

PROBLEMAS PROPUESTOS DE REGRESIÓN MULTINOMIAL
ANÁLISIS ESTADÍSTICO MULTIVARIANTE
GRADO EN CIENCIA E INGENIERÍA DE DATOS

PROBLEMA 1

El fichero **iris** de R, contiene los datos correspondientes a medidas de los pétalos y sépalos de tres variedades de flor de iris (setosa, virginica y versicolor).

Se desea realizar un análisis de Regresión Multinomial con el fin de predecir la variedad de la flor en función de las magnitudes de sus pétalos y sépalos. Se pide:

- 1) Recuperar los datos y realizar un estudio descriptivo previo atendiendo a nuestro objetivo.
- 2) Dividir el conjunto de datos en entrenamiento y prueba (70% entrenamiento, 30% prueba). Tomar semilla 123.
- 3) Con los datos de entrenamiento, obtener el modelo ajustado de Regresión Multinomial usando todos los predictores.
- 4) Con los datos de entrenamiento, aplicar los métodos de selección de regresores para comprobar si el modelo completo es reducible.
- 5) Con el modelo resultante del apartado anterior, obtener medidas de bondad del ajuste e indicar si el modelo es significativo.
- 6) Veamos ahora el problema de Regresión Multinomial como un problema de clasificación. Obtener la clase predicha para los datos del conjunto de prueba.
- 7) Obtener la matriz de confusión para el conjunto de prueba y medir la eficiencia del clasificador con el “accuracy” (número de aciertos en la clasificación dividido entre número de datos totales).