

9Es_A01571214_Lautaro_Coteja

A01571214 - Lautaro Coteja

2024-08-29

R Markdown

Actividad 9: ANOVA

Problema 1

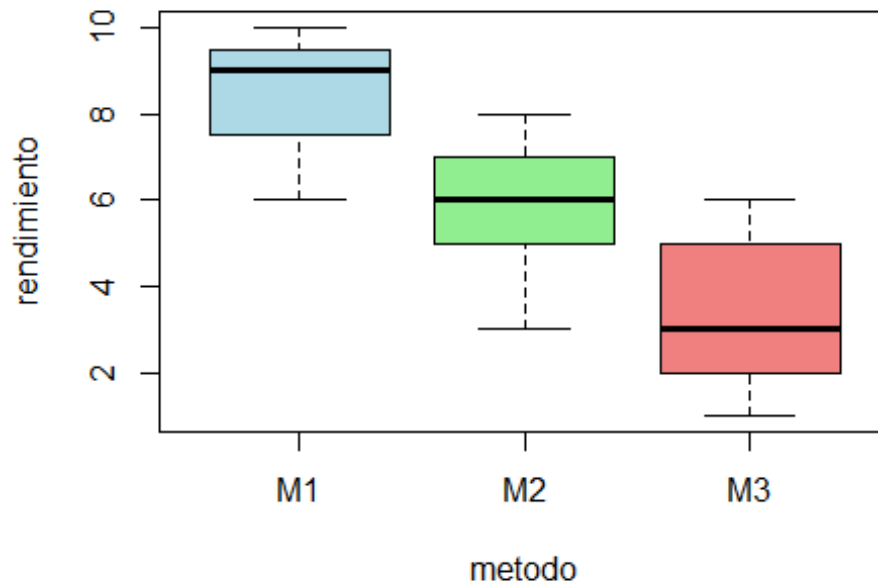
```
rendimiento = c(10, 7, 9, 9, 9, 10, 5, 7, 6, 6, 8, 4, 2, 6, 3, 5, 5, 3, 9, 7,
8, 8, 10, 6, 8, 3, 5, 6, 7, 7, 2, 6, 2, 1, 4, 3)
metodo = factor(c(rep("M1", 6), rep("M2", 6), rep("M3", 6), rep("M1", 6),
rep("M2", 6), rep("M3", 6)))
sexo = factor(c(rep("h", 18), rep("m", 18)))

media_por_metodo = tapply(rendimiento, metodo, mean)
media_por_metodo

##  M1  M2  M3
## 8.5 6.0 3.5

boxplot(rendimiento ~ metodo, col = c("lightblue", "lightgreen",
"lightcoral"), main = "Boxplot de Rendimiento por Metodo de Enseñanza")
```

Boxplot de Rendimiento por Metodo de Enseñanz



Interpretacion de Resultados

Es posible observar como Los diferentes metodos de enseñanza afectan el rendimiento de Los estudiantes. Estadísticamente, si Las medias de Los diferentes metodos son significativamente diferentes, esto indicaria que La metodologia de enseñanza tiene un impacto en el rendimiento de Los estudiantes. En el contexto del problema, este analisis podria ayudar a identificar cual de Los metodos es mas efectivo para mejorar el rendimiento en ciencias naturales.

Establecimiento de Hipotesis

Primera hipotesis:

H_0 : No hay efecto del metodo de enseñanza en el rendimiento.

H_1 : Algun metodo de enseñanza tiene un efecto significativo.

Segunda hipotesis:

H_0 : No hay efecto del genero en el rendimiento.

H_1 : Algun género tiene un efecto significativo.

Tercera hipotesis:

H_0 : No hay interaccion entre el metodo de enseñanza y el genero.

H_1 : Existe interaccion entre el metodo de enseñanza y el genero.

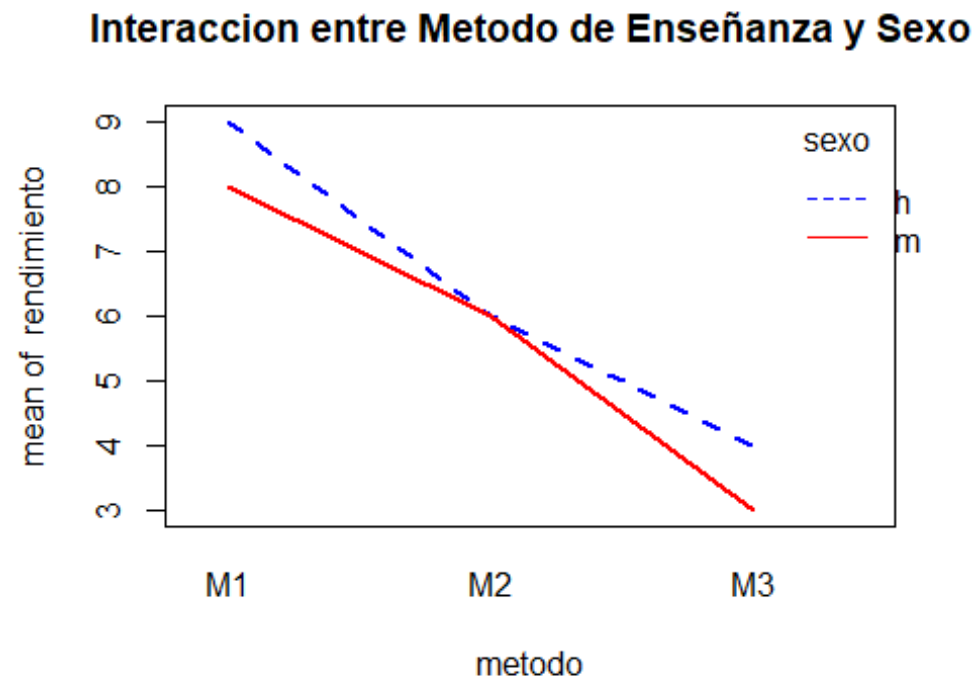
ANOVA con Interaccion

```
anova_interaccion = aov(rendimiento ~ metodo * sexo)
```

```
summary(anova_interaccion)
```

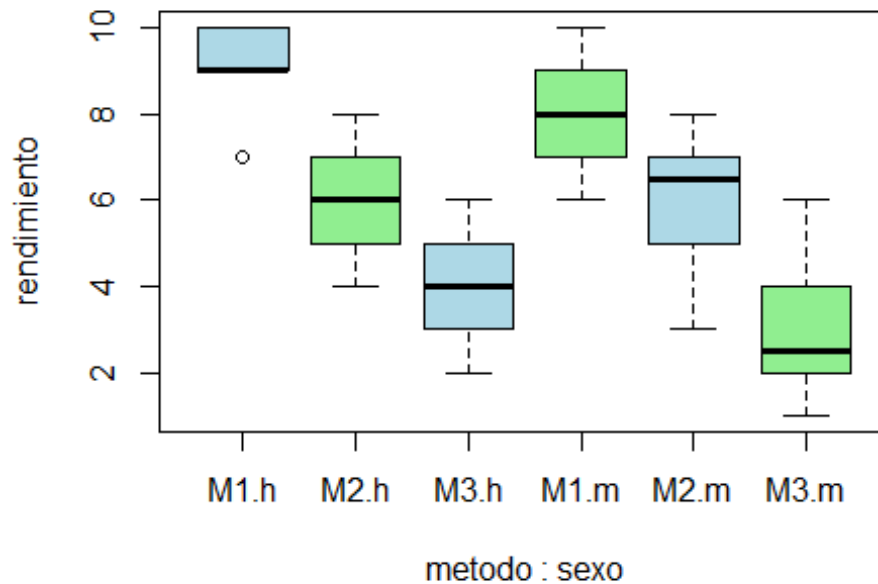
```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## metodo      2    150   75.00   32.143 3.47e-08 ***
## sexo        1      4    4.00    1.714   0.200
## metodo:sexo  2      2    1.00    0.429   0.655
## Residuals   30     70    2.33
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'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

```
interaction.plot(metodo, sexo, rendimiento, col = c("blue", "red"), lwd = 2,
main = "Interaccion entre Metodo de Enseñanza y Sexo")
```



```
boxplot(rendimiento ~ metodo * sexo, col = c("lightblue", "lightgreen"), main
= "Boxplot de Rendimiento por Metodo y Sexo")
```

Boxplot de Rendimiento por Metodo y Sexo



Interpretacion de Resultados

Se puede identificar si existe una interaccion significativa entre Los factores (genero y metodo de enseñanza). Si La interacción es significativa, significa que el efecto de un método de enseñanza en el rendimiento depende del género del estudiante. En el contexto del problema, esto podria indicar que diferentes metodologias de enseñanza son mas efectivas para diferentes generos, Lo que podria tener implicaciones importantes para La personalizacion de La enseñanza.

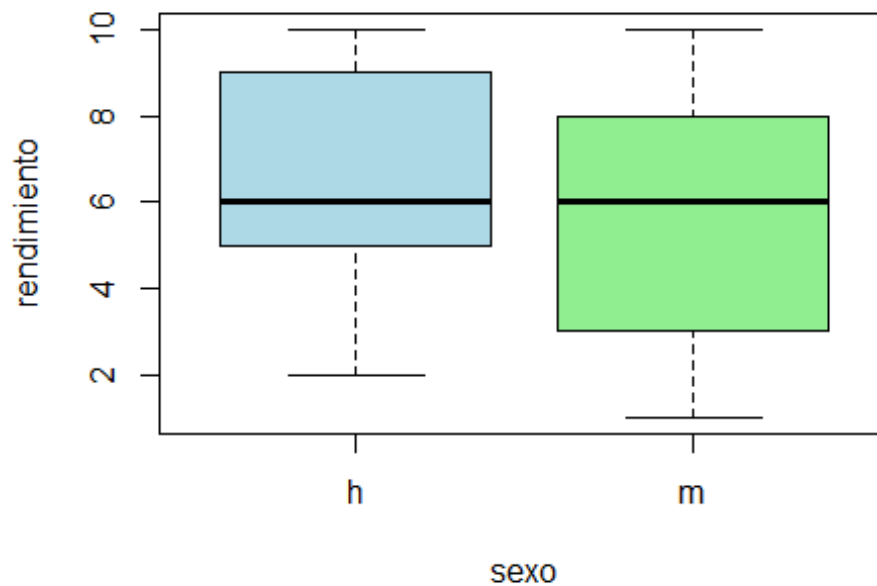
ANOVA sin Interaccion

```
anova_sin_interaccion = aov(rendimiento ~ metodo + sexo)
summary(anova_sin_interaccion)
```

```
##           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, col = c("lightblue", "lightgreen"), main =
"Boxplot de Rendimiento por Sexo")
```

Boxplot de Rendimiento por Sexo



Media por Sexo

```
media_por_sexo = tapply(rendimiento, sexo, mean)
media_por_sexo
```

```
##           h           m
## 6.333333 5.666667
```

Intervalos de confianza por sexo

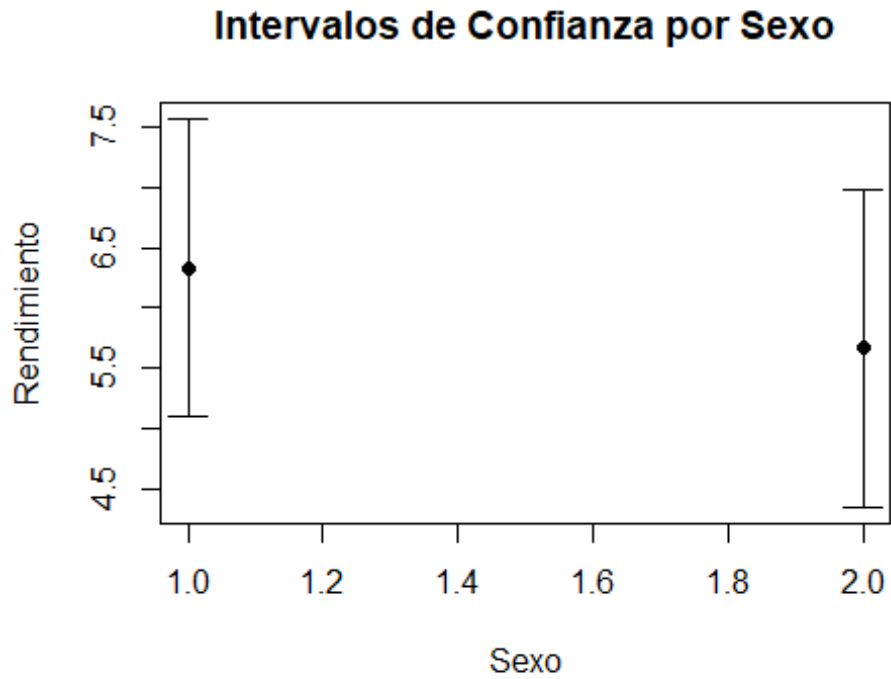
```
IC = tapply(rendimiento, sexo, function(x) t.test(x)$conf.int)
IC
```

```
## $h
## [1] 5.103347 7.563320
## attr("conf.level")
## [1] 0.95
##
## $m
## [1] 4.356505 6.976828
## attr("conf.level")
## [1] 0.95
```

```
lower_bounds = sapply(IC, `[`, 1)
upper_bounds = sapply(IC, `[`, 2)
means = sapply(IC, mean)
```

```
plot(1:length(IC), means, ylim = range(c(lower_bounds, upper_bounds)), pch =
19, xlab = "Sexo", ylab = "Rendimiento", main = "Intervalos de Confianza por
```

```
Sexo")
arrows(1:length(IC), lower_bounds, 1:length(IC), upper_bounds, angle = 90,
code = 3, length = 0.1)
```



```
# ANOVA para un solo factor (Metodo de Enseñanza)
anova_un_solo_factor = aov(rendimiento ~ metodo)
summary(anova_un_solo_factor)

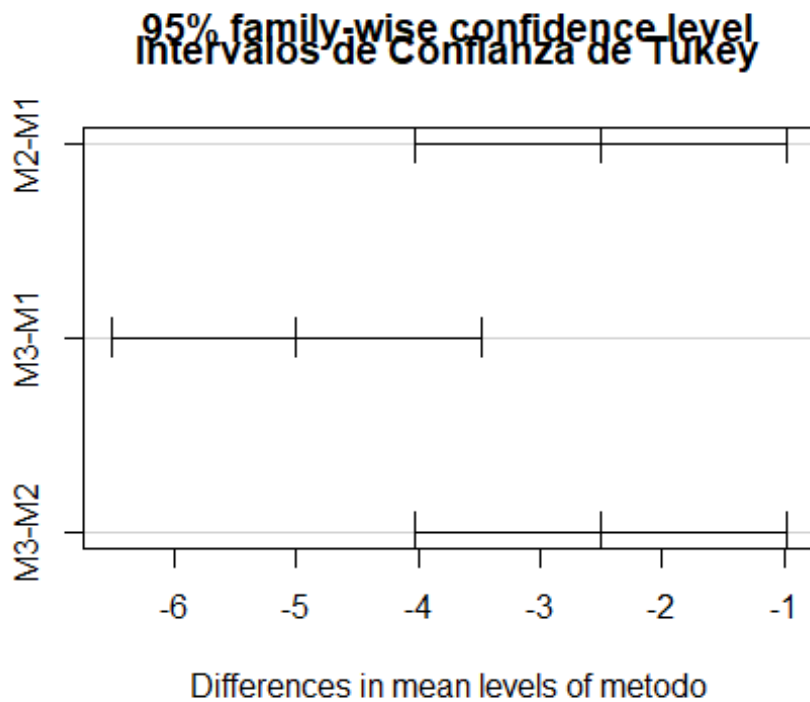
##              Df Sum Sq Mean Sq F value    Pr(>F)
## metodo         2    150    75.0   32.57 1.55e-08 ***
## Residuals     33     76     2.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Prueba de Tukey
tukey_test = TukeyHSD(anova_un_solo_factor)
tukey_test

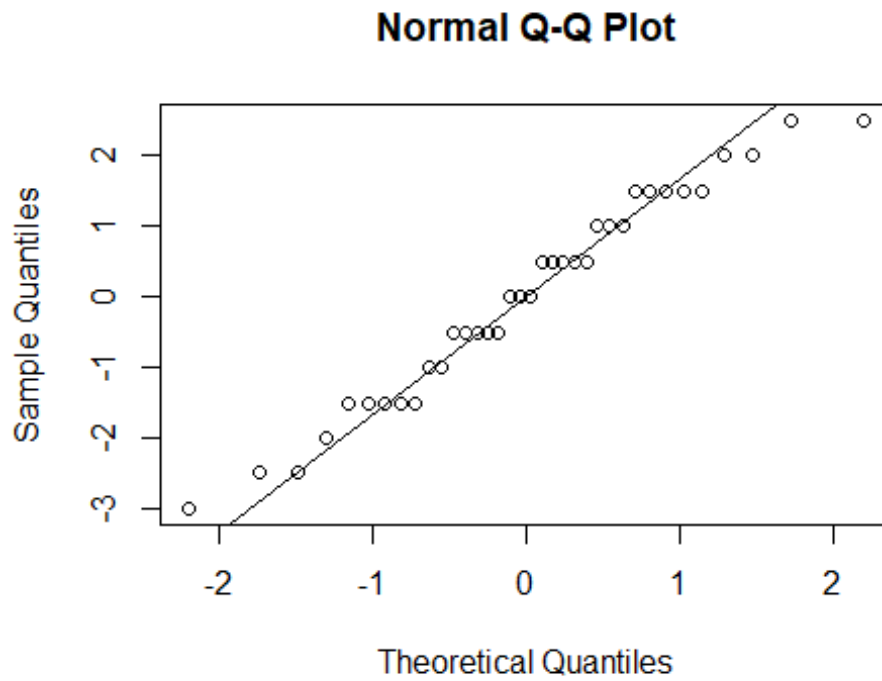
##      Tukey multiple comparisons of means
##      95% family-wise confidence level
##
## Fit: aov(formula = rendimiento ~ metodo)
##
## $metodo
##      diff      lwr      upr    p adj
## M2-M1 -2.5 -4.020241 -0.9797592 0.0008674
```

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

plot(tukey_test)
title(main = "Intervalos de Confianza de Tukey")
```

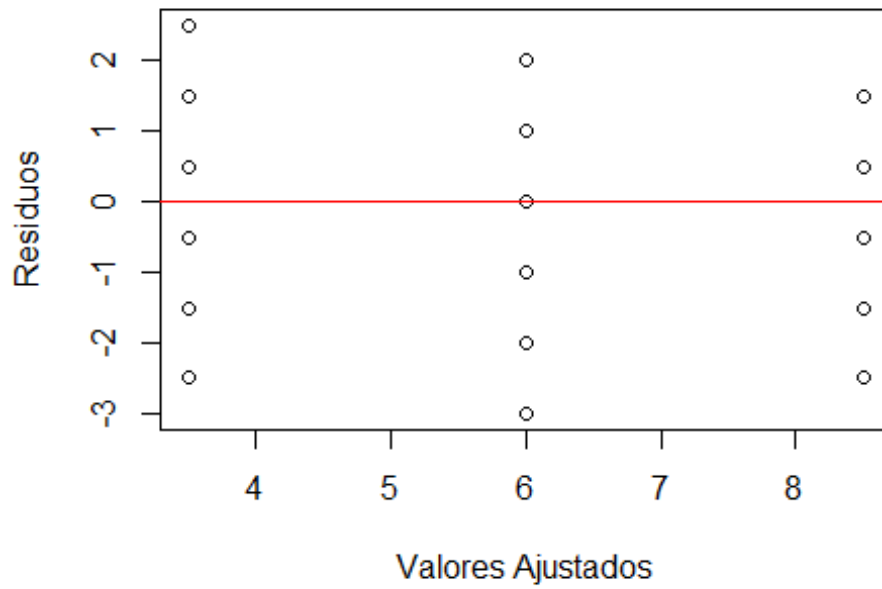


```
# Normalidad de Residuos
qqnorm(anova_un_solo_factor$residuals)
qqline(anova_un_solo_factor$residuals)
```



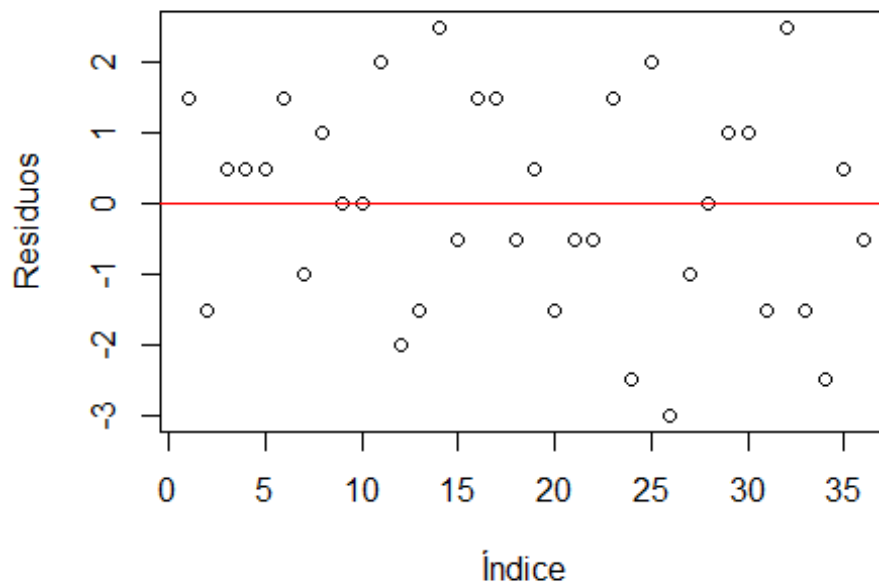
```
# Homocedasticidad
plot(anova_un_solo_factor$fitted.values, anova_un_solo_factor$residuals, main = "Homocedasticidad", xlab = "Valores Ajustados", ylab = "Residuos")
abline(h = 0, col = "red")
```


Homocedasticidad



```
# Independencia  
plot(anova_un_solo_factor$residuals, main = "Independencia de Residuos", ylab = "Residuos", xlab = "Índice")  
abline(h = 0, col = "red")
```

Independencia de Residuos



Conclusion

Se puede concluir que el metodo de enseñanza es un factor determinante en el rendimiento de los estudiantes, mientras que el genero puede no ser tan influyente. Además, la ausencia de una interacción significativa sugiere que el mejor metodo de enseñanza es efectivo independientemente del genero. Estos resultados proporcionan evidencia estadística para recomendar el uso de un metodo particular en la enseñanza de ciencias naturales para maximizar el rendimiento de todos los estudiantes.

Problema 2

```
vibracion = c(13.1, 13.2, 16.3, 15.8, 13.7, 14.3, 15.7, 15.8, 13.5, 12.5,
15.0, 14.8, 15.7, 16.4, 13.9, 14.3, 13.7, 14.2, 13.4, 13.8, 14.0, 14.3, 17.2,
16.7, 12.4, 12.3, 14.4, 13.9, 13.2, 13.1)
```

```
material = factor(rep(c("Acero", "Aluminio", "Plastico"), each = 10))
proveedor = factor(rep(1:5, times = 6))
```

Calcular la media de vibracion por material y por proveedor

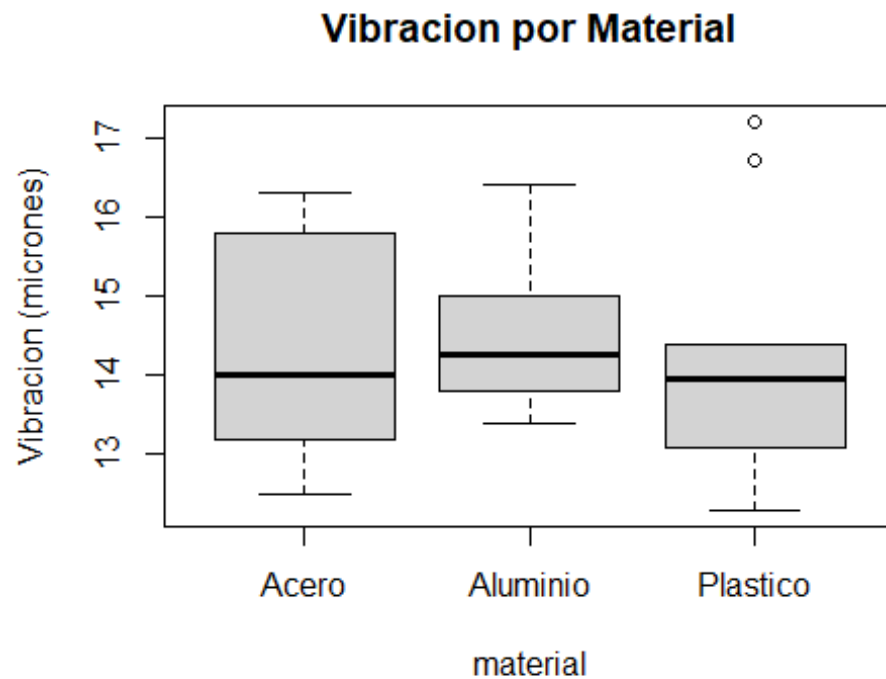
```
tapply(vibracion, material, mean)
```

```
##   Acero Aluminio Plastico
##  14.39   14.52   14.15
```

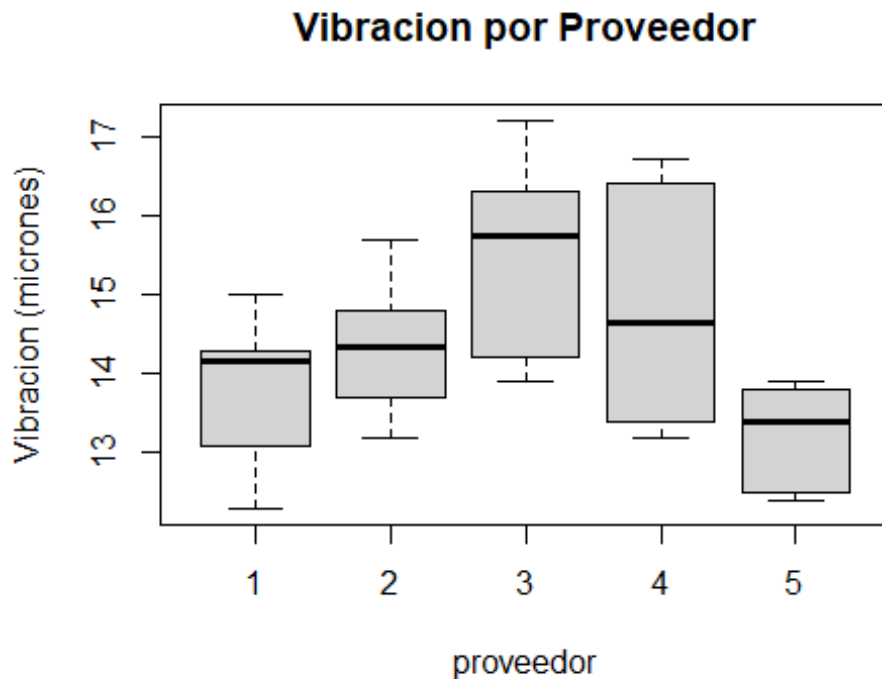
```
tapply(vibracion, proveedor, mean)
```

```
##          1          2          3          4          5
## 13.83333 14.35000 15.51667 14.83333 13.23333

# Boxplot por material
boxplot(vibracion ~ material, main = "Vibracion por Material", ylab =
"Vibracion (micrones)")
```



```
# Boxplot por proveedor
boxplot(vibracion ~ proveedor, main = "Vibracion por Proveedor", ylab =
"Vibracion (micrones)")
```



Interpretacion de Resultados

De acuerdo a la media del material, se observa que el material influye en la cantidad promedio de vibracion, y lo mismo pasa con la media del proveedor, donde se nota que la media de vibracion para cada proveedor puede revelar si algunos proveedores suministran cojinetes que resultan en mayor o menor vibracion.

Hipotesis

Hipotesis 1: H_0 No hay diferencia significativa en la vibracion promedio entre los materiales de la carcasa. H_1 Si existe alguna diferencia significativa

Hipotesis 2: H_0 No hay diferencia significativa en la vibracion promedio entre los proveedores de cojinetes. H_1 Si existe algun proveedor que produce una vibracion significativamente diferente

Hipotesis 3: H_0 No hay interaccion significativa entre el material de la carcasa y el proveedor en terminos de vibracion. H_1 Si existe una interaccion significativa entre los dos, lo que significa que afectan a la vibracion.

ANOVA

Con Interaccion

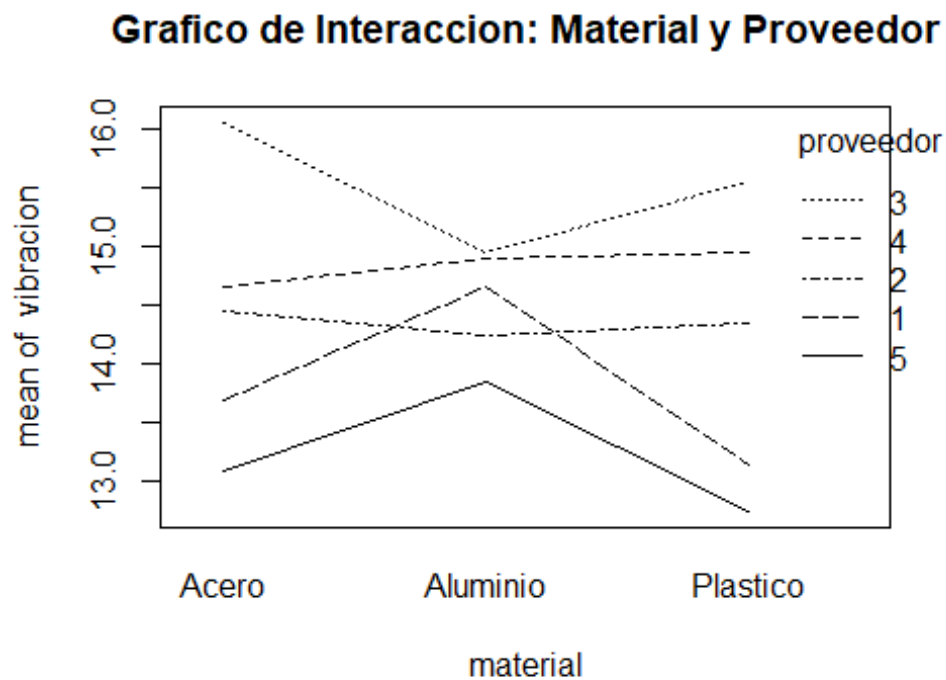
```
anova_inter = aov(vibracion ~ material * proveedor)
```

```
summary(anova_inter)
```

##		Df	Sum Sq	Mean Sq	F value	Pr(>F)
##	material	2	0.705	0.352	0.195	0.8248
##	proveedor	4	18.651	4.663	2.583	0.0797 .
##	material:proveedor	8	4.219	0.527	0.292	0.9579

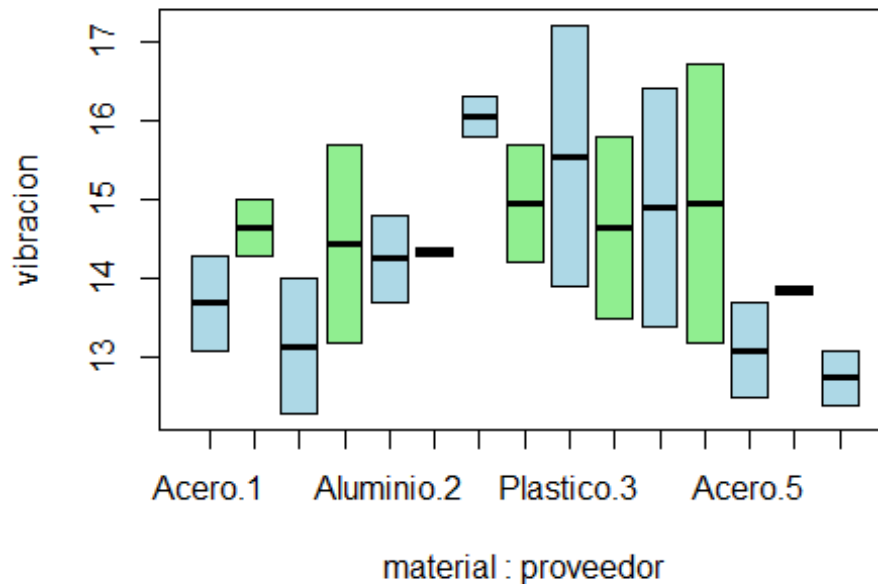
```
## Residuals      15 27.080    1.805
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Grafica
interaction.plot(material, proveedor, vibracion, main = "Grafico de
Interaccion: Material y Proveedor")
```



```
# Boxplot para visualizar la interaccion entre material y proveedor
boxplot(vibracion ~ material * proveedor, col = c("lightblue", "lightgreen"),
main = "Boxplot de Vibracion por Material y Proveedor")
```

Boxplot de Vibracion por Material y Proveedor



Sin Interaccion

```
anova_sin_inter = aov(vibracion ~ material + proveedor)
summary(anova_sin_inter)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## material    2  0.705   0.352   0.259 0.7741
## proveedor    4 18.651   4.663   3.427 0.0245 *
## Residuals   23 31.299   1.361
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

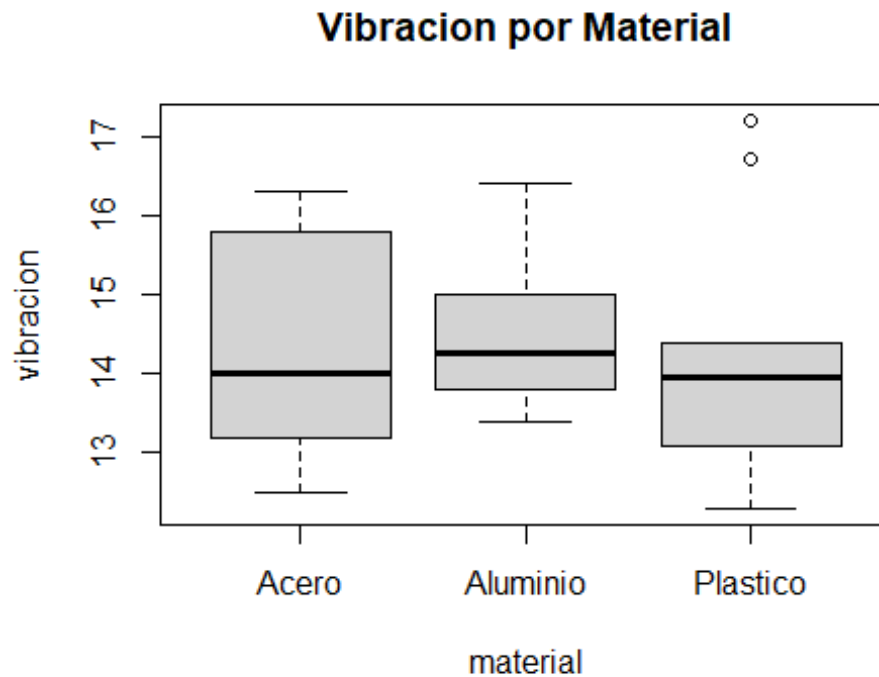
Intervalos de Confianza y Boxplots

Media y Boxplot por sexo y metodo

```
tapply(vibracion, material, mean)
```

```
##      Acero Aluminio Plastico
##      14.39    14.52    14.15
```

```
boxplot(vibracion ~ material, main = "Vibracion por Material")
```

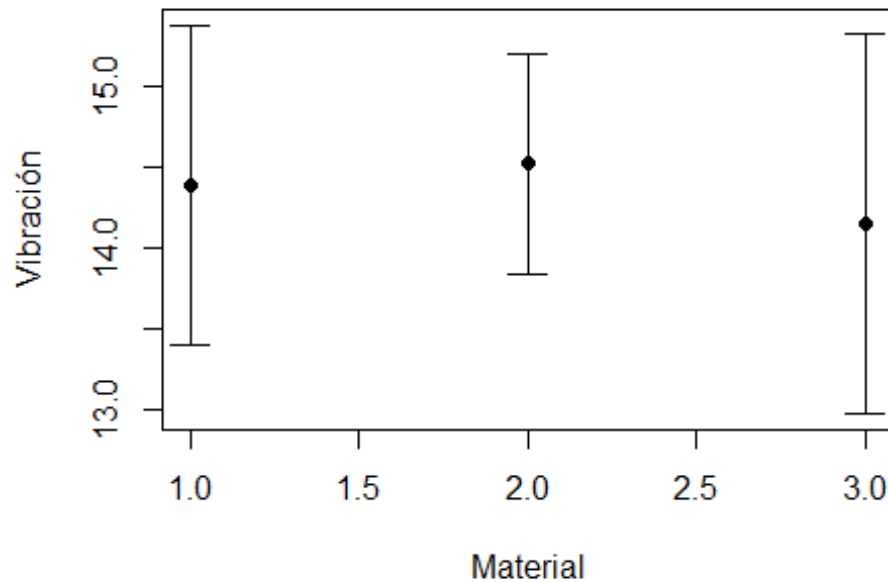


```
# Intervalos de confianza por sexo
ci_material = tapply(vibracion, material, function(x) t.test(x)$conf.int)
ci_proveedor = tapply(vibracion, proveedor, function(x) t.test(x)$conf.int)

lower_bounds_material = sapply(ci_material, `[`, 1)
upper_bounds_material = sapply(ci_material, `[`, 2)
means_material = sapply(ci_material, mean)

# Graficos de los intervalos de confianza
plot(1:length(ci_material), means_material, ylim =
range(c(lower_bounds_material, upper_bounds_material)), pch = 19, xlab =
"Material", ylab = "Vibración", main = "Intervalos de Confianza por
Material")
arrows(1:length(ci_material), lower_bounds_material, 1:length(ci_material),
upper_bounds_material, angle = 90, code = 3, length = 0.1)
```

Intervalos de Confianza por Material



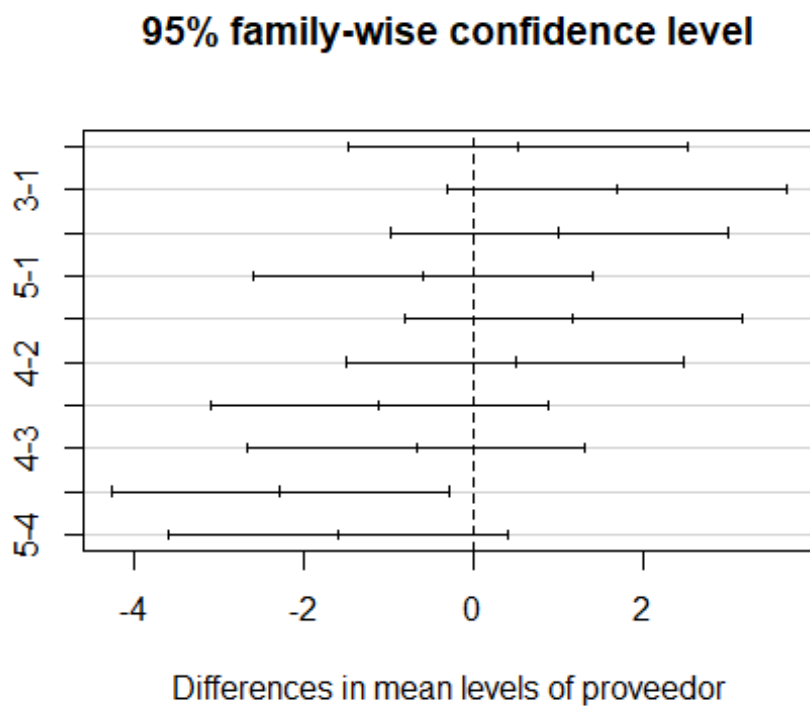
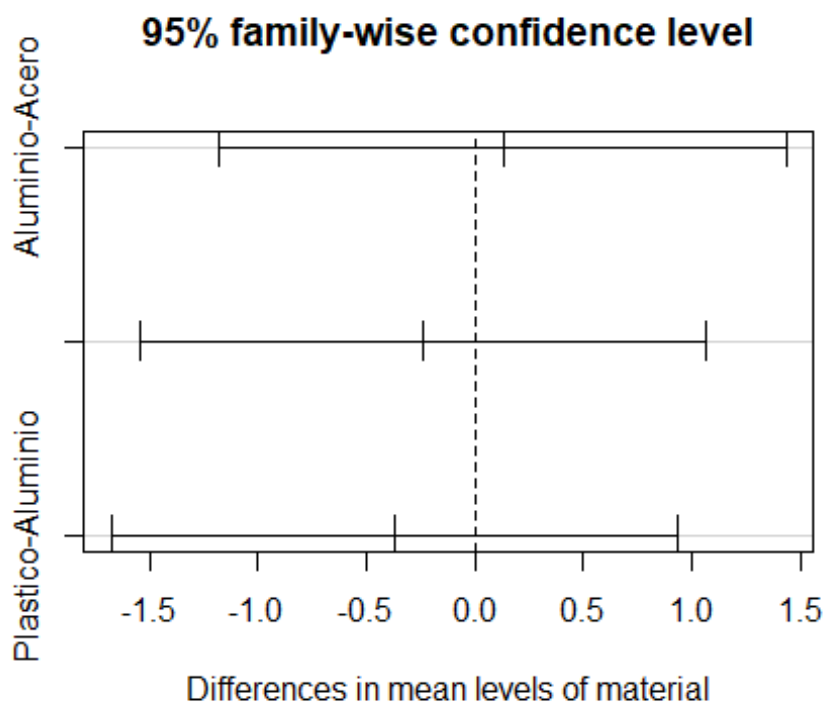
Interpretacion de Resultados

Podemos notar que el material y el proveedor si afectan a La vibracion. Y la interaccion enseña como combinaciones especificas de material y proveedor afectan a la vibracion.

Tukey

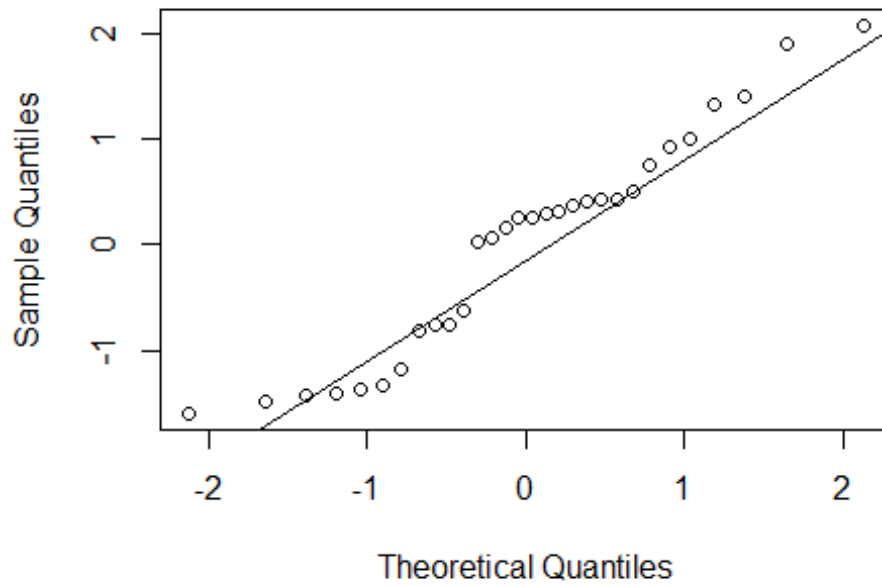
```
tukey <- TukeyHSD(anova_sin_inter)
```

```
plot(tukey)
```

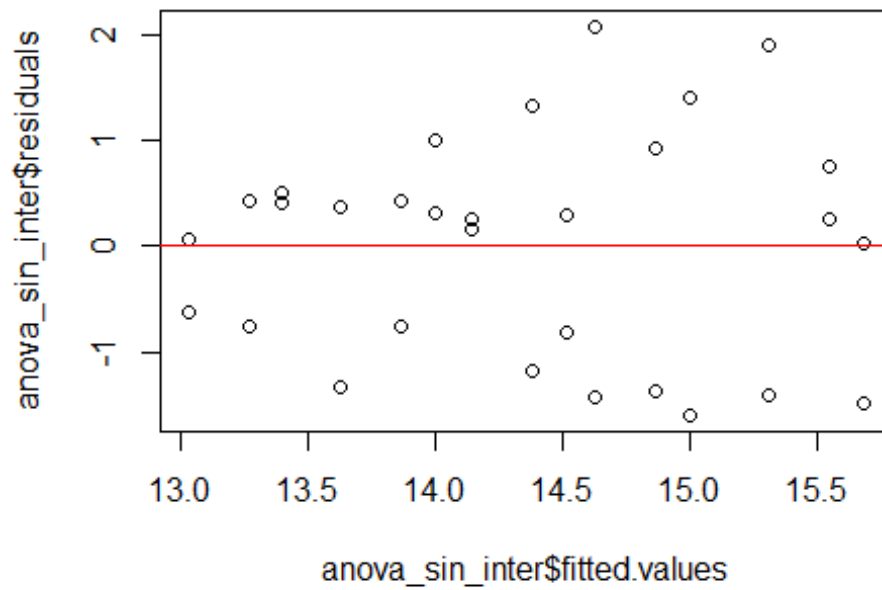



```
# Validez
# Normalidad
qqnorm(anova_sin_inter$residuals)
qqline(anova_sin_inter$residuals)
```

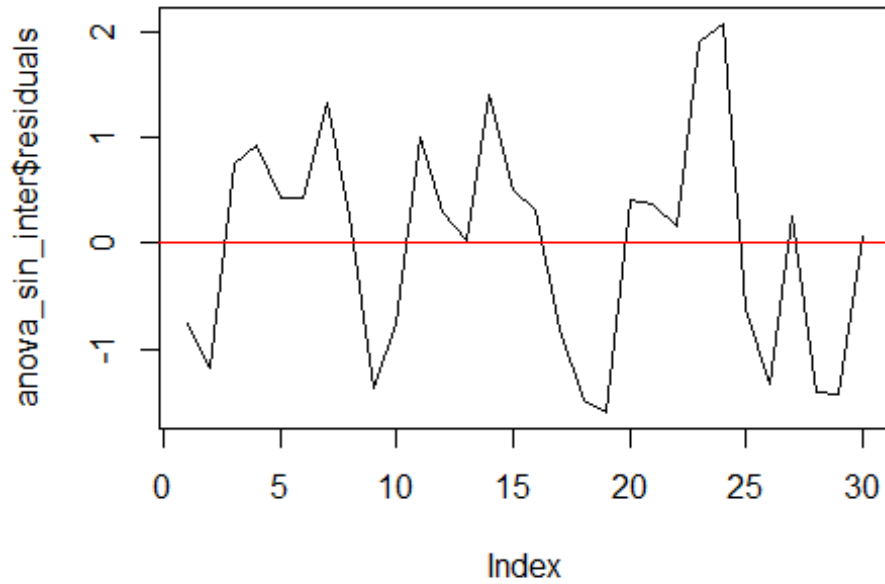
Normal Q-Q Plot



```
# Homocedasticidad  
plot(anova_sin_inter$fitted.values, anova_sin_inter$residuals)  
abline(h=0, col="red")
```



```
# Independencia
plot(anova_sin_inter$residuals, type = "l")
abline(h=0, col="red")
```



```
# Relacion lineal entre variables (Coeficiente de determinacion)
summary(lm(vibracion ~ material + proveedor))$r.squared
```

```
## [1] 0.3821168
```

Interpretacion de Resultados

La normalidad da validez a los resultados del ANOVA, la Homocedasticidad dice que la varianza de los errores es consistente a lo largo de los niveles de los factores. En la dependencia, si los residuos no muestran una tendencia clara cuando se grafican en el orden de observación, significa que son independientes, lo que valida los resultados. Y el coeficiente por su alto valor indica que el modelo funciona correctamente.

Conclusion

Basado en todos los análisis realizados, se puede concluir que los materiales y los proveedores, efectivamente, sí influyen significativamente en la vibración de los motores. Las gráficas de interacción y los boxplots proporcionan una visualización clara de cómo estos factores interactúan y afectan la vibración. En base a esto, ayuda a la selección de materiales y proveedores que minimicen la vibración, mejorando la calidad y durabilidad de los motores.