**Ejercicio Propuesto 2 (Resuelto)**

**Un laboratorio de reciclaje controla la calidad de los plásticos utilizados en bolsas. Se desea contrastar si existe variabilidad en la calidad de los plásticos que hay en el mercado. Para ello, se eligen al azar cuatro plásticos y se les somete a una prueba para medir el grado de resistencia a la degradación ambiental. De cada plástico elegido se han seleccionado ocho muestras y los resultados de la variable que mide la resistencia son los de la tabla adjunta.**

|  |  |
| --- | --- |
| Calidadplásticos | Resistencia |
| Plástico A | **135** | **175** | **97** | **169** | **213** | **171** | **115** | **143** |
| Plástico B | **275** | **170** | **154** | **133** | **219** | **187** | **220** | **185** |
| Plástico C | **169** | **239** | **184** | **222** | **253** | **179** | **280** | **193** |
| Plástico D | **115** | **105** | **93** | **85** | **120** | **74** | **87** | **63** |

**Figura 1: Tabla de datos del Ejercicio Propuesto2.doc**

**¿Qué conclusiones se deducen de este experimento?**

#### ****Solución:****

Los cuatro tipos de plásticos analizados corresponden a una selección aleatoria de 4 conjuntos de observaciones extraídas aleatoriamente del total de diferentes tipos de plásticos que hay en el mercado, entre los cuales debemos observar si existen o no diferencias significativas. Nos encontramos por tanto ante un diseño unifactorial completamente aleatorio con efectos aleatorios.

En este modelo, se supone que las variables ***τi*** son variables aleatorias normales independientes con media 0 y varianza común $σ\_{τ}^{2}.$

La mecánica del Análisis de la Varianza es la misma que en el modelo de efectos fijos. En este modelo, carece de sentido probar la hipótesis que se refiere a los efectos de los tratamientos individuales. Si las medias poblacionales en el conjunto mayor son iguales, no variarán los efectos del tratamiento τi, es decir, $σ\_{τ}^{2}$ . Así en el modelo de efectos aleatorios, la hipótesis de medias iguales se contrasta considerando:

$ H\_{0}≡ σ\_{T}^{2} $= 0 vs $ H\_{1}≡ σ\_{T}^{2} \ne $ 0

Si no se rechaza ***H0***, significa que no hay variedad en los efectos de los tratamientos

El problema planteado se modeliza a través de un **diseño unifactorial totalmente aleatorizado de efectos aleatorios equilibrado.**

* **Variable respuesta:** ***Resistencia a la degradación ambiental.***
* **Factor**: **Tipo de plástico.**
* **Modelo equilibrado:** Cada uno de los niveles del factor tienen el mismo número de observaciones.
* **Tamaño del experimento:** Número total de observaciones (40 unidades experimentales).

Para realizar este supuesto en R debemos introducir primero los datos de forma correcta. Podemos introducir los datos directamente en R de forma manual o introducirlos previamente en un archivo de texto o Excel y leerlos en R.

En este caso lo hacemos en un archivo de texto:

Resistencia Plástico

135 P1

175 P1

97 P1

169 P1

213 P1

171 P1

115 P1

143 P1

275 P2

170 P2

154 P2

133 P2

219 P2

187 P2

220 P2

185 P2

169 P3

239 P3

184 P3

222 P3

253 P3

179 P3

280 P3

193 P3

115 P4

120 P4

105 P4

74 P4

93 P4

87 P4

85 P4

63 P4

Figura 2: Tabla de datos del ejercicio propuesto2.txt

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.

***Nota****: La ruta hasta llegar al fichero varía en función del ordenador. Utilizar la  orden****setwd()*** *para situarse en el directorio de trabajo*

> setwd("C:/Users/Usuario.Usuario-PC/Desktop/Datos")

C:\Users\Usuario.Usuario-PC\Desktop\Datos
> propuesto2<- read.table("propuesto2.txt", header = TRUE)

 Resistencia Plástico

1 135 p1

2 175 p1

3 97 p1

4 169 p1

5 213 p1

6 171 p1

7 115 p1

8 143 p1

9 275 p2

10 170 p2

11 154 p2

12 133 p2

13 219 p2

14 187 p2

15 220 p2

16 185 p2

17 169 p3

18 239 p3

19 184 p3

20 222 p3

21 253 p3

22 179 p3

23 280 p3

24 193 p3

25 115 p4

26 120 p4

27 105 p4

28 74 p4

29 93 p4

30 87 p4

31 85 p4

32 63 p4

Debemos transformar la variable referente a los niveles del factor como factor para poder hacer los cálculos de forma adecuada:

**> propuesto2$Plástco <- factor(propuesto2$Plástico)**

**> propuesto2$Plástco**

 **[1] p1 p1 p1 p1 p1 p1 p1 p1 p2 p2 p2 p2 p2 p2 p2 p2 p3 p3 p3 p3 p3 p3 p3 p3 p4**

**[26] p4 p4 p4 p4 p4 p4 p4**

**Levels: p1 p2 p3 p4**

Para calcular la tabla ANOVA primero hacemos uso de la función “aov” de la siguiente forma:

**> mod <- aov(Resistencia ~ Plástico, data = propuesto2)**

donde:

* **Resistencia**: Nombre de la columna de las observaciones.
* **Plástico:** Nombre de la columna en la que están representados los tratamientos.
* data= data.frame en el que están guardados los datos.

**> mod <- aov(Resistencia ~ Plástico, data = propuesto2)**

**> mod**

**Call:**

 **aov(formula = Resistencia ~ Plástico, data = propuesto2)**

**Terms:**

 **Plástico Residuals**

**Sum of Squares 69072.12 37410.75**

**Deg. of Freedom 3 28**

**Residual standard error: 36.55268**

**Estimated effects may be unbalanced**

y posteriormente mostramos un resumen de los resultados con la función “summary” (verdadera tabla ANOVA):

**> summary(mod)**

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

**Plástico 3 69072 23024 17.23 1.55e-06 \*\*\***

**Residuals 28 37411 1336**

**---**

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

Esta tabla muestra los resultados del contraste planteado. El valor del estadístico de contraste es igual a 17.23 que deja a la derecha un p-valor de 1.55e-06, rechazando la Hipótesis nula tanto a un nivel de significación del 5% como del 1%. Podemos concluir que los datos muestran evidencias de variabilidad en la resistencia para la degradación ambiental según el tipo de plástico empleado en la fabricación de la bolsa.

En el modelo de efectos aleatorios no se necesitan llevar a cabo más contrastes incluso aunque la hipótesis nula sea rechazada. Es decir, en el caso de rechazar H0 no hay que realizar comparaciones múltiples para comprobar que medias son distintas, ya que el propósito del experimento es hacer un planteamiento general relativo a las poblaciones de las que se extraen las muestras.

En este caso, ***R*** no tiene ninguna función que nos permita calcular la varianza de tratamientos, por lo que tenemos que calcularla a mano:

En el modelo de efectos aleatorios las variables ***τi*** y ***uij*** son independientes, por lo tanto la varianza de cualquier observación de la muestra, es decir, la varianza total, es la suma de las varianzas

$$σ\_{T}^{2}=σ\_{τ}^{2}+σ^{2}$$

* **Varianza de los residual:** $\hat{σ}^{2}=S\_{R}^{2}=37411$
* **Varianza entre los factores**

$$\hat{σ}\_{τ}^{2}=\frac{N\*\left(I-1\right)}{N^{2}-\sum\_{i}^{}n\_{i}^{2}}\left(S\_{Tr}^{2}-S\_{R}^{2}\right)=\frac{32\*3}{32^{2}-4\*8^{2}}\left(69072-37411\right)=3957.625$$

* **Varianza Total:** $\hat{σ}\_{T}^{2}=\hat{σ}\_{τ}^{2}+\hat{σ}^{2}=3957.625+37411=41368.625$

Por lo tanto, la varianza total (41368.625) se descompone en una parte atribuible a la diferencia entre los fabricantes (3957.625) y otra procedente de la variabilidad existente dentro de ellos (37411). Comprobamos que en dicha varianza tienen mayor peso la variación dentro los fabricantes, en porcentaje un 90.433 % frente a la variación entre ellos, que representa el 9.566 % del total.

Editor de R

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

propuesto2<-read.table("propuesto2.txt", header = TRUE)

propuesto2

propuesto2$Plástco <- factor(propuesto2$Plástico)

propuesto2$Plástco

mod <- aov(Resistencia ~ Plástico, data = propuesto2)

mod

summary(mod)