

Ayudantía MLE

Sea Y una variable aleatoria con dos posibles resultados: $\mathbb{P}(Y = 1) = p$ y $\mathbb{P}(Y = 0) = 1 - p$. Un modelo logístico expresa p en función de un vector $x \in \mathbb{R}^n$ de variables explicativas mediante:

$$p = \frac{e^{<a,x>+b}}{1 + e^{<a,x>+b}}$$

donde $a \in \mathbb{R}^n$ y $b \in \mathbb{R}$ son parámetros a calibrar a partir de una muestra independiente $(Y_i, x_i) : i = 1, \dots, m$. El método de máxima verosimilitud escoge los parámetros a, b que maximizan la probabilidad $\prod_{i \in I_1} p_i \cdot \prod_{i \in I_0} (1 - p_i)$ de los resultados observados, donde I_1, I_0 denotan las observaciones con $Y_i = 1$ e $Y_i = 0$ respectivamente y $p_i = \frac{e^{<a,x_i>+b}}{1 + e^{<a,x_i>+b}}$.

1. Grafique el negativo de la log-verosimilitud para $a \in [-4, 4]$, y $b \in [-8, 8]$, con $a, b \in \mathbb{R}$ **para los siguientes valores de x e y :**

x	1.1	1.5	3.3	5.5
y	0	0	1	1

Cuadro 1: valores de x e y para gráfica

2. Considere la siguiente tabla de observaciones en que la primera fila representa la cantidad de radiación en un tratamiento de cancer y la segunda fila indica si el resultado del tratamiento fue exitoso o no

Radiación (x)	0.5	1.0	1.5	2.2	2.5	3.1	3.0	4.5	4.0	4.5	4.8	5.0
Resultado (y)	0	0	0	1	0	1	0	1	1	1	1	1

Utilice un método de *newton-raphson* o el paquete *scipy.optimize* para encontrar los parámetros óptimos $(a, b) \in \mathbb{R}^2$. Utilice los valores de a y b encontrados para expresar la probabilidad de éxito p del tratamiento en función de la intensidad de radiación (x) . Grafique esta función.