

---

# Prueba de Evaluación Continua – 1

---

Análisis Multivariante de Datos (Curso 2019-20)

## Descripción de datos multivariantes, análisis gráfico y detección de valores atípicos

### **Introducción**

Durante las primeras semanas de curso nos hemos iniciado en el análisis multivariante. Tras una introducción al tema y un repaso de conceptos de álgebra matricial, hemos estudiado diversas formas de describir los datos multivariantes y representarlos gráficamente. También hemos comenzado el estudio de los valores atípicos (*outliers*). En esta PEC se trabaja la aplicación de los conocimientos adquiridos durante estas primeras semanas de curso y se propone la ampliación del estudio de los valores atípicos.

### **Conjunto de datos**

Se proporciona a los estudiantes un conjunto de datos multivariante al que nos referiremos como AM20. Este conjunto de datos será el que los estudiantes tendrán que analizar.

### **Objetivos**

A partir del conjunto de datos AM20 se pide realizar las siguientes tareas:

1. Describir el conjunto de datos
  - a. Usar medidas de centralización y analizarlas
  - b. Usar medidas de variabilidad y analizarlas
  - c. Estudiar las dependencias lineales entre las variables
2. Representar gráficamente los datos
3. Estudiar la existencia de valores atípicos
  - a. Identificar y aplicar diversas (al menos dos) técnicas
  - b. Comparar los resultados de las técnicas estudiadas
  - c. Analizar de forma crítica los resultados

### **Evaluación**

Para evaluar la PEC se tendrán en cuenta los siguientes conceptos:

- 50% Aplicación de conocimientos adquiridos
- 20% Capacidad para usar R para resolver el problema planteado
- 30% Originalidad (especialmente para la Tarea 3), corrección y calidad de la solución y la documentación

### **¿Porqué evaluamos así?**

En esta PEC se da un peso del 50% a la aplicación de conocimientos adquiridos durante las primeras semanas de curso, es decir, si el estudiante demuestra saber aplicar los conocimientos relativos a la descripción de datos multivariantes, su análisis gráfico y la detección de valores atípicos habrá superado la PEC. (Para obtener la máxima calificación en este apartado no es necesario usar R).

Dado que el planteamiento de la asignatura es aplicado, se valora hasta un 20% la capacidad de usar R (u otro software convenido, a priori con el profesor) para resolver el problema planteado. No es necesario tener un nivel elevado de programación en R sino demostrar que se saben usar algunas de sus librerías para tratar datos multivariantes y describirlos, analizarlos gráficamente y detectar valores atípicos.

Finalmente, un 30% de la calificación recae en la originalidad, corrección y calidad de la solución propuesta y la documentación entregada por el estudiante. Dado que se está cursando un Máster, se espera que el estudiante demuestre capacidad de resolver problemas de forma original y sea crítico con la calidad y corrección de la solución y la documentación. Se recomienda usar LaTeX para la documentación (aunque no es obligatorio).

### **Fechas importantes**

- |                                      |                     |
|--------------------------------------|---------------------|
| • Publicación del Enunciado:         | 01 de Abril de 2019 |
| • Entrega (Fecha límite):            | 22 de Abril de 2019 |
| • Publicación de la Solución:        | 29 de Abril de 2019 |
| • Publicación de las calificaciones: | 06 de Mayo de 2019  |