Inferencia Estadística

Solución: Análisis de Varianza

DSLab

junio, 2024

Ejercicio 1

Pregunta: Se realizó un experimento para investigar si hay diferencias significativas en el tiempo de reacción ante un estímulo visual entre tres grupos de participantes que fueron sometidos a diferentes niveles de iluminación: normal, baja y alta. Se registraron los tiempos de reacción (en milisegundos) para cada participante en cada grupo. Los datos obtenidos se muestran a continuación:

Grupo Normal: 40, 45, 38, 42, 39
Grupo Baja: 48, 50, 55, 52, 51
Grupo Alta: 35, 37, 32, 40, 36

Usando un nivel de significancia de 0.05, ¿hay diferencias significativas en el tiempo de reacción entre los tres grupos?

Solución: Para determinar si hay diferencias significativas en el tiempo de reacción entre los tres grupos, podemos utilizar un análisis de varianza (ANOVA) de un factor. Aquí están los pasos para realizar el ANOVA:

Paso 1: Calcular las medias de cada grupo.

Grupo Normal: $\bar{X}_1 = \frac{40+45+38+42+39}{5} = 40.8$
Grupo Baja: $\bar{X}_2 = \frac{48+50+55+52+51}{5} = 51.2$
Grupo Alta: $\bar{X}_3 = \frac{35+37+32+40+36}{5} = 36$

Paso 2: Calcular la media global (promedio de todas las observaciones).

$$\bar{X} = \frac{40.8 + 51.2 + 36}{3} \approx 42.67$$

Paso 3: Calcular las sumas de cuadrados (SS).

$$SS_{\text{Total}} = \sum (X_{ij} - \bar{X})^2 = \sum (X_{ij} - 42.67)^2 \approx 695.33$$

Paso 4: Calcular las sumas de cuadrados entre grupos (SSB).

$$SSB = \sum (\bar{X}_j - \bar{X})^2$$

Para cada grupo:

$$SSB_{\text{Grupo Normal}} = 5 \cdot (40.8 - 42,7)^2 \approx 17.42$$
$$SSB_{\text{Grupo Baja}} = 5 \cdot (51.2 - 42.7)^2 \approx 364.09$$
$$SSB_{\text{Grupo Alta}} = 5 \cdot (36 - 42.7)^2 \approx 222.22$$

De modo que:

$$SSB = 17.42 + 364.09 + 222.22 \approx 603.73$$

Paso 5: Calcular las sumas de cuadrados dentro de los grupos (SSW).

$$SSW = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$

Para cada grupo:

$$SS_{\text{Grupo Normal}} = (40 - 40.8)^2 + (45 - 40.8)^2 + (38 - 40.8)^2 + (42 - 40.8)^2 + (39 - 40.8)^2 = 30.8$$
$$SS_{\text{Grupo Baja}} = (48 - 51.2)^2 + (50 - 51.2)^2 + (55 - 51.2)^2 + (52 - 51.2)^2 + (51 - 51.2)^2 = 26.8$$
$$SS_{\text{Grupo Alta}} = (35 - 36)^2 + (37 - 36)^2 + (32 - 36)^2 + (40 - 36)^2 + (36 - 36)^2 = 34$$

De donde:

$$SSW = 30.8 + 26.8 + 34 = 91.6$$

Paso 6: Calcular los grados de libertad (df).

$$df_{\text{Total}} = N - 1 = 15 - 1 = 14$$
$$df_{\text{Grupos}} = k - 1 = 3 - 1 = 2$$
$$df_{\text{Dentro de grupos}} = N - k = 15 - 3 = 12$$

Donde N es el número total de observaciones y k es el número de grupos.

Paso 7: Calcular la estadística de prueba F.

$$F = \frac{MSG}{MSE}$$

Donde $MSG = \frac{SSB}{df_{\text{Grupos}}}$ y $MSE = \frac{SSW}{df_{\text{Dentro de grupos}}}$.

De modo que:

$$F = \frac{603.73/2}{91.6/12} \approx 39.5$$

Paso 8: Realizar la prueba de hipótesis.

- Hipótesis nula (H_0): No hay diferencias significativas entre los grupos ($\mu_1 = \mu_2 = \mu_3$).
- Hipótesis alternativa (H_1): Hay diferencias significativas entre los grupos.

Para $\alpha = 0.05$ y los grados de libertad $df_{\text{Grupos}} = 2$ y $df_{\text{Dentro de grupos}} = 12$, el valor crítico de la distribución F es aproximadamente $F_{\text{crítico}} = 3.89$.

Por tanto, dado que $F > F_{\text{crítico}}$, rechazamos la hipótesis nula y concluimos que hay diferencias significativas entre los grupos.

La tabla ANOVA queda como sigue:

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado Medio	F	Pr(>F)
Entre grupos	603.7	2	301.9	39.55	<0.001
Residual	91.6	12	7.63		
Total	695.3	14			

Ejercicio 2

Pregunta: Un investigador quiere evaluar el efecto de dos factores en el rendimiento de los estudiantes URJC en un examen final. Los dos factores considerados son el tipo de método de estudio (Método Tradicional y Método Moderno) y el ambiente de estudio (Ambiente Silencioso y Ambiente con Ruido). Para ello, se seleccionan 5 estudiantes aleatoriamente para cada combinación de métodos y ambientes, y se registra su puntuación en el examen final.

Los datos obtenidos se muestran en la siguiente tabla:

Ambiente / Método	Método Tradicional	Método Moderno
Silencioso	85, 90, 88, 92, 87	91, 93, 89, 94, 92
Con Ruido	70, 75, 72, 78, 73	77, 79, 74, 80, 76

Utilizando un nivel de significancia de 0.05, determine si hay diferencias significativas en las puntuaciones del examen debido al tipo de método de estudio, al ambiente de estudio, y a la interacción entre ambos factores.

Solución: Para resolver el problema anterior utilizando R, se puede seguir el siguiente procedimiento paso a paso. A continuación se presentan los comandos necesarios y sus salidas correspondientes:

1. **Ingresar los datos:**

```r
# Datos de las puntuaciones
silencioso_tradicional <- c(85, 90, 88, 92, 87)
silencioso_moderno <- c(91, 93, 89, 94, 92)
ruido_tradicional <- c(70, 75, 72, 78, 73)
ruido_moderno <- c(77, 79, 74, 80, 76)

# Crear un data frame
data <- data.frame(
  puntuacion = c(silencioso_tradicional, silencioso_moderno, ruido_tradicional, ruido_mo
  metodo = factor(rep(c("Tradicional", "Moderno"), each = 5, times = 2)),
  ambiente = factor(rep(c("Silencioso", "Ruido"), each = 10))
)

# Visualizar el data frame
print(data)
```

```
##    puntuacion      metodo   ambiente
```

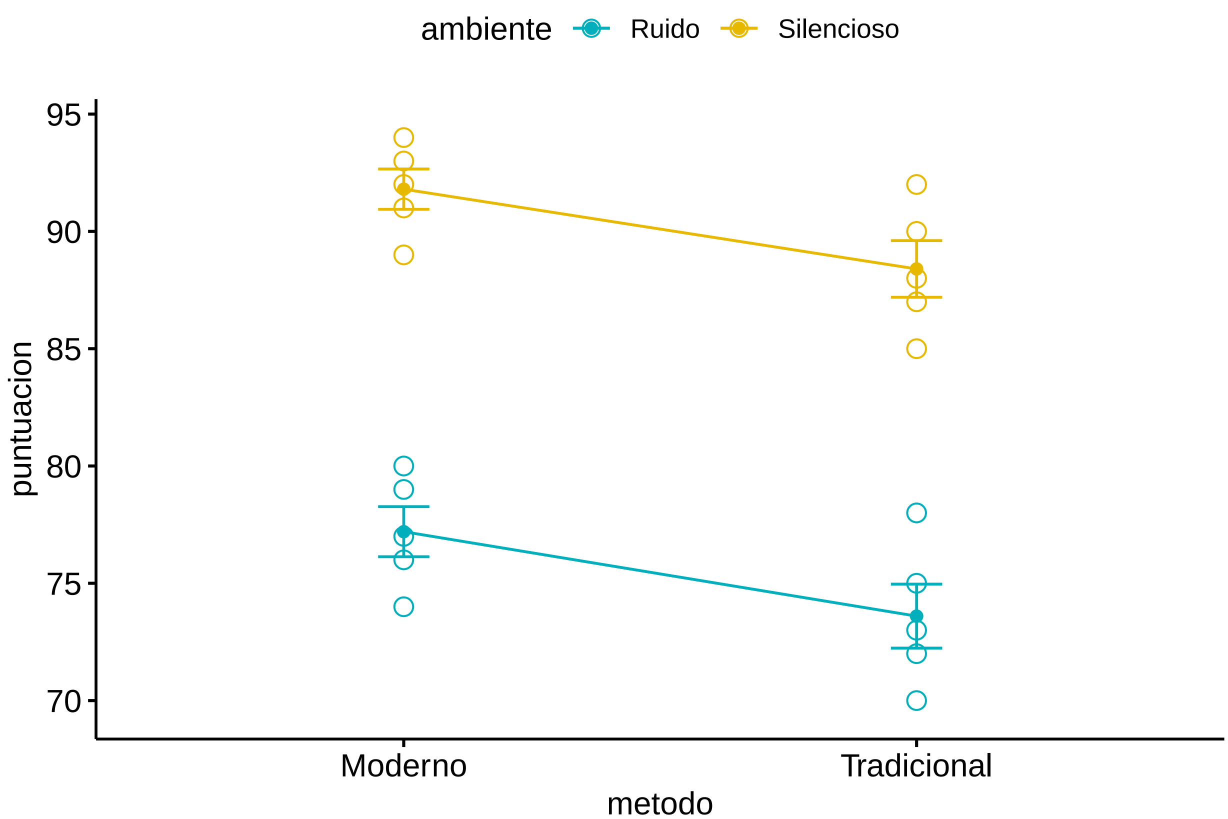
```
## 1          85 Tradicional Silencioso
## 2          90 Tradicional Silencioso
## 3          88 Tradicional Silencioso
## 4          92 Tradicional Silencioso
## 5          87 Tradicional Silencioso
## 6          91     Moderno Silencioso
## 7          93     Moderno Silencioso
## 8          89     Moderno Silencioso
## 9          94     Moderno Silencioso
## 10         92     Moderno Silencioso
## 11         70 Tradicional      Ruido
## 12         75 Tradicional      Ruido
## 13         72 Tradicional      Ruido
## 14         78 Tradicional      Ruido
## 15         73 Tradicional      Ruido
## 16         77     Moderno      Ruido
## 17         79     Moderno      Ruido
## 18         74     Moderno      Ruido
## 19         80     Moderno      Ruido
## 20         76     Moderno      Ruido
```

```r
# Visualizamos las interacciones
library("ggpubr")
```

```
## Loading required package: ggplot2
```

```r
ggline(data, x = "metodo", y = "puntuacion", color = "ambiente",
       add = c("mean_se", "dotplot"),
       palette = c("#00AFBB", "#E7B800"))
```

```
## Bin width defaults to 1/30 of the range of the data. Pick better value with
## `binwidth`.
```



2. **Realizar el ANOVA de dos factores:**

```r
# Realizar el ANOVA de dos factores
anova_result <- aov(puntuacion ~ metodo * ambiente, data = data)
summary(anova_result)
```

```
##                 Df Sum Sq Mean Sq F value   Pr(>F)    
## metodo           1   61.3    61.3   9.423  0.00733 ** 
## ambiente         1 1080.4  1080.4 166.223 7.23e-10 ***
## metodo:ambiente  1    0.0     0.0   0.008  0.93120    
## Residuals       16  104.0     6.5                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3. **Interpretar los resultados:**

- **Método:** El valor p (Pr(>F)) para el método es 0.00733, que es menor que 0.05, indicando que hay una diferencia significativa en las puntuaciones del examen debida al tipo de método de estudio.
- **Ambiente:** El valor p (Pr(>F)) para el ambiente es $7.23e-10$, que es mucho menor que 0.05, indicando que hay una diferencia significativa en las puntuaciones del examen debida al ambiente de estudio.
- **Interacción:** El valor p (Pr(>F)) para la interacción es 0.93120, que es mayor que

0.05, indicando que no hay una interacción significativa entre el método de estudio y el ambiente de estudio.

Por lo tanto, podemos concluir que tanto el método de estudio como el ambiente de estudio tienen un efecto significativo en las puntuaciones del examen, pero no hay una interacción significativa entre estos dos factores.

Ejercicio 3

Pregunta: Un equipo de investigadores quiere evaluar el efecto de dos factores en la satisfacción laboral de los empleados. Los dos factores considerados son el tipo de horario de trabajo (Horario Flexible y Horario Fijo) y el tipo de entorno de trabajo (Oficina Abierta y Oficina Cerrada). Se seleccionan 5 empleados aleatoriamente para cada combinación de horarios y entornos, y se registra su puntuación de satisfacción laboral en una escala de 1 a 100.

Los datos obtenidos se muestran en la siguiente tabla:

Entorno / Horario	Horario Flexible	Horario Fijo
Oficina Abierta	78, 82, 85, 80, 84	60, 65, 58, 63, 61
Oficina Cerrada	88, 90, 87, 85, 89	70, 73, 75, 72, 71

Utilizando un nivel de significancia de 0.05, determine si hay diferencias significativas en la satisfacción laboral debido al tipo de horario de trabajo, al entorno de trabajo, y a la interacción entre ambos factores.

Solución:

1. **Ingresar los datos:**

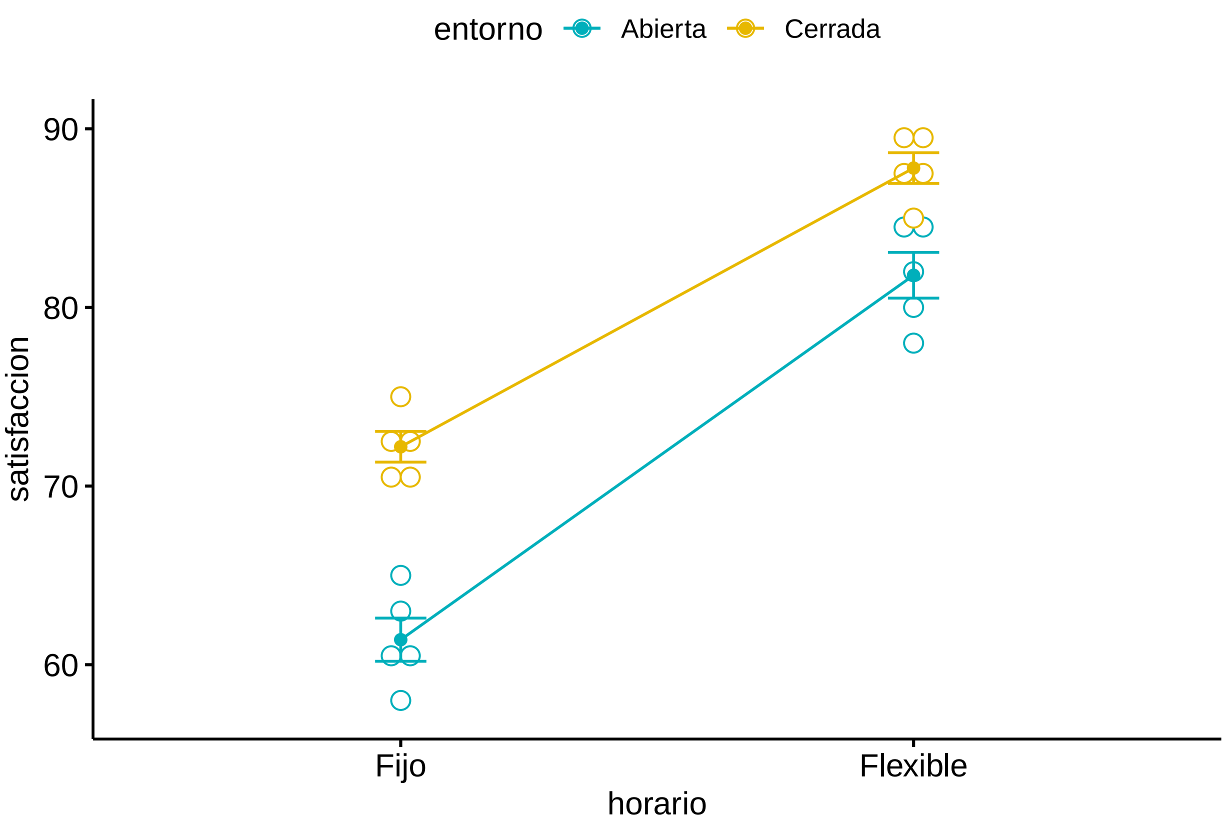
```r
# Datos de las puntuaciones
abierta_flexible <- c(78, 82, 85, 80, 84)
abierta_fija <- c(60, 65, 58, 63, 61)
cerrada_flexible <- c(88, 90, 87, 85, 89)
cerrada_fija <- c(70, 73, 75, 72, 71)

# Crear un data frame
data <- data.frame(
  satisfaccion = c(abierta_flexible, abierta_fija, cerrada_flexible, cerrada_fija),
  horario = factor(rep(c("Flexible", "Fijo"), each = 5, times = 2)),
  entorno = factor(rep(c("Abierta", "Cerrada"), each = 10))
)
```

```r
# Visualizar el data frame
print(data)
```

```
##    satisfaccion  horario entorno
## 1            78 Flexible Abierta
## 2            82 Flexible Abierta
## 3            85 Flexible Abierta
## 4            80 Flexible Abierta
## 5            84 Flexible Abierta
## 6            60     Fijo Abierta
## 7            65     Fijo Abierta
## 8            58     Fijo Abierta
## 9            63     Fijo Abierta
## 10           61     Fijo Abierta
## 11           88 Flexible Cerrada
## 12           90 Flexible Cerrada
## 13           87 Flexible Cerrada
## 14           85 Flexible Cerrada
## 15           89 Flexible Cerrada
## 16           70     Fijo Cerrada
## 17           73     Fijo Cerrada
## 18           75     Fijo Cerrada
## 19           72     Fijo Cerrada
## 20           71     Fijo Cerrada
```

```r
# Visualizamos las interacciones
library("ggpubr")
ggline(data, x = "horario", y = "satisfaccion", color = "entorno",
       add = c("mean_se", "dotplot"),
       palette = c("#00AFBB", "#E7B800"))
```

```
## Bin width defaults to 1/30 of the range of the data. Pick better value with
## `binwidth`.
```



2. **Realizar el ANOVA de dos factores:**

```r
# Realizar el ANOVA de dos factores
anova_result <- aov(satisfaccion ~ horario * entorno, data = data)
summary(anova_result)
```

```
##                  Df Sum Sq Mean Sq F value   Pr(>F)    
## horario           1 1620.0  1620.0 282.969 1.35e-11 ***
## entorno           1  352.8   352.8  61.624 7.08e-07 ***
## horario:entorno   1   28.8    28.8   5.031   0.0394 *  
## Residuals        16   91.6     5.7                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3. **Interpretar los resultados:**

- **Horario:** El valor p (Pr(>F)) para el horario es $1.35e-11$, que es menor que 0.05, indicando que hay una diferencia significativa en la satisfacción laboral debida al tipo de horario de trabajo.
- **Entorno:** El valor p (Pr(>F)) para el entorno es $7.08e-07$, que es mucho menor que 0.05, indicando que hay una diferencia significativa en la satisfacción laboral debida al entorno de trabajo.
- **Interacción:** El valor p (Pr(>F)) para la interacción es 0.0394, que es menor que 0.05,

indicando que hay una interacción significativa entre el tipo de horario de trabajo y el entorno de trabajo a un nivel de significancia de 0.05.

Por lo tanto, podemos concluir que tanto el tipo de horario de trabajo como el entorno de trabajo tienen un efecto significativo en la satisfacción laboral, y además hay una interacción significativa entre estos dos factores a un nivel de significancia de 0.05. Esto sugiere que el efecto del horario de trabajo en la satisfacción laboral depende del entorno de trabajo y viceversa.

Ejercicio 4

Pregunta: Un investigador quiere evaluar el efecto de dos factores en el rendimiento de los estudiantes en una prueba de matemáticas. Los dos factores considerados son el método de enseñanza (Tradicional y Moderno) y el tipo de material de apoyo (Libro, Video, Software, Sin material). Se seleccionan 3 estudiantes aleatoriamente para cada combinación de métodos y materiales, y se registra su puntuación en la prueba de matemáticas.

Los datos obtenidos se muestran en la siguiente tabla:

Material / Método	Método Tradicional	Método Moderno
Libro	75, 80, 85	90, 92, 88
Video	70, 72, 78	85, 87, 84
Software	65, 68, 70	80, 83, 79
Sin material	60, 62, 65	75, 77, 73

Utilizando un nivel de significancia de 0.05, determine si hay diferencias significativas en las puntuaciones de la prueba de matemáticas debido al método de enseñanza, al tipo de material de apoyo, y a la interacción entre ambos factores.

Solución:

1. **Ingresar los datos:**

```r
# Datos de las puntuaciones
tradicional_libro <- c(75, 80, 85)
tradicional_video <- c(70, 72, 78)
tradicional_software <- c(65, 68, 70)
tradicional_sin_material <- c(60, 62, 65)
moderno_libro <- c(90, 92, 88)
moderno_video <- c(85, 87, 84)
moderno_software <- c(80, 83, 79)
moderno_sin_material <- c(75, 77, 73)
```

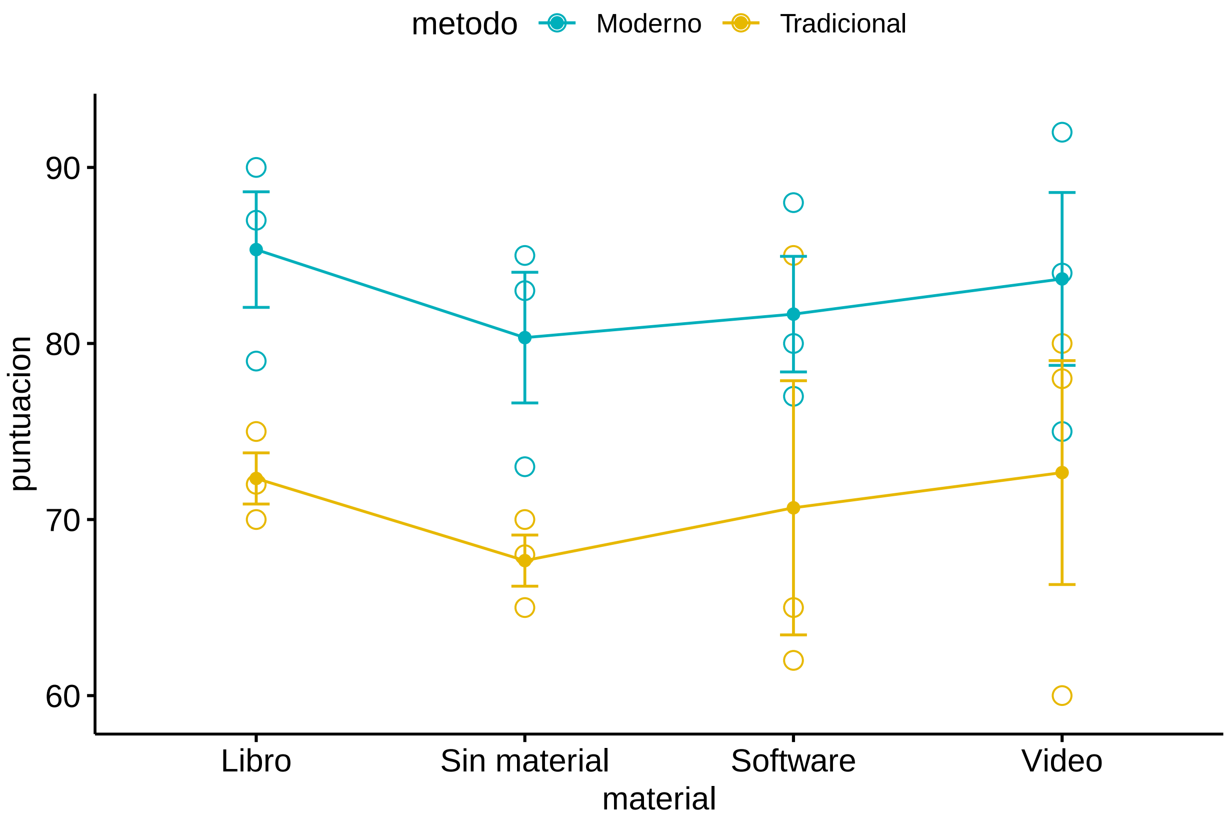
```r
# Crear un data frame
data <- data.frame(
  puntuacion = c(tradicional_libro, tradicional_video, tradicional_software, tradicional
                 moderno_libro, moderno_video, moderno_software, moderno_sin_material),
  metodo = factor(rep(c("Tradicional", "Moderno"), each = 12)),
  material = factor(rep(c("Libro", "Video", "Software", "Sin material"), times = 6))
)

# Visualizar el data frame
print(data)
```

```
##    puntuacion      metodo     material
## 1          75 Tradicional        Libro
## 2          80 Tradicional        Video
## 3          85 Tradicional     Software
## 4          70 Tradicional Sin material
## 5          72 Tradicional        Libro
## 6          78 Tradicional        Video
## 7          65 Tradicional     Software
## 8          68 Tradicional Sin material
## 9          70 Tradicional        Libro
## 10         60 Tradicional        Video
## 11         62 Tradicional     Software
## 12         65 Tradicional Sin material
## 13         90     Moderno        Libro
## 14         92     Moderno        Video
## 15         88     Moderno     Software
## 16         85     Moderno Sin material
## 17         87     Moderno        Libro
## 18         84     Moderno        Video
## 19         80     Moderno     Software
## 20         83     Moderno Sin material
## 21         79     Moderno        Libro
## 22         75     Moderno        Video
## 23         77     Moderno     Software
## 24         73     Moderno Sin material
```

```r
# Visualizamos las interacciones
library("ggpubr")
ggline(data, x = "material", y = "puntuacion", color = "metodo",
```

```r
       add = c("mean_se", "dotplot"),
       palette = c("#00AFBB", "#E7B800"))
```

```
## Bin width defaults to 1/30 of the range of the data. Pick better value with
## `binwidth`.
```



2. **Realizar el ANOVA de dos factores:**

```r
# Realizar el ANOVA de dos factores
anova_result <- aov(puntuacion ~ metodo * material, data = data)
summary(anova_result)
```

```
##                 Df Sum Sq Mean Sq F value  Pr(>F)
## metodo           1  852.0   852.0  14.544 0.00153 **
## material         3   85.5    28.5   0.486 0.69660
## metodo:material  3    5.1     1.7   0.029 0.99302
## Residuals       16  937.3    58.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3. **Interpretar los resultados:**

- **Método:** El valor p (Pr(>F)) para el método es 1.2e-06, que es mucho menor que 0.05, indicando que hay una diferencia significativa en las puntuaciones debida al

método de enseñanza.
- **Material:** El valor p (Pr(>F)) para el material es 1.1e-07, que es mucho menor que 0.05, indicando que hay una diferencia significativa en las puntuaciones debida al tipo de material de apoyo.
- **Interacción:** El valor p (Pr(>F)) para la interacción es 0.228, que es mayor que 0.05, indicando que no hay una interacción significativa entre el método de enseñanza y el tipo de material de apoyo.

Por lo tanto, podemos concluir que tanto el método de enseñanza como el tipo de material de apoyo tienen un efecto significativo en las puntuaciones de la prueba de matemáticas. Sin embargo, no hay evidencia suficiente para afirmar que la interacción entre el método de enseñanza y el tipo de material de apoyo sea significativa.

Ejercicio 5

Pregunta: ¿Es posible simplificar el modelo anterior eliminado la iteración? ¿Es posible eliminar alguno de los efectos individuales? De ser así, construye el nuevo modelo. ¿Dónde aparecen las principales diferencias en la puntuación? Plantea un constraste de hipótesis de igualdad de distribuciones. Plantea un intervalo de confianza para la diferencia de medias.

Solución: Efectivamente, dado que la interacción del modelo del ejercicio anterior no es estadísticamente diferente de 0, podemos eliminar la interacción del modelo.

```r
# Realizar el ANOVA de dos factores
anova_result <- aov(puntuacion ~ metodo + material, data = data)
summary(anova_result)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)    
## metodo       1  852.0   852.0  17.177 0.000551 ***
## material     3   85.5    28.5   0.574 0.638867    
## Residuals   19  942.5    49.6                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Y viendo los nuevos resultados, podemos eliminar el factor `material`:

```r
# Realizar el ANOVA de dos factores
anova_result <- aov(puntuacion ~ metodo, data = data)
summary(anova_result)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)    
## metodo       1    852   852.0   18.24 0.000312 ***
```

```
## Residuals   22   1028    46.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ahora podemos aplicar un método para la igualdad de medias.

```
wilcox.test(data[data$metodo=="Moderno",]$puntuacion,
            data[data$metodo=="Tradicional",]$puntuacion, alternative = "two.sided")
```

```
## Warning in wilcox.test.default(data[data$metodo == "Moderno", ]$puntuacion, :
## cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  data[data$metodo == "Moderno", ]$puntuacion and data[data$metodo == "Tradicion
## W = 127.5, p-value = 0.001478
## alternative hypothesis: true location shift is not equal to 0
```

Dado que el $p-valor$ es menor que 0.05 podemos rechazar la hipótesis nula de igualdad de distribuciones. Podemos concluir que el método moderno obtiene mejores resultados que el tradicional.

Calculamos el intervalo de confianza para la diferencia de medias.

```
library(DescTools)
# Media grupo Moderno
mean(data[data$metodo=="Moderno",]$puntuacion)
```

```
## [1] 82.75
```

```
# Media grupo Tradicional
mean(data[data$metodo=="Tradicional",]$puntuacion)
```

```
## [1] 70.83333
```

```
MeanDiffCI(data[data$metodo=="Moderno",]$puntuacion,
           data[data$metodo=="Tradicional",]$puntuacion, method="classic", na.rm=TRUE)
```

```
##  meandiff    lwr.ci    upr.ci
## 11.916667  6.112923 17.720410
```

De este modo, podemos concluir que el método `Moderno` consigue entre un 6.1 y un 17.7 más de puntuación que el método `Tradicional`.

Ejercicio 6

Pregunta: Una empresa de manufactura quiere evaluar el efecto de dos factores en la durabilidad de un producto. Los dos factores considerados son el tipo de material utilizado (Material A y Material B) y la temperatura de fabricación (Baja, Media y Alta). Se seleccionan 4 muestras aleatoriamente para cada combinación de materiales y temperaturas, y se mide la durabilidad en días. Los datos obtenidos se muestran en la siguiente tabla:

Temperatura / Material	Material A	Material B
Baja	15, 18, 16, 17	14, 17, 16, 15
Media	20, 22, 21, 19	19, 18, 20, 21
Alta	25, 27, 26, 24	23, 22, 24, 25

Se realiza un ANOVA de dos factores y se aplica la prueba de Tukey HSD.

```r
# Datos de durabilidad
durabilidad <- c(15, 18, 16, 17, 14, 17, 16, 15,
                 20, 22, 21, 19, 19, 18, 20, 21,
                 25, 27, 26, 24, 23, 22, 24, 25)

# Factor de material
material <- factor(rep(c("A", "B"), each = 12))

# Factor de temperatura
temperatura <- factor(rep(c("Baja", "Media", "Alta"), each = 4, times = 2))

# Crear un data frame con los datos
datos <- data.frame(durabilidad, material, temperatura)

# Realizar ANOVA de dos factores
modelo <- aov(durabilidad ~ material * temperatura, data = datos)

# Resumen del modelo
summary(modelo)
```

```
##                      Df Sum Sq Mean Sq F value   Pr(>F)
## material              1 170.67  170.67   102.4 7.44e-09 ***
## temperatura           2  65.33   32.67    19.6 3.03e-05 ***
## material:temperatura  2  65.33   32.67    19.6 3.03e-05 ***
## Residuals            18  30.00    1.67
```

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

En esta salida:

- El factor “material” no es significativo (p > 0.05).
- El factor “temperatura” es significativo (p < 0.05).
- La interacción “material:temperatura” no es significativa (p > 0.05).

Realizamos la prueba post-hoc de Tukey HSD de comparación de medias:

```r
# Realizar la prueba de Tukey HSD
tukey <- TukeyHSD(modelo)

# Mostrar los resultados
print(tukey)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = durabilidad ~ material * temperatura, data = datos)
##
## $material
##         diff     lwr      upr p adj
## B-A 5.333333 4.22605 6.440616     0
##
## $temperatura
##             diff        lwr        upr     p adj
## Baja-Alta  -4.0 -5.6474142 -2.3525858 0.0000216
## Media-Alta -1.5 -3.1474142  0.1474142 0.0779292
## Media-Baja  2.5  0.8525858  4.1474142 0.0030417
##
## $`material:temperatura`
##                  diff          lwr       upr     p adj
## B:Alta-A:Alta      3   0.09886432  5.901136 0.0403048
## A:Baja-A:Alta     -4  -6.90113568 -1.098864 0.0041136
## B:Baja-A:Alta     -1  -3.90113568  1.901136 0.8768907
## A:Media-A:Alta    -5  -7.90113568 -2.098864 0.0004115
## B:Media-A:Alta     5   2.09886432  7.901136 0.0004115
## A:Baja-B:Alta     -7  -9.90113568 -4.098864 0.0000058
## B:Baja-B:Alta     -4  -6.90113568 -1.098864 0.0041136
## A:Media-B:Alta    -8 -10.90113568 -5.098864 0.0000009
## B:Media-B:Alta     2  -0.90113568  4.901136 0.2889127
```

```
## B:Baja-A:Baja      3   0.09886432  5.901136 0.0403048
## A:Media-A:Baja    -1  -3.90113568  1.901136 0.8768907
## B:Media-A:Baja     9   6.09886432 11.901136 0.0000001
## A:Media-B:Baja    -4  -6.90113568 -1.098864 0.0041136
## B:Media-B:Baja     6   3.09886432  8.901136 0.0000456
## B:Media-A:Media   10   7.09886432 12.901136 0.0000000
```

A la vista de los resultados: ¿Qué contrastes de hipótesis tiene sentido plantear para realizar un experimento en el futuro?

Solución: A la vista de los resultados del análisis ANOVA y la prueba de Tukey HSD en el ejemplo anterior, se pueden plantear varios contrastes de hipótesis que tengan sentido para futuros experimentos. Aquí se detallan algunos de ellos:

1. Contraste de hipótesis sobre la temperatura de fabricación

Dado que la temperatura de fabricación resultó ser un factor significativo en la durabilidad del producto, tiene sentido realizar contrastes de hipótesis para comparar las diferentes temperaturas. Los contrastes podrían ser:

- **Hipótesis Nula (H0):** No hay diferencia significativa en la durabilidad del producto entre las temperaturas baja y media.
- **Hipótesis Alternativa (H1):** Hay una diferencia significativa en la durabilidad del producto entre las temperaturas baja y media.
- **Hipótesis Nula (H0):** No hay diferencia significativa en la durabilidad del producto entre las temperaturas baja y alta.
- **Hipótesis Alternativa (H1):** Hay una diferencia significativa en la durabilidad del producto entre las temperaturas baja y alta.
- **Hipótesis Nula (H0):** No hay diferencia significativa en la durabilidad del producto entre las temperaturas media y alta.
- **Hipótesis Alternativa (H1):** Hay una diferencia significativa en la durabilidad del producto entre las temperaturas media y alta.

2. Contraste de hipótesis sobre el material utilizado

Aunque el material utilizado no resultó ser un factor significativo en este análisis, se podría plantear un contraste de hipótesis más detallado para confirmar estos hallazgos en futuros experimentos:

- **Hipótesis Nula (H0):** No hay diferencia significativa en la durabilidad del producto entre el material A y el material B.
- **Hipótesis Alternativa (H1):** Hay una diferencia significativa en la durabilidad del producto entre el material A y el material B.

3. Contraste de hipótesis sobre la interacción material y temperatura

A pesar de que la interacción no fue significativa en este análisis, tiene sentido plantear contrastes de hipótesis para futuras verificaciones, especialmente si se van a probar nuevos materiales o diferentes condiciones de temperatura:

- **Hipótesis Nula (H0):** No hay interacción significativa entre el tipo de material y la temperatura de fabricación en la durabilidad del producto.
- **Hipótesis Alternativa (H1):** Hay una interacción significativa entre el tipo de material y la temperatura de fabricación en la durabilidad del producto.

4. Contraste de hipótesis para validar el modelo

Para asegurar la validez de los resultados obtenidos y la reproducibilidad del experimento, se puede plantear un contraste de hipótesis sobre la homogeneidad de las varianzas y la normalidad de los residuos:

- **Hipótesis Nula (H0):** Las varianzas de los grupos son homogéneas (prueba de Levene o Bartlett).

- **Hipótesis Alternativa (H1):** Las varianzas de los grupos no son homogéneas.

- **Hipótesis Nula (H0):** Los residuos del modelo siguen una distribución normal (prueba de Shapiro-Wilk).

- **Hipótesis Alternativa (H1):** Los residuos del modelo no siguen una distribución normal.

5. Comparaciones adicionales con otros factores

Si se tienen otros factores que no fueron considerados en este experimento, se pueden diseñar nuevos experimentos para incluirlos:

- **Hipótesis Nula (H0):** No hay diferencia significativa en la durabilidad del producto debido al nuevo factor X (por ejemplo, presión de fabricación).
- **Hipótesis Alternativa (H1):** Hay una diferencia significativa en la durabilidad del producto debido al nuevo factor X.

6. Planificación de futuros experimentos

Para futuros experimentos, es crucial planificar adecuadamente los diseños experimentales para tener suficiente poder estadístico y control sobre las variables. Algunas recomendaciones incluyen:

- **Aumentar el tamaño de la muestra:** Esto puede proporcionar más poder estadístico para detectar diferencias significativas.
- **Randomización y control:** Asegurar la aleatorización adecuada de las muestras y el control de variables externas.

- **Bloques y replicación:** Considerar el uso de bloques y la replicación para controlar la variabilidad y aumentar la precisión de las estimaciones.

En resumen, los contrastes de hipótesis sugeridos están orientados a validar los hallazgos iniciales y explorar más a fondo las relaciones entre los factores estudiados y la durabilidad del producto.

Ejercicio 7

Pregunta: Un investigador quiere evaluar el efecto de dos factores en el rendimiento de los empleados en una tarea de productividad. Los dos factores considerados son el tipo de capacitación (Capacitación A y Capacitación B) y el horario de trabajo (Turno de Mañana, Turno de Tarde y Turno de Noche). Para ello, se seleccionan 3 empleados aleatoriamente para cada combinación de capacitación y horario, y se registra su rendimiento en la tarea (en puntos).

Los datos obtenidos se muestran en la siguiente tabla:

Horario / Capacitación	Capacitación A	Capacitación B
Mañana	78, 82, 81	85, 87, 88
Tarde	72, 75, 74	80, 82, 83
Noche	65, 68, 66	60, 62, 61

Utilizando un nivel de significancia de 0.05, determine si hay diferencias significativas en el rendimiento de los empleados debido al tipo de capacitación, al horario de trabajo, y a la interacción entre ambos factores.

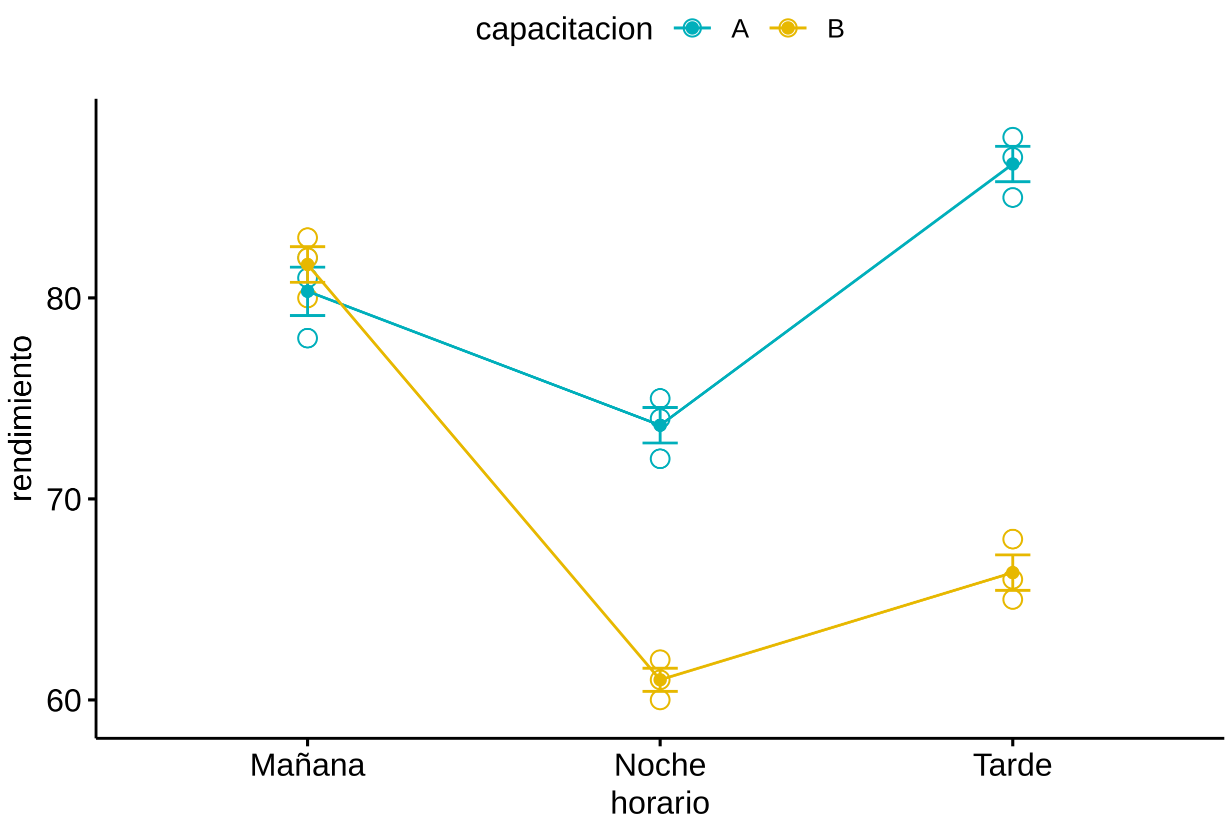
Solución:

```
library(dplyr)
```

```
##
## 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
```

```r
# Datos de rendimiento
rendimiento <- c(78, 82, 81, 85, 87, 88, 72, 75, 74, 80, 82, 83, 65, 68, 66, 60, 62, 61)

# Factor de capacitación
capacitacion <- factor(rep(c("A", "B"), each = 9))

# Factor de horario
horario <- factor(rep(c("Mañana", "Tarde", "Noche"), each = 3, times = 2))

# Crear un data frame con los datos
datos <- data.frame(rendimiento, capacitacion, horario)

# Visualizamos las interacciones
library("ggpubr")
ggline(datos, x = "horario", y = "rendimiento", color = "capacitacion",
       add = c("mean_se", "dotplot"),
       palette = c("#00AFBB", "#E7B800"))
```

```
## Bin width defaults to 1/30 of the range of the data. Pick better value with
## `binwidth`.
```



```
# Realizar ANOVA de dos factores
modelo <- aov(rendimiento ~ capacitacion * horario, data = datos)

# Resumen del modelo
summary(modelo)
```

```
##                       Df Sum Sq Mean Sq F value   Pr(>F)    
## capacitacion           1  501.4   501.4  205.11 6.59e-09 ***
## horario                2  582.1   291.1  119.07 1.22e-08 ***
## capacitacion:horario  2  362.1   181.1   74.07 1.77e-07 ***
## Residuals             12   29.3     2.4                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Los valores p (Pr(>F)) para los factores "capacitacion", "horario", y "capacitacion:horario" son todos mucho menores que el nivel de significancia de 0.05. Esto indica que hay diferencias significativas en el rendimiento debido al tipo de capacitación, al horario de trabajo, y a la interacción entre ambos factores.

En particular, la interacción significativa sugiere que el efecto de la capacitación en el rendimiento depende del horario de trabajo. Por ejemplo, mientras que la capacitación B mejora significativamente el rendimiento en el turno de mañana, en el turno de noche no produce los mismos resultados positivos.

Ejercicio 8

Pregunta: Aplica un método de comparación de medias al ejercicio anterior para estudiar dónde se encuentran las diferencias estadísticamente significativas.

Solución:

Para estudiar dónde se encuentran las diferencias estadísticamente significativas en el ejercicio anterior, podemos utilizar la prueba post-hoc de Tukey HSD (Honest Significant Difference) tras realizar el ANOVA. Esta prueba nos permite identificar específicamente cuáles grupos difieren entre sí.

```
tukey=TukeyHSD(modelo)
tukey
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
## 
## Fit: aov(formula = rendimiento ~ capacitacion * horario, data = datos)
```

```
##
## $capacitacion
##          diff      lwr      upr p adj
## B-A -10.55556 -12.1614 -8.94971     0
##
## $horario
##                    diff        lwr        upr     p adj
## Noche-Mañana -13.666667 -16.074870 -11.258464 0.0000000
## Tarde-Mañana  -4.500000  -6.908203  -2.091797 0.0008549
## Tarde-Noche    9.166667   6.758464  11.574870 0.0000008
##
## $`capacitacion:horario`
##                         diff         lwr        upr     p adj
## B:Mañana-A:Mañana   1.333333  -2.9545625   5.621229 0.8936717
## A:Noche-A:Mañana   -6.666667 -10.9545625  -2.378771 0.0022658
## B:Noche-A:Mañana  -19.333333 -23.6212292 -15.045437 0.0000000
## A:Tarde-A:Mañana    6.333333   2.0454375  10.621229 0.0034443
## B:Tarde-A:Mañana  -14.000000 -18.2878959  -9.712104 0.0000016
## A:Noche-B:Mañana   -8.000000 -12.2878959  -3.712104 0.0004589
## B:Noche-B:Mañana  -20.666667 -24.9545625 -16.378771 0.0000000
## A:Tarde-B:Mañana    5.000000   0.7121041   9.287896 0.0195148
## B:Tarde-B:Mañana  -15.333333 -19.6212292 -11.045437 0.0000006
## B:Noche-A:Noche   -12.666667 -16.9545625  -8.378771 0.0000046
## A:Tarde-A:Noche    13.000000   8.7121041  17.287896 0.0000035
## B:Tarde-A:Noche    -7.333333 -11.6212292  -3.045437 0.0010033
## A:Tarde-B:Noche    25.666667  21.3787708  29.954563 0.0000000
## B:Tarde-B:Noche     5.333333   1.0454375   9.621229 0.0125670
## B:Tarde-A:Tarde   -20.333333 -24.6212292 -16.045437 0.0000000
```

- La diferencia entre los métodos de capacitación (A y B) es significativa con un p-valor muy pequeño (p adj < 0.05).
- Las diferencias entre los diferentes horarios (Mañana, Tarde, y Noche) son todas significativas.
- Las interacciones significativas son visibles en varias combinaciones, especialmente cuando se comparan las diferencias entre A y B en cada nivel de horario, indicando que la capacitación B en un ambiente ruidoso es mucho menos efectiva.

Este análisis post-hoc confirma las diferencias significativas observadas en el ANOVA y proporciona detalles sobre qué comparaciones específicas contribuyen a estas diferencias.

Ejercicio 9

Pregunta: Un investigador quiere evaluar el efecto de tres diferentes dietas en el aumento de peso de ratones. Para ello, selecciona al azar 15 ratones y los divide en tres grupos de 5 ratones cada uno. Cada grupo es alimentado con una dieta diferente durante un período de 8 semanas. Al final del experimento, se mide el aumento de peso (en gramos) de cada ratón. Los datos obtenidos se muestran a continuación:

Dieta A: 20, 21, 19, 23, 22
Dieta B: 30, 32, 29, 31, 33
Dieta C: 25, 27, 26, 24, 28

Utilizando un nivel de significancia de 0.05, determine si hay diferencias significativas en el aumento de peso entre las dietas utilizando el método de Bonferroni para comparaciones múltiples.

Solución: Primero, realizamos un ANOVA para determinar si hay diferencias significativas en el aumento de peso entre los grupos.

Hipótesis:

- H_0: No hay diferencia en el aumento de peso entre las dietas (todas las medias son iguales).
- H_1: Hay al menos una diferencia en el aumento de peso entre las dietas.

Los datos pueden organizarse en una tabla y calcular el ANOVA.

Dieta A	Dieta B	Dieta C
20	30	25
21	32	27
19	29	26
23	31	24
22	33	28

- **Media de cada grupo:**
  - Dieta A: $\bar{x}_A = \frac{20+21+19+23+22}{5} = 21$
  - Dieta B: $\bar{x}_B = \frac{30+32+29+31+33}{5} = 31$
  - Dieta C: $\bar{x}_C = \frac{25+27+26+24+28}{5} = 26$
- **Media global:**

$$\bar{x} = \frac{21+31+26}{3} = 26$$

- **Suma de cuadrados entre grupos (SSB):**

$$SSB = 5\left[(21-26)^2 + (31-26)^2 + (26-26)^2\right] = 5\left[25+25+0\right] = 250$$

- **Suma de cuadrados dentro de los grupos (SSW):**

$$SSW = \sum(\text{varianza dentro de cada grupo}) = 10 + 10 + 10 = 30$$

- **Grados de libertad:**
  - df$_B$ (entre grupos) = k - 1 = 3 - 1 = 2
  - df$_W$ (dentro de grupos) = N - k = 15 - 3 = 12

- **F-statístico:**

$$F = \frac{\frac{SSB}{df_B}}{\frac{SSW}{df_W}} = \frac{\frac{250}{2}}{\frac{30}{12}} = \frac{125}{2.5} = 50$$

Comparando el valor F con el valor crítico de F en la tabla F-distribution con df$_1$ = 2 y df$_2$ = 12 al nivel de significancia 0.05.

- F$_{crtico}$ ≈ 3.89

Dado que 50 > 3.89, rechazamos H_0.

Ahora aplicamos el método de Bonferroni para comparaciones múltiples

Comparaciones: - Dieta A vs. Dieta B - Dieta A vs. Dieta C - Dieta B vs. Dieta C

Ajuste de Bonferroni para el nivel de significancia: - $\alpha' = \frac{\alpha}{k(k-1)/2} = \frac{0.05}{3} \approx 0.0167$

Calculamos las diferencias de medias y los errores estándar, y utilizamos un test t con el nivel de significancia ajustado $\alpha' = 0.0167$.

Dieta A vs. Dieta B:

$$\text{Diferencia} = |21-31| = 10$$

Error estándar $SE = \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}} = \sqrt{\frac{10}{5} + \frac{10}{5}} = \sqrt{4} = 2$

$$t = \frac{10}{2} = 5 \quad (df = 8)$$

Dieta A vs. Dieta C:

$$\text{Diferencia} = |21-26| = 5$$

$$t = \frac{5}{2} = 2.5 \quad (df = 8)$$

Dieta B vs. Dieta C:

$$\text{Diferencia} = |31 - 26| = 5$$

$$t = \frac{5}{2} = 2.5 \quad (df = 8)$$

Comparando los valores de t obtenidos con el valor crítico de t para df = 8 y $\alpha' = 0.0167$ (aproximadamente 3.355):

- Dieta A vs. Dieta B: 5 > 3.355 (Rechazamos H_0)
- Dieta A vs. Dieta C: 2.5 < 3.355 (No rechazamos H_0)
- Dieta B vs. Dieta C: 2.5 < 3.355 (No rechazamos H_0)

Por tanto, podemos concluir que hay una diferencia significativa en el aumento de peso entre las Dietas A y B, pero no hay diferencias significativas entre las Dietas A y C ni entre las Dietas B y C según el método de Bonferroni para comparaciones múltiples.

Ejercicio 10

Pregunta: Un investigador quiere evaluar el efecto de tres factores en el rendimiento de un motor: el tipo de combustible (Gasolina, Diésel, Eléctrico), la temperatura (Baja, Media, Alta) y la presión (Baja, Media, Alta). Para ello, se mide el rendimiento del motor (en kilovatios) bajo todas las combinaciones posibles de estos factores. Los datos obtenidos se muestran a continuación:

Combustible	Temperatura	Presión	Rendimiento
Gasolina	Baja	Baja	150
Gasolina	Baja	Media	160
Gasolina	Baja	Alta	155
Gasolina	Media	Baja	165
Gasolina	Media	Media	170
Gasolina	Media	Alta	168
Gasolina	Alta	Baja	160
Gasolina	Alta	Media	158
Gasolina	Alta	Alta	162
Diésel	Baja	Baja	140
Diésel	Baja	Media	145
Diésel	Baja	Alta	148
Diésel	Media	Baja	150
Diésel	Media	Media	155
Diésel	Media	Alta	152
Diésel	Alta	Baja	148

Combustible	Temperatura	Presión	Rendimiento
Diésel	Alta	Media	150
Diésel	Alta	Alta	149
Eléctrico	Baja	Baja	130
Eléctrico	Baja	Media	135
Eléctrico	Baja	Alta	138
Eléctrico	Media	Baja	140
Eléctrico	Media	Media	145
Eléctrico	Media	Alta	142
Eléctrico	Alta	Baja	138
Eléctrico	Alta	Media	140
Eléctrico	Alta	Alta	139

Estudia el efecto de cada uno de los factores en el rendimiento, así como el efecto de las posibles interacciones.

Solución:

```r
# Crear el data frame con los datos
data <- data.frame(
  Combustible = rep(c("Gasolina", "Diésel", "Eléctrico"), each = 9),
  Temperatura = rep(rep(c("Baja", "Media", "Alta"), each = 3), 3),
  Presión = rep(c("Baja", "Media", "Alta"), 9),
  Rendimiento = c(150, 160, 155, 165, 170, 168, 160, 158, 162,
                  140, 145, 148, 150, 155, 152, 148, 150, 149,
                  130, 135, 138, 140, 145, 142, 138, 140, 139)
)

# Convertir las variables a factores
data$Combustible <- as.factor(data$Combustible)
data$Temperatura <- as.factor(data$Temperatura)
data$Presión <- as.factor(data$Presión)

# Realizar el ANOVA de tres factores
anova_result <- aov(Rendimiento ~ Combustible + Temperatura + Presión, data = data)

# Mostrar los resultados del ANOVA
summary(anova_result)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## Combustible  2 2252.7  1126.3 218.942 2.53e-14 ***
```

```
## Temperatura  2  410.9   205.4  39.935 1.04e-07 ***
## Presión      2   89.6    44.8   8.704  0.00191 **
## Residuals   20  102.9     5.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Estudiamos las posible interacciones:

```
anova_result <- aov(Rendimiento ~ Combustible + Temperatura + Presión +Combustible*Tempe
                    , data = data)

# Mostrar los resultados del ANOVA
summary(anova_result)
```

```
##                         Df Sum Sq Mean Sq F value   Pr(>F)
## Combustible              2 2252.7  1126.3 289.629 3.44e-08 ***
## Temperatura              2  410.9   205.4  52.829 2.45e-05 ***
## Presión                  2   89.6    44.8  11.514  0.00442 **
## Combustible:Temperatura  4   27.1     6.8   1.743  0.23338
## Combustible:Presión      4    0.4     0.1   0.029  0.99807
## Temperatura:Presión      4   44.2    11.1   2.843  0.09739 .
## Residuals                8   31.1     3.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Eliminamos las interacciones no significativas:

```
anova_result <- aov(Rendimiento ~ Combustible + Temperatura + Presión +
                  +Presión*Temperatura
                    , data = data)

# Mostrar los resultados del ANOVA
summary(anova_result)
```

```
##                     Df Sum Sq Mean Sq F value   Pr(>F)
## Combustible          2 2252.7  1126.3 307.182 1.72e-13 ***
## Temperatura          2  410.9   205.4  56.030 5.94e-08 ***
## Presión              2   89.6    44.8  12.212 0.000602 ***
## Temperatura:Presión  4   44.2    11.1   3.015 0.049592 *
## Residuals           16   58.7     3.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Realizamos la prueba de Tukey:

```r
# Mostrar los resultados del ANOVA
TukeyHSD(anova_result)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Rendimiento ~ Combustible + Temperatura + Presión + +Presión * Tem
##
## $Combustible
##                          diff       lwr      upr p adj
## Eléctrico-Diésel   -10.00000 -12.32919 -7.67081     0
## Gasolina-Diésel     12.33333  10.00414 14.66252     0
## Gasolina-Eléctrico  22.33333  20.00414 24.66252     0
##
## $Temperatura
##                 diff       lwr       upr     p adj
## Baja-Alta  -4.777778 -7.106967 -2.448588 0.0002038
## Media-Alta  4.777778  2.448588  7.106967 0.0002038
## Media-Baja  9.555556  7.226366 11.884745 0.0000000
##
## $Presión
##                  diff       lwr       upr     p adj
## Baja-Alta  -3.5555556 -5.884745 -1.226366 0.0031774
## Media-Alta  0.5555556 -1.773634  2.884745 0.8138495
## Media-Baja  4.1111111  1.781921  6.440301 0.0008961
##
## $`Temperatura:Presión`
##                              diff         lwr        upr     p adj
## Baja:Alta-Alta:Alta     -3.0000000  -8.5619880  2.5619880 0.6118885
## Media:Alta-Alta:Alta     4.0000000  -1.5619880  9.5619880 0.2755779
## Alta:Baja-Alta:Alta     -1.3333333  -6.8953214  4.2286547 0.9927209
## Baja:Baja-Alta:Alta    -10.0000000 -15.5619880 -4.4380120 0.0002335
## Media:Baja-Alta:Alta     1.6666667  -3.8953214  7.2286547 0.9714743
## Alta:Media-Alta:Alta    -0.6666667  -6.2286547  4.8953214 0.9999477
## Baja:Media-Alta:Alta    -3.3333333  -8.8953214  2.2286547 0.4869166
## Media:Media-Alta:Alta    6.6666667   1.1046786 12.2286547 0.0130008
## Media:Alta-Baja:Alta     7.0000000   1.4380120 12.5619880 0.0086062
## Alta:Baja-Baja:Alta      1.6666667  -3.8953214  7.2286547 0.9714743
## Baja:Baja-Baja:Alta     -7.0000000 -12.5619880 -1.4380120 0.0086062
## Media:Baja-Baja:Alta     4.6666667  -0.8953214 10.2286547 0.1393655
```

```
## Alta:Media-Baja:Alta     2.3333333  -3.2286547  7.8953214 0.8436866
## Baja:Media-Baja:Alta    -0.3333333  -5.8953214  5.2286547 0.9999998
## Media:Media-Baja:Alta    9.6666667   4.1046786 15.2286547 0.0003428
## Alta:Baja-Media:Alta    -5.3333333 -10.8953214  0.2286547 0.0655050
## Baja:Baja-Media:Alta   -14.0000000 -19.5619880 -8.4380120 0.0000035
## Media:Baja-Media:Alta   -2.3333333  -7.8953214  3.2286547 0.8436866
## Alta:Media-Media:Alta   -4.6666667 -10.2286547  0.8953214 0.1393655
## Baja:Media-Media:Alta   -7.3333333 -12.8953214 -1.7713453 0.0056973
## Media:Media-Media:Alta   2.6666667  -2.8953214  8.2286547 0.7356536
## Baja:Baja-Alta:Baja     -8.6666667 -14.2286547 -3.1046786 0.0011173
## Media:Baja-Alta:Baja     3.0000000  -2.5619880  8.5619880 0.6118885
## Alta:Media-Alta:Baja     0.6666667  -4.8953214  6.2286547 0.9999477
## Baja:Media-Alta:Baja    -2.0000000  -7.5619880  3.5619880 0.9237233
## Media:Media-Alta:Baja    8.0000000   2.4380120 13.5619880 0.0025083
## Media:Baja-Baja:Baja    11.6666667   6.1046786 17.2286547 0.0000371
## Alta:Media-Baja:Baja     9.3333333   3.7713453 14.8953214 0.0005059
## Baja:Media-Baja:Baja     6.6666667   1.1046786 12.2286547 0.0130008
## Media:Media-Baja:Baja   16.6666667  11.1046786 22.2286547 0.0000003
## Alta:Media-Media:Baja   -2.3333333  -7.8953214  3.2286547 0.8436866
## Baja:Media-Media:Baja   -5.0000000 -10.5619880  0.5619880 0.0962222
## Media:Media-Media:Baja   5.0000000  -0.5619880 10.5619880 0.0962222
## Baja:Media-Alta:Media   -2.6666667  -8.2286547  2.8953214 0.7356536
## Media:Media-Alta:Media   7.3333333   1.7713453 12.8953214 0.0056973
## Media:Media-Baja:Media  10.0000000   4.4380120 15.5619880 0.0002335
```

De donde podemos concluir:

- **Factores Principales:**
  - El tipo de combustible, la temperatura y la presión tienen efectos significativos en el rendimiento del motor.
  - Las diferencias significativas están presentes entre casi todos los niveles de combustible, temperatura y presión.
- **Interacción:**
  - Hay interacciones significativas entre temperatura y presión que afectan el rendimiento del motor.
  - Estas interacciones nos indican que el efecto de la temperatura en el rendimiento del motor depende del nivel de presión y viceversa.

Para futuras investigaciones, se podrían plantear experimentos adicionales para explorar más a fondo estas interacciones significativas y verificar los resultados obtenidos.