**Ejercicio Propuesto 3 (Resuelto)**

**Se realiza un estudio para comprobar la efectividad en el retraso del crecimiento de bacterias utilizando cuatro soluciones diferentes para lavar los envases de la leche. El análisis se realiza en el laboratorio y sólo se pueden realizar seis pruebas en un mismo día. Como los días son una fuente de variabilidad potencial, el investigador decide utilizar un diseño aleatorizado por bloques, pero al recopilar las observaciones durante seis días no ha sido posible aplicar todos los tratamientos en cada día, sino que sólo se han podido aplicar dos de las cuatro soluciones cada día.**

|  |  |
| --- | --- |
|  | **Días** |
| **Soluciones** | **1** | **2** | **3** | **4** | **5** | **6** |
| **1** | 12 | 24 | 31 |  |  |  |
| **2** | 21 |  |  |  | 20 | 21 |
| **3** |  |  | 19 | 18 |  | 19 |
| **4** |  | 15 |  | 19 | 47 |  |

**Para un nivel de significación del 5%:**

1. **¿Son las diferentes soluciones fuente de variación? ¿Y los días?**
2. **Utilizando el método de Newman-Keuls, ¿qué tipo de solución produce un mayor efecto en el crecimiento de bacterias?**

**Solución:**

Este experimento se modeliza mediante un diseño de bloques incompletos balanceados, siendo ambos factores de efectos fijos.

Para realizar este supuesto en RStudio debemos introducir primero los datos de forma correcta. Podemos introducir los datos directamente en RStudio de forma manual o introducirlos previamente en un archivo de texto o Excel y leerlos en RStudio. En este caso lo hacemos en un archivo de texto: En el menú principal de **RStudio** elegimos **Session/Set Working Directory/Choose Directory** o bien con las teclas **Ctrl+Shift+H**, seleccionamos el directorio de trabajo donde están los datos. En ese directorio elegimos el fichero "supuesto4.txt y se muestran los datos en la siguiente ventana de RStudio



Para cargar los datos utilizamos la función “read.table” indicando el nombre del archivo (que debe de estar en el directorio de trabajo) e indicando además que tiene cabecera.

propuesto3 <- read.table("propuesto3.txt", header = TRUE)

propuesto3

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

propuesto3$soluciones = factor(propuesto3$soluciones)

propuesto3$días = factor(propuesto3$días)

1. **¿Son las diferentes soluciones fuente de variación? ¿Y los días?**

Para poder analizar los datos mediante un diseño BIB debemos instalar y cargar dos paquetes de R especializados en este tipo de diseños:

library(daewr)
library(AlgDesign)

La función “BIBsize(t , k)” de la librería daewr nos permite saber si el diseño puede realizarse.  Calcula los parámetros del diseño donde: t = número de niveles del factor tratamiento; k = número de tratamientos por bloque. Ejecutamos:

BIBsize(t = 4 , k = 3)



 **1. Realizaremos el análisis evaluando primero el efecto de los tratamientos y después el de los bloques utilizando dos funciones diferentes.**

**Para evaluar el efecto de los tratamientos**, la suma de cuadrados de tratamientos debe ajustarse por bloques, por lo tanto primero se introducen los bloques y después los tratamientos.

mod1 <- aov(y ~ días + soluciones, data = propuesto3)

mod1

summary(mod1)



**Para evaluar el efecto de los bloques**, la suma de cuadrados de bloques debe ajustarse por los tratamientos, por lo tanto primero se introducen los tratamientos y después los bloques:

mod2 <- aov(y ~ soluciones + días, data = propuesto3)

mod2

summary(mod2)



Observamos como ni el tipo de solución (p-valor=0.834>0.05, no podemos rechazar la hipótesis nula) ni los diferentes días (p-valor=0.720>0.05) influyen significativamente en el crecimiento de bacterias

**También podemos realizar el análisis evaluando tanto para los tratamientos como para los bloques ejecutando solo una función.**

**Para ello,** Para ello necesitamos instalar y cargar el paquete “car”, mediante la función “install.packages” para instalarlo y “library” para cargarlo.

install.packages("car")

library(car)

library(carData)

car::Anova(mod1, type="III")



Observamos cómo ni el tipo de solución (p-valor=0.818>0.05, no podemos rechazar la hipótesis nula) ni los diferentes días (p-valor=0.7118>0.05) influyen significativamente en el crecimiento de bacterias.

1. **Utilizando el método de Newman-Keuls, ¿qué tipo de solución produce un mayor efecto en el crecimiento de bacterias?**

**contraste <- SNK.test(mod2, "soluciones", console = TRUE, main = "Contraste de Newman-Keuls")**

1. ****

Solo hay un subgrupo (letra a), coincidiendo con la conclusión obtenida en el apartado anterior (sin diferencia significativa entre las soluciones). También se observa que la solución donde la cantidad de bacterias es mayor en la cuarta solución (27 de media) y es menor en la tercera solución (18.67 de media).