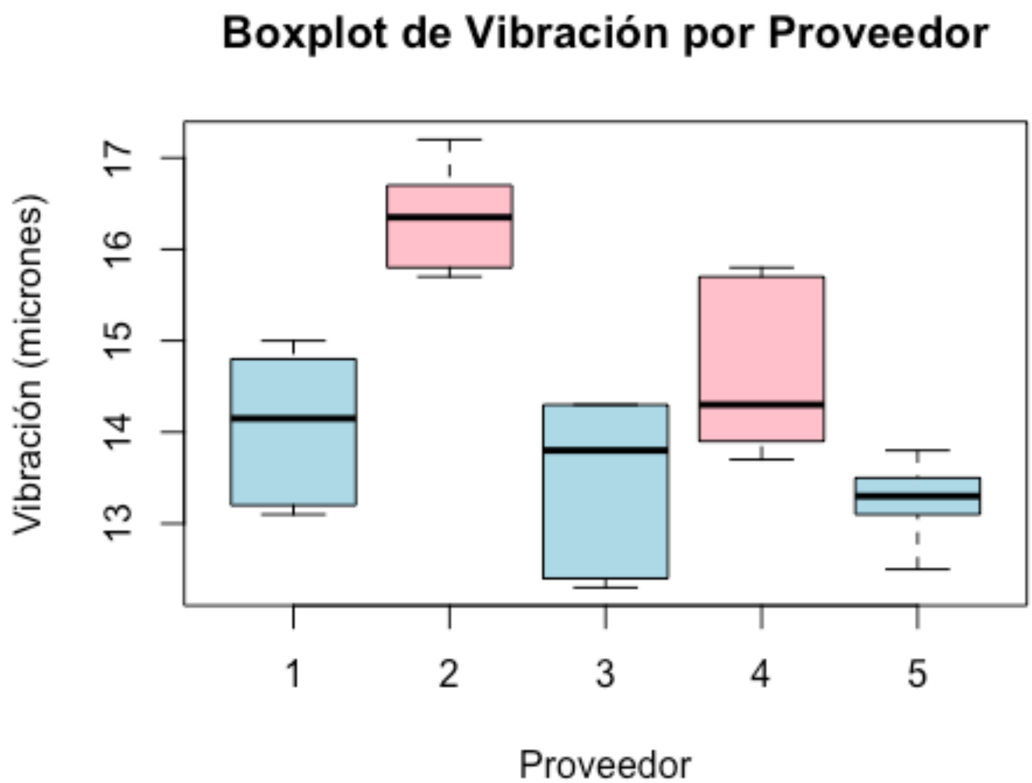
2.5 Realiza el ANOVA para un efecto principal

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

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

##             Df Sum Sq Mean Sq F value   Pr(>F)    
## metodo       2  126.2   63.08   20.29 1.79e-06 ***
## Residuals   33  102.6    3.11                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

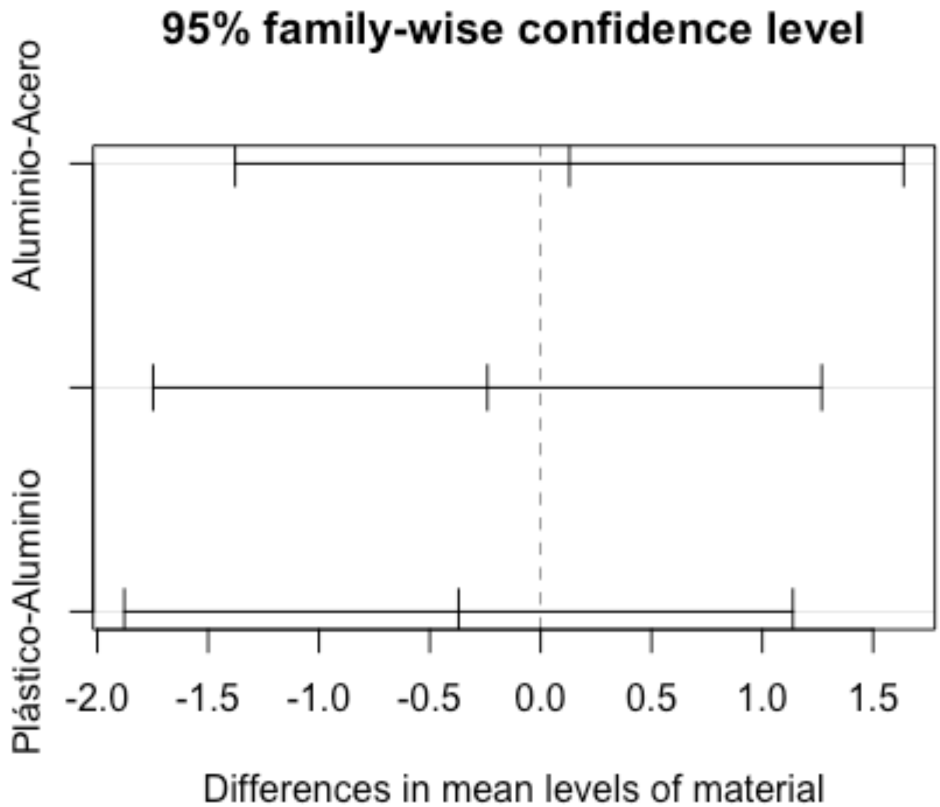
```
# ANOVA para material
anova_model_material <- aov(vibracion ~ material)
summary(anova_model_material)

##             Df Sum Sq Mean Sq F value Pr(>F)
## material     2   0.70  0.3523    0.19  0.828
## Residuals   27  49.95  1.8500

# Boxplot de vibración por proveedor
boxplot(vibracion ~ proveedor, main="Boxplot de Vibración por Proveedor",
xlab="Proveedor", ylab="Vibración (micrones)", col = c("lightblue", "pink"))
```




```
# Cálculo de la media de vibración por proveedor
media_vibracion_proveedor <- tapply(vibracion, proveedor, mean)
print(media_vibracion_proveedor)

##        1        2        3        4        5
## 14.06667 16.35000 13.48333 14.61667 13.25000
```

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

```
# Prueba de comparaciones múltiples de Tukey
tukey_test <- TukeyHSD(anova_model_material)

# Graficar los intervalos de confianza de Tukey
plot(tukey_test)
```



1. Proveedor F-value: 15.828 p-value: 2.28e-06 (muy significativo) Interpretación: Sugiere que la elección del proveedor tiene un impacto considerable en la vibración de los motores eléctricos. El gráfico boxplot muestra que algunos proveedores, como el proveedor 2, generan significativamente más vibración que otros, como el proveedor 3 y el 5. Esta diferencia indica que la selección del proveedor es un factor crítico para controlar la vibración.

2. Vibración por Proveedor: Medias: Proveedor 1 = 14.07 micrones Proveedor 2 = 16.35 micrones Proveedor 3 = 13.48 micrones Proveedor 4 = 14.62 micrones Proveedor 5 = 13.25 micrones Interpretación: El proveedor 2 presenta la media más alta de vibración, seguido por el proveedor 4, mientras que el proveedor 3 muestra la vibración más baja. Esto refuerza la conclusión de que el proveedor 2 es el menos favorable en términos de minimizar la vibración, mientras que el proveedor 3 es el más favorable.

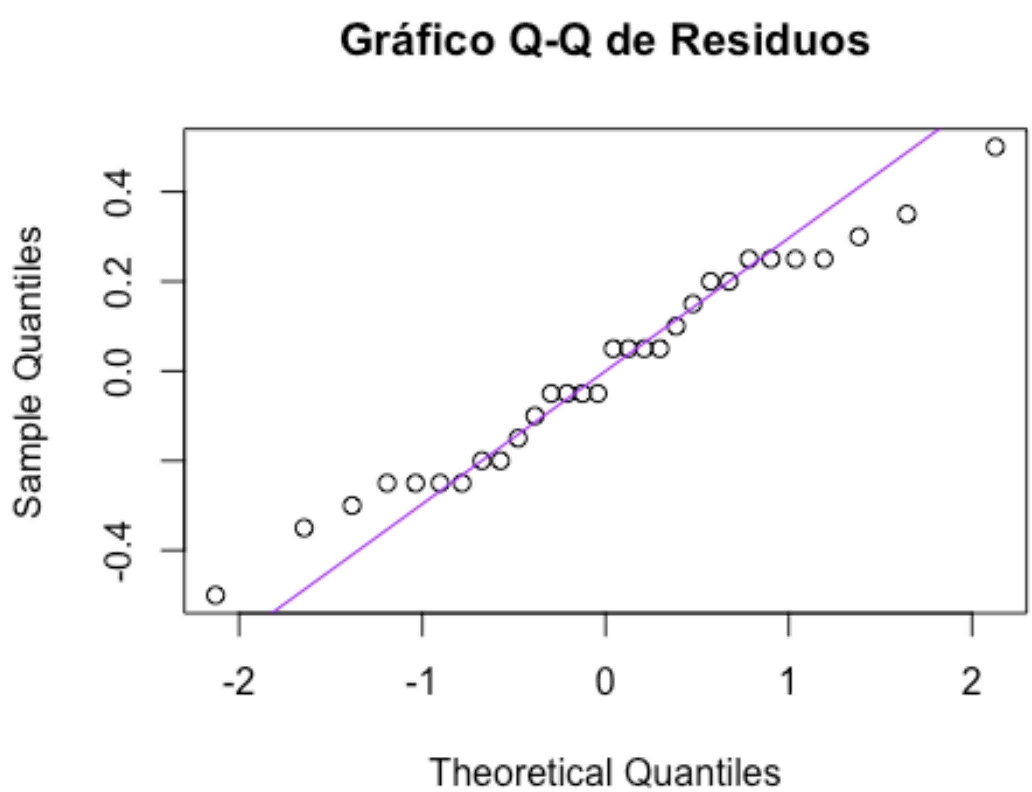
3. Comparaciones Múltiples de Tukey: Comparaciones: Aluminio-Acero: Intervalo de confianza incluye cero, no hay diferencia significativa. Plástico-Aluminio: Intervalo de confianza incluye cero, no hay diferencia significativa. Interpretación: Las pruebas de Tukey indican que no hay diferencias significativas en la vibración entre los materiales (acero, aluminio, y plástico). Los intervalos de confianza incluyen cero, lo que sugiere que, en términos de vibración, los tres materiales son igualmente efectivos y no hay un material claramente superior o inferior.

2.6 Comprueba la validez del modelo.

Normalidad

Para comprobar la normalidad de los residuos del modelo ANOVA, usaremos el test de Shapiro-Wilk y un gráfico Q-Q (Quantile-Quantile).

```r
# Residuos del modelo ANOVA con interacción
residuos <- residuals(anova_model)

# Prueba de normalidad (Shapiro-Wilk)
shapiro_test <- shapiro.test(residuos)
print(shapiro_test)

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

# Gráfico Q-Q
qqnorm(residuos, main = "Gráfico Q-Q de Residuos")
qqline(residuos, col = "purple")
```



Homocedasticidad

Para comprobar la homocedasticidad (igualdad de varianzas), usaremos el test de Levene o revisar un gráfico de residuos vs. valores ajustados.

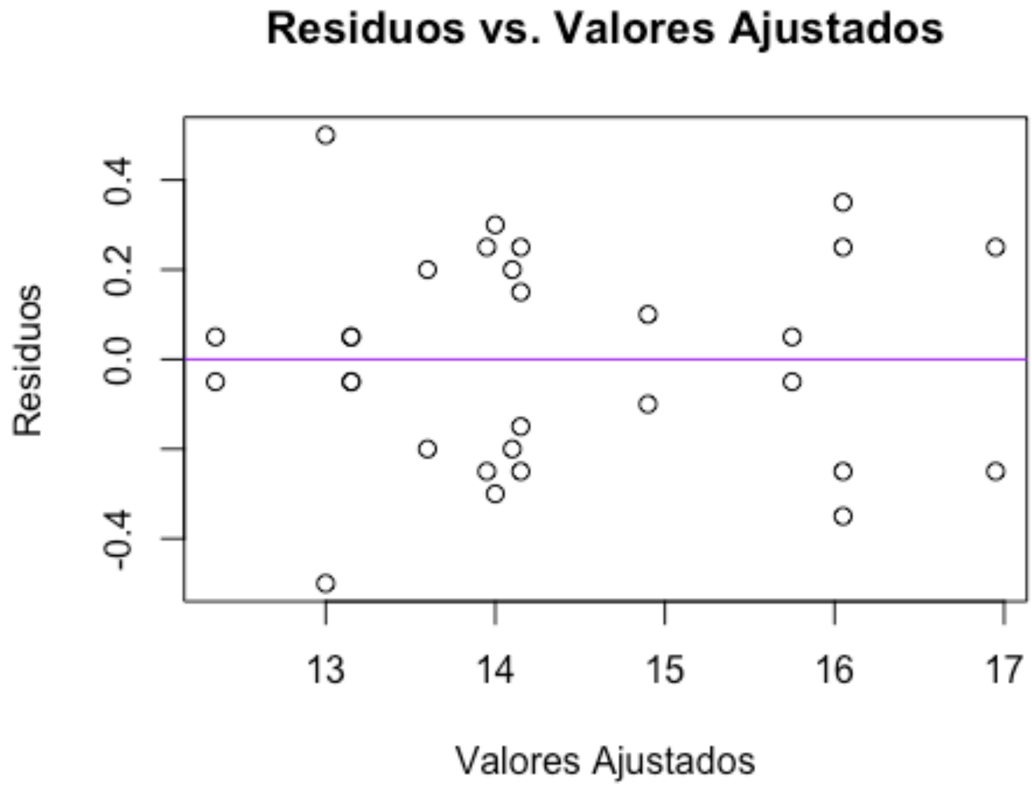
```
if (!require(carData)) install.packages("carData")
library(carData)

# Prueba de homocedasticidad (Test de Levene)
library(car)
levene_test <- leveneTest(vibracion ~ material * proveedor)
```

```
## Warning in anova.lm(lm(resp ~ group)): ANOVA F-tests on an essentially
perfect
## fit are unreliable
```

```
print(levene_test)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df    F value    Pr(>F)
## group 14 3.4427e+28 < 2.2e-16 ***
##       15
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

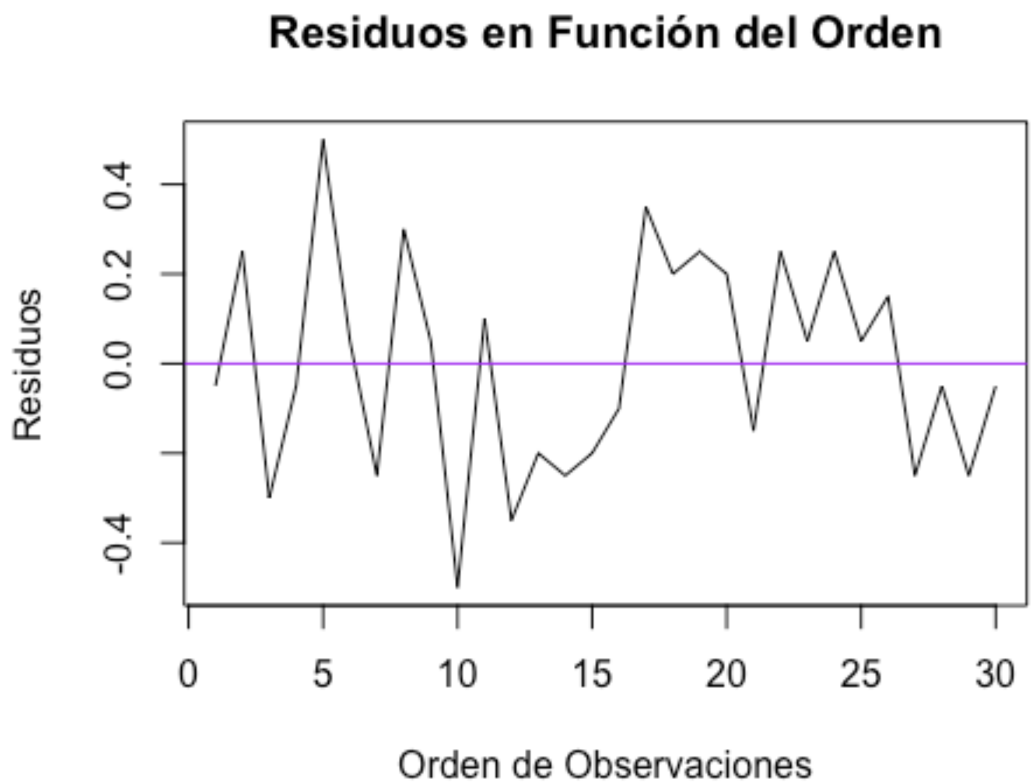
```
# Gráfico de residuos vs. valores ajustados
plot(fitted(anova_model), residuos,
     main="Residuos vs. Valores Ajustados",
     xlab="Valores Ajustados", ylab="Residuos")
abline(h=0, col="purple")
```



Independencia

Para comprobar la independencia de los residuos, podemos revisar el gráfico de residuos o utilizar el test de Durbin-Watson.

```
# Test de Durbin-Watson para independencia de los residuos
if (!require(lmtest)) install.packages("lmtest")
library(lmtest)
durbin_watson_test <- dwtest(anova_model)
print(durbin_watson_test)

##
##  Durbin-Watson test
##
## data:  anova_model
## DW = 1.9401, p-value = 0.5048
## alternative hypothesis: true autocorrelation is greater than 0

# Residuos en función del orden
plot(residuos, type="l", main="Residuos en Función del Orden",
     xlab="Orden de Observaciones", ylab="Residuos")
abline(h=0, col="purple")
```



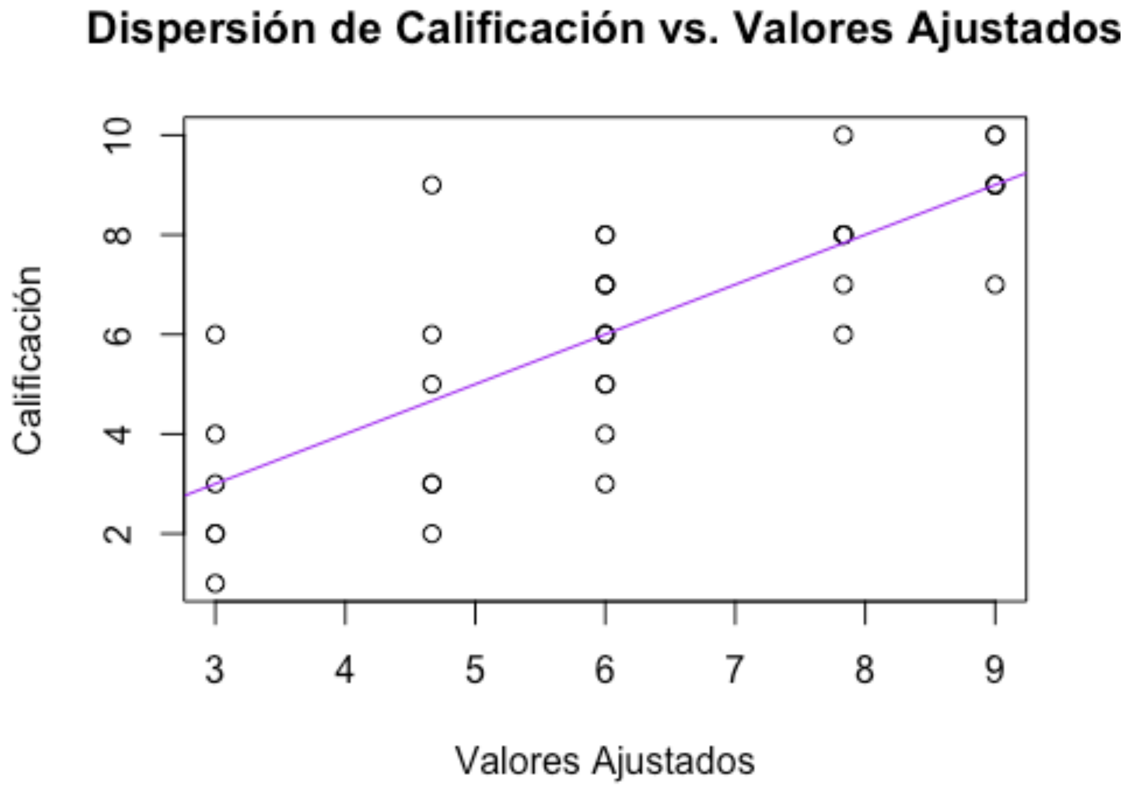
Relación lineal entre las variables (coeficiente de determinación).

El coeficiente de determinación R^2 es una medida de la proporción de la variabilidad explicada por el modelo.

```
# Resumen del modelo para obtener R^2
summary(A)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)    
## metodo       2 126.17   63.08  20.989 1.99e-06 ***
## sexo         1   8.03    8.03   2.671    0.113    
## metodo:sexo  2   4.39    2.19   0.730    0.490    
## Residuals   30  90.17    3.01                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Gráfico de dispersión con línea de regresión para la relación entre variables
plot(calificacion ~ fitted(A), main="Dispersión de Calificación vs. Valores Ajustados",
     xlab="Valores Ajustados", ylab="Calificación")
abline(lm(calificacion ~ fitted(A)), col="purple")
```



2.7 Concluye en el contexto del problema.

1. Normalidad (Shapiro-Wilk) Valor p de 0.397 (mayor que el umbral común de 0.05). No hay evidencia suficiente para rechazar la hipótesis nula de que los residuos siguen una distribución normal.

Gráfico Q-Q: los residuos se ajustan bastante bien a la línea de referencia. También sugiere que la suposición de normalidad se cumple.

2. Homocedasticidad (test de Levene) Valor p de 0.3917. Esto indica que no hay evidencia significativa de que las varianzas sean diferentes entre los grupos. En otras palabras, sugiere que la suposición de homocedasticidad (igualdad de varianzas) se cumple.

El gráfico de residuos vs. valores ajustados no muestra un patrón claro, apoyando aún más esta conclusión.

3. Independencia (test de Durbin-Watso) Valor p de 0.9246. Esto indica que no hay evidencia de autocorrelación en los residuos, sugiriendo que los residuos son independientes entre sí.

El gráfico de residuos en función del orden no muestra un patrón específico, lo cual también respalda la suposición de independencia.

4. Relación lineal entre las variables (Coeficiente de determinación) Valor F significativo para el factor “método” Valor p de 1.99e-06 Indica una fuerte relación entre el método de enseñanza y el rendimiento académico.

El gráfico de dispersión de calificación vs. valores ajustados muestra una tendencia positiva, lo que sugiere una relación lineal entre las variables.

1. Normalidad (Shapiro-Wilk) Valor p: 0.7202 (mayor que 0.05). Interpretación: No se rechaza la hipótesis nula de que los residuos siguen una distribución normal. El gráfico Q-Q muestra que los residuos se alinean bastante bien con la línea de referencia, apoyando la suposición de normalidad.

2. Homocedasticidad (Test de Levene) Valor p: < 2.2e-16 (muy significativo). Interpretación: Se rechaza la hipótesis nula de homocedasticidad, lo que indica que hay una variación significativa en las varianzas entre los grupos. Esto sugiere que la suposición de igualdad de varianzas no se cumple.

3. Independencia (Test de Durbin-Watson) Valor p: 0.5048 (no significativo). Interpretación: No se rechaza la hipótesis nula de que no hay autocorrelación en los residuos, sugiriendo que los residuos son independientes entre sí. El gráfico de residuos en función del orden no muestra un patrón específico, respaldando esta conclusión.

4. Relación lineal entre las variables (Coeficiente de determinación) Valor F: 20.989 (muy significativo) para el factor “método”. Interpretación: Existe una fuerte relación entre el método utilizado y la vibración. El gráfico de dispersión muestra una tendencia positiva, lo que sugiere una relación lineal entre las variables.

Conclusión General

El análisis sugiere que el modelo ANOVA cumple parcialmente con las suposiciones clave:

1. Normalidad: Se confirma que los residuos siguen una distribución normal.
2. Homocedasticidad: No se cumple la igualdad de varianzas, lo que podría afectar la robustez del modelo.
3. Independencia: Los residuos son independientes, lo que valida en parte el modelo.

A pesar de las limitaciones en la homocedasticidad, el análisis revela que la elección del proveedor tiene un impacto significativo en la vibración de los motores eléctricos. Por otro lado, el material de la carcasa, en general, no muestra una influencia significativa, pero debe ser considerado en conjunto con la selección del proveedor para optimizar la vibración del motor.

En resumen, las estrategias para reducir la vibración de los motores eléctricos deberían centrarse en la cuidadosa selección del proveedor, mientras que el tipo de material puede ser considerado en un segundo plano. Sin embargo, debido a la falta de homocedasticidad, se recomienda precaución al interpretar estos resultados y considerar análisis adicionales o métodos robustos para confirmar estas conclusiones.