

## Regresión Logística. Ejercicios

*Abdelmalik Moujahid, Iñaki Inza y Pedro Larrañaga*  
*Departamento de Ciencias de la Computación e Inteligencia Artificial*  
*Universidad del País Vasco–Euskal Herriko Unibertsitatea*

1. Supongamos que se define una la variable  $Y$  que registra la aparición (o no) de alguna enfermedad coronaria (EC) para varones no menores de 50 años y durante un cierto período de observación:  $Y = 1$ , si el sujeto la desarrolló durante el período,  $Y = 0$  si transitó dicho tiempo sin su aparición. Admitamos también que se consideran 3 variables que se miden al comienzo del período y que presumiblemente influyen en el proceso de que determina cuál de los dos desenlaces es el que ocurre. Estas variables son:  $X_1$  es la edad del sujeto,  $X_2$  es el hábito de fumar: 1 si fumaba; 0 en caso contrario, y  $X_3$  la tensión arterial sistólica. Imaginemos que una vez estimados los parámetros, el modelo de regresión logística, escrito en formato *logit*, es el siguiente:

$$\text{logit}(p) = -6,614 + 0,075X_1 + 0,312X_2 + 0,018X_3$$

- Utilizando las nociones de riesgo relativo (RR) y risk odds ratio (ROR) responder a la siguiente pregunta: ¿Cuanto más probable de desarrollar una enfermedad coronaria tiene un sujeto de 65 años, fumador y cuya tensión sistólica sea de 175, que uno de 58 años, no fumador y con una tensión sistólica de 150? Interpretar los resultados.
2. Dado un modelo de regresión logística, demostrar que el parámetro  $C\beta_i$  asociado a una variable no dicotómica  $X_i$  expresa el cambio que se produce en  $\ln(OR)$  cuando  $X_i$  aumenta de  $C$  unidades y el resto de la variables permanecen inalteradas.
  3. Supongamos que para el modelo logístico considerado en el ejercicio 1 tenemos la siguiente información:  
la lejanía del modelo completo (con las 3 variables) es:

$$-2\ln V = 1504,53$$

y la del modelo nulo, igual a:

$$-2\ln V_0 = 1550,40$$

donde  $V_0$  es la verosimilitud del modelo nulo y  $V_1$  es la verosimilitud correspondiente al modelo completo.

- Teniendo en cuenta que el percentil 95 de Chi-cuadrado con 3 grados de libertad es de 7.81, utilizar el test de la razón de verosimilitud para testear el modelo.