**Ejercicio Propuesto 9 (Resuelto)**

**La cotinina es uno de los principales metabolitos de la nicotina. Actualmente se le considera el mejor indicador de la exposición al humo de tabaco. Se ha realizado un estudio con distintas marcas de tabaco distinguiendo principalmente entre negro y rubio para detectar las posibles diferencias en el nivel de nicotina de personas expuestas al humo de tabaco. Para ello, se han analizado personas de distintas edades (niños, jóvenes y adultos) y se ha distinguido entre mujeres y hombres. Se han obtenido los datos de la siguiente tabla sobre el nivel de nicotina en miligramos por mililitro.**

|  |  |
| --- | --- |
|  | **Sexo** |
|  | **Hombres** | **Mujeres** |
|  | **Tabaco** | **Tabaco** |
| **Edades** | **Rubio** | **Negro** | **Rubio** | **Negro** |
| **Niños** | **110****240** | **360****125** | **230****219** | **141****123** |
| **Jóvenes** | **112****239** | **252****455** | **655****432** | **873****256** |
| **Adultos** | **652****451** | **354****701** | **653****259** | **198****343** |

**Responder a las siguientes cuestiones:**

1. **Identificar el diseño adecuado a este experimento, escribir el modelo matemático y explicar los distintos elementos que intervienen.**
2. **Contrastar la hipótesis nula de no interacción entre los factores. Adecuar el modelo al resultado de las interacciones y contrastar los efectos principales.**
3. **¿Hay diferencias significativas en el nivel de nicotina en las distintas edades?¿En qué edad el nivel de nicotina es mayor?**
4. **¿El tipo de tabaco es un factor determinante en el nivel de nicotina?**
5. **Comparar el nivel medio de nicotina entre las mujeres y los hombres. ¿Se detectan diferencias significativas?**

**Solución:**

**Responder a las siguientes cuestiones:**

**1. Identificar el diseño adecuado a este experimento, escribir el modelo matemático y explicar los distintos elementos que intervienen.**

Es un diseño factorial de tres factores con replicación, la variable respuesta $y\_{ijkl}$ viene descrita por la siguiente ecuación

Modelo estadístico del diseño factorial de cuatro factores con replicación

$$y\_{ijkl }=μ+τ\_{i }+β\_{j}+γ\_{k}+(τβ)\_{ij}+(τγ)\_{ik}+(βγ)\_{jk}+(τβγ)\_{ijk}+u\_{ijkl} , i=1,2 ; j=1,2; k=1,2,3; l=1,2$$

* ***yijkl***: Nivel de nicotina
* ***µ*:** Efecto constante, común a todos los niveles de los factores, denominado media global.
* ***τi***: Efecto medio producido por el tipo de tabaco *i*,  (∑iτi = 0)
* ***βj***: Efecto medio producido por los distintos sexos *j*,  (∑j βj  = 0)
* $γ\_{k}: $Efecto medio producido por las distintas edades *k*, (∑k $γ\_{k}$  = 0)
* ***(τβ)ij***: Efecto medio producido por la interacción entre el tipo de tabaco i y el sexo j,   (∑i (τβ)ij = ∑j (τβ)ij = 0)
* $(τγ)\_{ik}$: Efecto medio producido por la interacción entre el tipo de tabaco i y la edad k,   (∑i $(τγ)\_{ik}$= ∑k $(τγ)\_{ik}$= 0)
* $(βγ)\_{jk}$: Efecto medio producido por la interacción entre el sexo j y la edad k,   (∑j $(βγ)\_{jk}$= ∑k $(βγ)\_{jk}$= 0)
* $(τβγ)\_{ijk}: $Efecto medio producido por la interacción entre el tipo de tabaco i , el sexo j y la edad k,   (∑i $(τβγ)\_{ijk}$= ∑j $(τβγ)\_{ijk}$=∑k $(τβγ)\_{ijk}$ 0)
* ***uijkl***: Vv aa. Independientes con distribución N(0,σ).
* **Variable respuesta: *Nivel de nicotina***
* **Factor:** ***Tipo de tabaco*** (dos niveles)
* **Factor:** ***Sexo*** (dos niveles)
* **Factor:** ***Edad*** ( tres niveles)
* Los tres factores de **efectos fijos.**
* **Tamaño del experimento:** Número total de observaciones (24).

Para realizar este supuesto en *R* vamos a introducir los datos en un archivo de texto, donde:

* **Variable respuesta: Nivel\_nicotina**
* **Factor**: **sexo**que tiene dos niveles. Es un factor de **efectos fijos** ya que viene decidido que niveles concretos se van a utilizar.
* **Factor**: ***Tabaco*** con dos niveles y de **efectos fijos**.
* **Factor**: ***Edades*** con tres niveles y de **efectos fijos**.
* **Tamaño del experimento:** Número total de observaciones (24).

**Nivel\_nicotina Sexo Tabaco Edades**

**110 Hombres Rubio Niños**

**360 Hombres Negro Niños**

**230 Mujeres Rubio Niños**

**141 Mujeres Negro Niños**

**240 Hombres Rubio Niños**

**125 Hombres Negro Niños**

**219 Mujeres Rubio Niños**

**123 Mujeres Negro Niños**

**112 Hombres Rubio Jóvenes**

**252 Hombres Negro Jóvenes**

**655 Mujeres Rubio Jóvenes**

**873 Mujeres Negro Jóvenes**

**239 Hombres Rubio Jóvenes**

**455 Hombres Negro Jóvenes**

**432 Mujeres Rubio Jóvenes**

**256 Mujeres Negro Jóvenes**

**652 Hombres Rubio Adultos**

**354 Hombres Negro Adultos**

**653 Mujeres Rubio Adultos**

**198 Mujeres Negro Adultos**

**451 Hombres Rubio Adultos**

**701 Hombres Negro Adultos**

**259 Mujeres Rubio Adultos**

**343 Mujeres Negro Adultos**

> setwd("C:/Users/Usuario/Desktop/Datos") \*directorio de trabajo donde estén los datos

> propuesto9<-read.table("propuesto8.txt", header = TRUE)

> propuesto9

 Nivel\_nicotina Sexo Tabaco Edades

1 110 Hombres Rubio Niños

2 360 Hombres Negro Niños

3 230 Mujeres Rubio Niños

4 141 Mujeres Negro Niños

5 240 Hombres Rubio Niños

6 125 Hombres Negro Niños

7 219 Mujeres Rubio Niños

8 123 Mujeres Negro Niños

9 112 Hombres Rubio Jóvenes

10 252 Hombres Negro Jóvenes

11 655 Mujeres Rubio Jóvenes

12 873 Mujeres Negro Jóvenes

13 239 Hombres Rubio Jóvenes

14 455 Hombres Negro Jóvenes

15 432 Mujeres Rubio Jóvenes

16 256 Mujeres Negro Jóvenes

17 652 Hombres Rubio Adultos

18 354 Hombres Negro Adultos

19 653 Mujeres Rubio Adultos

20 198 Mujeres Negro Adultos

21 451 Hombres Rubio Adultos

22 701 Hombres Negro Adultos

23 259 Mujeres Rubio Adultos

24 343 Mujeres Negro Adultos

A continuación debemos transformar tanto la columna de los tratamiento como la de los bloques en un factor para podemos realizar los cálculos posteriores adecuadamente.

> propuesto9$Sexo <- factor(propuesto9$Sexo)

> propuesto9$Sexo

 [1] Hombres Hombres Mujeres Mujeres Hombres Hombres Mujeres Mujeres Hombres

[10] Hombres Mujeres Mujeres Hombres Hombres Mujeres Mujeres Hombres Hombres

[19] Mujeres Mujeres Hombres Hombres Mujeres Mujeres

Levels: Hombres Mujeres

> propuesto9$Tabaco <- factor(propuesto9$Tabaco)

> propuesto9$Tabaco

 [1] Rubio Negro Rubio Negro Rubio Negro Rubio Negro Rubio Negro Rubio Negro

[13] Rubio Negro Rubio Negro Rubio Negro Rubio Negro Rubio Negro Rubio Negro

Levels: Negro Rubio

> propuesto9$Edades <- factor(propuesto9$Edades)

> propuesto9$Edades

 [1] Niños Niños Niños Niños Niños Niños Niños Niños Jóvenes

[10] Jóvenes Jóvenes Jóvenes Jóvenes Jóvenes Jóvenes Jóvenes Adultos Adultos

[19] Adultos Adultos Adultos Adultos Adultos Adultos

Levels: Adultos Jóvenes Niños

**2. Contrastar la hipótesis nula de no interacción entre los factores. Adecuar el modelo al resultado de las interacciones y contrastar los efectos principales.**

Primero vamos a obtener la tabla ANOVA

> mod <- aov(Nivel\_nicotina~ Sexo + Tabaco + Edades+ Sexo \* Tabaco + Sexo \* Edades + Tabaco \* Edades + Sexo \* Tabaco \* Edades , data = propuesto9 )

> mod

Call:

 aov(formula = Nivel\_nicotina ~ Sexo + Tabaco + Edades + Sexo \*

 Tabaco + Sexo \* Edades + Tabaco \* Edades + Sexo \* Tabaco \*

 Edades, data = propuesto9)

Terms:

 Sexo Tabaco Edades Sexo:Tabaco Sexo:Edades

Sum of Squares 4565.0 210.0 306192.2 38160.4 227044.1

Deg. of Freedom 1 1 2 1 2

 Tabaco:Edades Sexo:Tabaco:Edades Residuals

Sum of Squares 41848.1 5.3 448698.5

Deg. of Freedom 2 2 12

Residual standard error: 193.3689

Estimated effects may be unbalanced

> summary(mod)

 Df Sum Sq Mean Sq F value Pr(>F)

Sexo 1 4565 4565 0.122 0.7328

Tabaco 1 210 210 0.006 0.9415

Edades 2 306192 153096 4.094 0.0441 \*

Sexo:Tabaco 1 38160 38160 1.021 0.3323

Sexo:Edades 2 227044 113522 3.036 0.0857 .

Tabaco:Edades 2 41848 20924 0.560 0.5857

Sexo:Tabaco:Edades 2 5 3 0.000 0.9999

Residuals 12 448698 37392

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

El único efecto significativo son las distintas edades.

Al no ser significativa ninguna de las interacciones, realizamos de nuevo el ANOVA sufrimiento la interacción entre los tres factores

> mod1 <- aov(Nivel\_nicotina~ Sexo + Tabaco + Edades+ Sexo \* Tabaco + Sexo \* Edades + Tabaco \* Edades, data = propuesto9 )

> mod1

Call:

 aov(formula = Nivel\_nicotina ~ Sexo + Tabaco + Edades + Sexo \*

 Tabaco + Sexo \* Edades + Tabaco \* Edades, data = propuesto9)

Terms:

 Sexo Tabaco Edades Sexo:Tabaco Sexo:Edades

Sum of Squares 4565.0 210.0 306192.2 38160.4 227044.1

Deg. of Freedom 1 1 2 1 2

 Tabaco:Edades Residuals

Sum of Squares 41848.1 448703.7

Deg. of Freedom 2 14

Residual standard error: 179.0259

Estimated effects may be unbalanced

> summary(mod1)

 Df Sum Sq Mean Sq F value Pr(>F)

Sexo 1 4565 4565 0.142 0.7115

Tabaco 1 210 210 0.007 0.9366

Edades 2 306192 153096 4.777 0.0262 \*

Sexo:Tabaco 1 38160 38160 1.191 0.2936

Sexo:Edades 2 227044 113522 3.542 0.0569 .

Tabaco:Edades 2 41848 20924 0.653 0.5357

Residuals 14 448704 32050

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

De nuevo el único efecto significativo es la Edad. Realizamos de nuevo el ANOVA suprimiendo la interacción de orden 2 Sexo\*Tabaco

> mod2 <- aov(Nivel\_nicotina~ Sexo + Tabaco + Edades+ Sexo \* Edades + Tabaco \* Edades, data = propuesto9 )

> mod2

Call:

 aov(formula = Nivel\_nicotina ~ Sexo + Tabaco + Edades + Sexo \*

 Edades + Tabaco \* Edades, data = propuesto9)

Terms:

 Sexo Tabaco Edades Sexo:Edades Tabaco:Edades Residuals

Sum of Squares 4565.0 210.0 306192.2 227044.1 41848.1 486864.1

Deg. of Freedom 1 1 2 2 2 15

Residual standard error: 180.16

Estimated effects may be unbalanced

> summary(mod2)

 Df Sum Sq Mean Sq F value Pr(>F)

Sexo 1 4565 4565 0.141 0.7129

Tabaco 1 210 210 0.006 0.9369

Edades 2 306192 153096 4.717 0.0258 \*

Sexo:Edades 2 227044 113522 3.498 0.0567 .

Tabaco:Edades 2 41848 20924 0.645 0.5388

Residuals 15 486864 32458

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

De nuevo el mismo resultado, suprimimos la iteración Tabaco\*Edades

> mod3 <- aov(Nivel\_nicotina~ Sexo + Tabaco + Edades+ Sexo \* Edades, data = propuesto9 )

> mod3

Call:

 aov(formula = Nivel\_nicotina ~ Sexo + Tabaco + Edades + Sexo \*

 Edades, data = propuesto9)

Terms:

 Sexo Tabaco Edades Sexo:Edades Residuals

Sum of Squares 4565.0 210.0 306192.2 227044.1 528712.2

Deg. of Freedom 1 1 2 2 17

Residual standard error: 176.354

Estimated effects may be unbalanced

> summary(mod3)

 Df Sum Sq Mean Sq F value Pr(>F)

Sexo 1 4565 4565 0.147 0.7064

Tabaco 1 210 210 0.007 0.9355

Edades 2 306192 153096 4.923 0.0206 \*

Sexo:Edades 2 227044 113522 3.650 0.0480 \*

Residuals 17 528712 31101

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Son significativos el efecto de la Edad y el efecto de la interacción del Sexo por la Edad

**3. ¿Hay diferencias significativas en el nivel de nicotina en las distintas edades? ¿En qué edad el nivel de nicotina es mayor?**

**¿Hay diferencias significativas en el nivel de nicotina en las distintas edades?**

> summary(mod)

 Df Sum Sq Mean Sq F value Pr(>F)

Sexo 1 4565 4565 0.122 0.7328

Tabaco 1 210 210 0.006 0.9415

Edades 2 306192 153096 4.094 0.0441 \*

Sexo:Tabaco 1 38160 38160 1.021 0.3323

Sexo:Edades 2 227044 113522 3.036 0.0857 .

Tabaco:Edades 2 41848 20924 0.560 0.5857

Sexo:Tabaco:Edades 2 5 3 0.000 0.9999

Residuals 12 448698 37392

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**Si hay diferencias significativas del nivel de nicotina en las distintas edades (P-valor es 0.0441)**

**¿En qué edad el nivel de nicotina es mayor?**

Realizamos un contraste de comparaciones múltiples. Por ejemplo LSD

Recordar que hay que cargar e instalar el paquete agricolae

> library(agricolae)

> LSD.test(mod3,"Edades", p.adj="bonferroni", console=TRUE)

Study: mod3 ~ "Edades"

LSD t Test for Nivel\_nicotina

P value adjustment method: bonferroni

Mean Square Error: 31100.72

Edades, means and individual ( 95 %) CI

 Nivel\_nicotina std r LCL UCL Min Max

Adultos 451.375 194.80755 8 319.82686 582.9231 198 701

Jóvenes 409.250 251.74349 8 277.70186 540.7981 112 873

Niños 193.500 85.57202 8 61.95186 325.0481 110 360

Alpha: 0.05 ; DF Error: 17

Critical Value of t: 2.654996

Minimum Significant Difference: 234.1095

Treatments with the same letter are not significantly different.

 Nivel\_nicotina groups

Adultos 451.375 a

Jóvenes 409.250 ab

Niños 193.500 b

El nivel de nicotina es mayor en los Adultos (451.375 miligramos por mililitro)**.**

**4.** **¿El tipo de tabaco es un factor determinante en el nivel de nicotina?**

> summary(mod)

 Df Sum Sq Mean Sq F value Pr(>F)

Sexo 1 4565 4565 0.122 0.7328

Tabaco 1 210 210 0.006 0.9415

Edades 2 306192 153096 4.094 0.0441 \*

Sexo:Tabaco 1 38160 38160 1.021 0.3323

Sexo:Edades 2 227044 113522 3.036 0.0857 .

Tabaco:Edades 2 41848 20924 0.560 0.5857

Sexo:Tabaco:Edades 2 5 3 0.000 0.9999

Residuals 12 448698 37392

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

El tipo de tabaco no es determinante en el nivel de nicotina (P-valor = 0.9415)

**5. Comparar el nivel medio de nicotina entre las mujeres y los hombres. ¿Se detectan diferencias significativas?**

No hay diferencias significativas en el nivel de nicotina entre las mujeres y los hombres (P-valor = 0.7328)

Apliquemos un contraste de comparaciones múltiples para ver donde es mayor

> SNK.test(mod3,"Sexo", console=TRUE)

Study: mod3 ~ "Sexo"

Student Newman Keuls Test

for Nivel\_nicotina

Mean Square Error: 31100.72

Sexo, means

 Nivel\_nicotina std r Min Max

Hombres 337.5833 198.3548 12 110 701

Mujeres 365.1667 239.1971 12 123 873

Alpha: 0.05 ; DF Error: 17

Critical Range

 2

151.8987

Means with the same letter are not significantly different.

 Nivel\_nicotina groups

Mujeres 365.1667 a

Hombres 337.5833 a

Las mujeres y los hombres forman un único grupo donde no se aprecian diferencias significativas. Es mayor el nivel de nicotina entre las mujeres (365.1667 miligramos por mililitro)

**Editor de R**

setwd("C:/Users/Usuario/Desktop/Datos")

propuesto9<-read.table("propuesto9.txt", header = TRUE)

propuesto9

propuesto9$Sexo <- factor(propuesto9$Sexo)

propuesto9$Sexo

propuesto9$Tabaco <- factor(propuesto9$Tabaco)

propuesto9$Tabaco

propuesto9$Edades <- factor(propuesto9$Edades)

propuesto9$Edades

mod <- aov(Nivel\_nicotina~ Sexo + Tabaco + Edades+ Sexo \* Tabaco + Sexo \* Edades + Tabaco \* Edades + Sexo \* Tabaco \* Edades , data = propuesto9 )

mod

summary(mod)

mod1 <- aov(Nivel\_nicotina~ Sexo + Tabaco + Edades+ Sexo \* Tabaco + Sexo \* Edades + Tabaco \* Edades, data = propuesto9 )

mod1

summary(mod1)

mod3 <- aov(Nivel\_nicotina~ Sexo + Tabaco + Edades+ Sexo \* Edades, data = propuesto9 )

mod3

summary(mod3)

library(agricolae)

LSD.test(mod3,"Edades", p.adj="bonferroni", console=TRUE)

SNK.test(mod3,"Sexo", console=TRUE)