

## LIBRERÍA *R*PART

La función principal de la librería *R*PART es `rpart`, cuya descripción es:

**rpart**(formula, data, weights, subset, na.action = na.rpart, method, model = FALSE, x = FALSE, y = TRUE, parms, control, cost, ...)

La explicación de los argumentos de esta función:

- **formula**: es la expresión que relacionará a la variable de clasificación en función de un conjunto de variables clasificadoras o predictoras. Por ejemplo:  $V1 \sim V2+V3+V4$
- **data**: se debe incluir el nombre del data frame
- **method**: Según la naturaleza de la variable explicada, podrían ser: "class" (variable nominal), "anova" (variable continua), "poisson" (variable discreta ordinal) y "exp" (variable exponencial, problemas de supervivencia)
- **weights**: argumento opcional que en su caso define ponderaciones
- **subset**: argumento opcional que indica si solo un subconjunto de filas del data frame serían utilizados en el ajuste del modelo
- **na.action**: por defecto actúa eliminando los casos en que la variable explicada tiene datos missing, pero mantiene aquellos en los que uno o más variables predictoras tienen valores faltantes.
- **x**: mantiene una copia de la matriz X en el resultado
- **y**: mantiene una copia de la variable explicada en el resultado.
- **parms**: parámetros opcionales según sea el tipo de variable explicada según lo indicado en el argumento method.
- **Control**: listado de opciones, habitualmente presentes en la función `rpart.control`. Un argumento control que utilizaremos será `cp`: umbral de complejidad
- **Cost**: especifica un vector no negativo de costs, uno por cada variable en el modelo

### **plot.rpart**

**ARGUMENTOS OBLIGATORIO**: objeto que representa al árbol

**ARGUMENTOS OPCIONALES**:

- **uniform**: Si es TRUE separación uniforme entre nodos
- **branch**: Desde el valor 0 (ramas en forma de V) a 1 (ramas cuadradas)
- **compress**: Alinea (FALSE) o no (TRUE) los nodos terminales
- **nspc**: espacio extra entre nodos padre e hijos

### **text.rpart**

**ARGUMENTOS OBLIGATORIO**: objeto que representa al árbol

**ARGUMENTOS OPCIONALES**:

- splits: Si TRUE (opción por defecto), las divisiones se etiquetan con el criterio de división
- label: etiquetan nodos
- all: Si TRUE todos los nodos etiquetados
- pretty: entero que refleja abreviaturas de las etiquetas de los factores de división
- use.n: Si TRUE especifica el tamaño muestral de cada clase en el nodo