

# Hoja 4 (d): Modelos lineales con R

## Estadística Computacional I. Grado en Estadística

Departamento de Estadística e Investigación Operativa. Universidad de Sevilla

### Ejercicio 1

ANOVA de un factor.

El fichero de datos “scores.txt” contiene la puntuación obtenida en una prueba de nivel de inglés para 40 alumnos. Se han considerado 4 academias, de cada una de las cuales han sido seleccionados aleatoriamente 10 alumnos. Se trata de estudiar si existen diferencias significativas entre las puntuaciones medias dependiendo de la academia.

### Ejercicio 2

ANOVA de dos factores.

A fin de investigar el efecto del fármaco Rhitalin sobre los niños hiperactivos se tomó una muestra de 4 niños para cada uno de los cruces de los dos siguientes factores: Tipo de niño (normal e hiperactivo) y medicamento administrado (Placebo y Rhitalin). Para cada niño se midió un índice de actividad.

```
indice<-c(50,45,55,52,67,60,58,65,70,72,68,
          75,51,57,40,55)
niño<- gl(2,8)
levels(niño)<- c("Normal","Hiperactivo")
tratamiento<- gl(2,4,16)
levels(tratamiento)<- c("Placebo","Rhitalin")
data.frame(niño,tratamiento,indice)
```

	niño	tratamiento	indice
## 1	Normal	Placebo	50
## 2	Normal	Placebo	45
## 3	Normal	Placebo	55
## 4	Normal	Placebo	52
## 5	Normal	Rhitalin	67
## 6	Normal	Rhitalin	60
## 7	Normal	Rhitalin	58
## 8	Normal	Rhitalin	65
## 9	Hiperactivo	Placebo	70
## 10	Hiperactivo	Placebo	72
## 11	Hiperactivo	Placebo	68
## 12	Hiperactivo	Placebo	75
## 13	Hiperactivo	Rhitalin	51
## 14	Hiperactivo	Rhitalin	57
## 15	Hiperactivo	Rhitalin	40
## 16	Hiperactivo	Rhitalin	55

### Ejercicio 3

Regresión Lineal Simple.

```
x <- c(18,23,25,35,65,54,34,56,72,
      19,23,42,18,39,37)
#x=Edad
y <-c(202,186,187,180,156,169,174,172,
     153,199,193,174,
     198,183,178)
#y=Máximo de "frecuencia cardíaca"
```

Nube puntos y superponer la recta de mínimos cuadrados.

### Ejercicio 4

Regresión Lineal Múltiple.

```
library(ISLR) #para acceder a Hitters
data(Hitters)
# ?Hitters
```

### Ejercicio 5

Regresión Lineal Múltiple.

El fichero de datos “Advertising.csv” contiene las ventas de un producto en 200 mercados diferentes junto con los presupuestos de publicidad en cada mercado en tres medios: televisión, radio y prensa.

El objetivo es construir un modelo de regresión lineal múltiple para predecir las ventas del producto en función de los gastos en publicidad.

Variables en el archivo: Caso; TV, Radio, Prensa (miles de dólares) y las Ventas (miles de unidades).

#### Apartado a

Incrementar en mil dólares el gasto publicitario en TV conlleva, por término medio, aumentar en  $0.046 \cdot 1000 = 46$  unidades las ventas del producto (suponiendo que el gasto publicitario en Radio no cambia).

Si se incrementa en mil dólares el gasto publicitario en la Radio, cabe esperar que las ventas aumenten en 188 unidades, suponiendo fijo el gasto en TV.

Los gastos publicitarios en TV y Radio explican el 89.72 % de la varianza de las Ventas del producto mediante este modelo.

### Ejercicio 6

Regresión cuadrática.

En 1609 Galileo demostró que la trayectoria de un cuerpo cayendo con una horizontal componente es una parábola. En el curso de ganar conocimiento de este hecho, estableció un experimento que midió dos variables, una altura y una distancia, produciendo los siguientes datos.

```
dist = c(253, 337,395,451,495,534,574)
height = c(100,200,300,450,600,800,1000)
```