

Departamento de Estadística e Investigación Operativa. Universidad de Sevilla  
**Análisis de Datos Multivariantes. Grado en Matemáticas.**

**Hoja 4. Prácticas de Inferencia Multivariante con R**

1. Se considera el fichero **pulmonary** de la librería **ICSNP**:
  - i) Cargar los datos y resumirlos numérica y gráficamente.
  - ii) Analizar la normalidad multivariante.
  - iii) Contrastar la hipótesis nula  $H_0: \underline{\mu}=(0,0,2)^t$ .
  
2. Para ilustrar las regiones de confianza para  $\underline{\mu}$ :
  - i) Generar y dibujar una muestra de tamaño 20 de una normal bivalente estándar con  $\rho=0.6$ .
  - ii) Dibujar regiones de confianza con la ayuda de la función **ellipse** de la librería **ellipse**.
  
3. Se consideran dos distribuciones normales multivariantes con parámetros respectivos  $\underline{\mu}_1=(0,0,0,0)^t$ ,  $\underline{\mu}_2=(0.5, 0.5, 0.5, 0.5)^t$ ,  $\Sigma_1=\Sigma_2=\text{diag}(1:4)$ .
  - i) Construir una matriz que contenga sendas muestras de tamaños 20 y 30 y guardar la identificación de la población correspondiente. Resumir los datos generados. (**set.seed(12345)**)
  - ii) Analizar la normalidad multivariante de ambas muestras y comparar las matrices de dispersión.
  - iii) Realizar el contrastes de hipótesis bilateral  $H_0: \underline{\mu}_1 - \underline{\mu}_2 = (-0.5, -0.5, -0.5, -0.5)^t$ .
  
4. Se quiere analizar la efectividad de un tratamiento para niños con problemas de rendimiento escolar. El fichero “Rendiescolar.dat” contiene cinco variables, de las cuales se van a tener en cuenta las variables 1, 4 y 5:

Grupo: 1 Tratamiento, 2 Control.  
Puntulec: Puntuación en un test de lectura.  
Puntuari: Puntuación en un test de aritmética.

  - i) Leer los datos en R.
  - ii) Dibujar la nube de puntos identificando los centroides.
  - iii) ¿Es efectivo el tratamiento?
  
5. El **data.frame** **painters** de la librería **MASS** contiene las puntuaciones que el crítico de arte de Piles asignó a 54 pintores clásico en cuatro características.
  - i) Leer los datos, y seleccionar los pintores de las escuelas Renacimiento y Veneciana.
  - ii) Obtener el gráfico de cajas y bigotes de cada característica según ambas escuelas.
  - iii) Generar las nubes de puntos de todos los pares de variables, distinguiendo las escuelas artísticas con colores diferentes e incluyendo los nombres de los artistas.

- iv) Realizar la comparación de medias univariante para cada característica.
- v) Contrastar la igualdad de vectores de medias.

**6.** Leer el fichero de R *iris*, un conjunto de datos clásicos del Análisis Multivariante.

- i) Analizar la normalidad para cada tipo de flor.
- ii) Contrastar la igualdad de vectores de medias.

**7.** El fichero “Simetria.dat” contiene las siguientes variables:

NIVEL=Nivel socioeconómico familiar (alto, medio-bajo)

LECTUPRI=Test de lectura en primavera

LECTUOTO=Test de lectura tras el verano

ARITPRI=Test de aritmética en primavera

ARITOTO=Test de aritmética tras verano

- i) Leer los datos en R y representar gráficamente y de forma simultánea los perfiles de las cuatro variables.
- ii) Realizar una transformación lineal apropiada para el contraste  $H_0: \mu_1=\mu_3$  y  $\mu_2=\mu_4$ . Representar gráficamente las variables transformadas.
- iii) Realizar e interpretar dicho contraste, llamado contraste de simetría.
- iv) ¿Se ha producido la misma disminución media en ambos niveles socioeconómicos utilizando de forma conjunta las puntuaciones en Lectura y Aritmética?

**8.** El fichero “InflamabilidadNoCab.txt” contiene una medida de la inflamabilidad de unos terrenos del Parque Natural de Doñana donde no pastaron cabras, recogidas durante 6 meses consecutivos.

- i) Leer los datos en R y representar gráficamente los perfiles de los distintos terrenos, superponiendo el perfil medio.
- ii) ¿Se puede considerar que la inflamabilidad media se mantiene constante?

**9.** El fichero “InflamabilidadCab.txt” contiene datos similares a los del problema anterior, pero referidos a terrenos donde sí pastaron cabras.

- i) Leer los datos y representarlos gráficamente.
- ii) Dibujar los perfiles medios para ambos tipos de terrenos.
- iii) Contrastar el paralelismo de perfiles.

**10.** Realizar un análisis de perfiles para los datos del fichero “Paralelismo.txt”.