



**UNIVERSITAS**  
*Miguel Hernández*

## Tema 2. Regresión Logística. Ejemplos.

José L. Sainz-Pardo Auñón

### **TÉCNICAS ESTADÍSTICAS PARA EL APRENDIZAJE II**

Máster Universitario en Estadística Computacional  
y Ciencia de Datos para la Toma de Decisiones.

# 1. Carga y Preparación de los Datos

- Descargar el archivo `logistica.xlsx` que contiene los datos para la regresión logística.
- Leer el archivo Excel en un DataFrame de pandas.
- Visualizar las primeras filas del conjunto de datos.
- Definir las variables independientes  $X_1$ ,  $X_2$ ,  $X_3$  y la variable dependiente  $Y$ .

## 2. Obtención de modelo de Regresión Logística.

- Divide los datos en un conjunto test del 30% y uno de entrenamiento del 70%
- Obtén a partir del conjunto de entrenamiento un modelo de Regresión Logística.
- ¿Cuáles son los coeficientes del modelo?
- ¿Cuál es la función sigmoide obtenida? ¿Y la función logit?
- Entiende e interpreta la función sigmoide y la función logit.

### 3. Tablas de confusión

- Obtén la tabla de confusión utilizando la librería sklearn.
- Obtén la tabla de confusión manualmente.

### 3. Tablas de confusión

- Obtén la tabla de confusión utilizando la librería sklearn.
- Obtén la tabla de confusión manualmente.

## 4. Probabilidades de $Y=1$

- Para cada dato del conjunto test, obtén la probabilidad que que pertenezca a la clase  $Y=1$  de forma 'manual'.
- Para cada dato del conjunto test, obtén la probabilidad que que pertenezca a la clase  $Y=1$  mediante sklearn.
- Según dichas probabilidades, ¿en qué clases clasificarías cada uno de los 5 primeros datos?

## 5. Curva ROC

- Obtén la Curva ROC.
- Analiza e interpreta la curva obtenida.

## 6. Obtención del modelo mediante statsmodels

- Obtén el modelo también mediante la librería statsmodels.
- Obtén un resumen estadístico de dicho modelo.
- ¿Te es útil algún aspecto de este resumen?





**UNIVERSITAS**  
*Miguel Hernández*