# 10 - Respuesta binaria

# Descripción del modelo

Se plantea el modelo

$$y_i = p_i + \varepsilon_i, \quad i = 1, \ldots, n,$$

donde

$$p_i = F(\boldsymbol{x}_i \cdot \boldsymbol{\beta}).$$

$F$ es la función logística. Escribimos $\boldsymbol{x}_i$ como una fila $1 \times m$. Frecuentemente $x_{i1} = 1$.

Así, escribiendo $\boldsymbol{\beta}$ como una columna $m \times 1$, su primera componente es el término constante.

# Estimación de parámetros: Máxima Verosimilitud

La verosimilitud para la observación $i$-ésima es:

$$\mathscr{L}_i(\boldsymbol{\beta}) = \begin{cases} p_i, & \text{si} \quad y_i = 1, \\ 1 - p_i, & \text{si} \quad y_i = 0. \end{cases}$$

que puede abreviarse, según es usual, como

$$\mathscr{L}_i(\boldsymbol{\beta}) = p_i^{y_i} \, (1 - p_i)^{(1 - y_i)}.$$

La verosimilitud $\mathscr{L}(\boldsymbol{\beta})$ es el producto de estos $n$ factores.

# Ecuaciones de verosimilitud

El logaritmo $\ell \equiv \ell(\boldsymbol{\beta}) = \log \mathscr{L}(\boldsymbol{\beta})$ es:

$$\ell = \sum_{i=1}^{n} [y_i \log p_i + (1 - y_i) \log(1 - p_i)]$$

Teniendo en cuenta que $p_i = F\left(\sum_{k=1}^{m} x_{ik}\,\beta_k\right)$,

$$\frac{\partial p_i}{\partial \beta_j} = p_i\,(1 - p_i)\,x_{ij},$$

las ecuaciones de verosimilitud son:

$$\frac{\partial \ell}{\partial \beta_j} = \sum_{i=1}^{n} (y_i - p_i)\,x_{ij} = 0, \quad j = 1, \ldots, m.$$

# Ecuaciones de verosimilitud

En notación matricial,

$$\frac{\partial \boldsymbol{p}}{\partial \boldsymbol{\beta}} = \boldsymbol{V} \cdot \boldsymbol{X},$$

donde

$$\boldsymbol{V} = \operatorname{diag}\{\boldsymbol{p}\,(1 - \boldsymbol{p})\}$$

es la matriz diagonal $n \times n$ con elementos $p_i\,(1 - p_i)$ en su diagonal principal.

$$\frac{\partial \ell}{\partial \boldsymbol{\beta}} = \boldsymbol{X}' \cdot (\boldsymbol{y} - \boldsymbol{p}).$$

# Segunda derivada

La segunda derivada de $\ell$ es:

$$\frac{\partial^2 \ell}{\partial \beta_j \, \partial \beta_k} = -\sum_{i=1}^{n} p_i \left(1 - p_i\right) x_{ik} \, x_{ik}$$

En notación matricial:

$$\frac{\partial^2 \ell}{\partial \boldsymbol{\beta} \, \partial \boldsymbol{\beta}'} = -\boldsymbol{X}' \cdot \boldsymbol{V} \cdot \boldsymbol{X}.$$

# Matriz de Información de Fisher

Las derivadas segundas no dependen de las observaciones $y_i$, por tanto, coinciden con sus esperanzas, las componentes $g_{jk}$ de la matriz de información de Fisher.

Así, la matriz de información de Fisher es:

$$\boldsymbol{G} = \boldsymbol{X}' \cdot \boldsymbol{V} \cdot \boldsymbol{X}.$$

# Cálculo de solución: IWLS (o IRLS)

Aproximaciones sucesivas.

A partir de un $\boldsymbol{\beta}(0)$ se calcula una segunda aproximación $\boldsymbol{\beta}(1)$, y así sucesivamente.

Aplicando el método de Newton-Raphson o su variante, el método de *Fisher scoring*, que en este caso coinciden, pues la segunda derivada no depende de las observaciones, se obtienen aproximaciones sucesivas $\boldsymbol{\beta}(r)$, $r \geq 0$.

# Cálculo de solución: IWLS (o IRLS)

En general, dado un $\boldsymbol{\beta} \equiv \boldsymbol{\beta}(r)$,

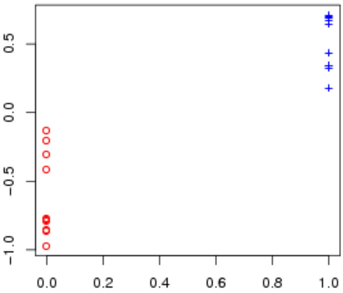
$$\begin{aligned}
\mathbf{p}(r) &= F(\mathbf{X} \cdot \boldsymbol{\beta}), \\
\mathbf{V}(r) &= \operatorname{diag}\{\mathbf{p}(r)(1 - \mathbf{p}(r))\}, \\
\mathbf{G}(r) &= \mathbf{X}' \cdot \mathbf{V}(r) \cdot \mathbf{X}, \\
\boldsymbol{\delta}(r) &= \mathbf{G}^{-1} \cdot \mathbf{X}' \cdot (\mathbf{y} - \mathbf{p}(r)), \\
\boldsymbol{\beta}(r+1) &= \boldsymbol{\beta}(r) + \boldsymbol{\delta}(r).
\end{aligned}$$

# 10 - Respuesta binaria

# Enunciado del teorema. Versión geométrica

Si la nube de puntos $\boldsymbol{x}_i$ correspondientes a los $y_i = 1$ queda <u>completamente separada</u> de la nube de puntos $\boldsymbol{x}_i$ correspondientes a los $y_i = 0$, entonces no tiene solución la ecuación de verosimilitud.

# Ilustración



# A modo de precaución

Es importante tener en cuenta este problema cuando se calcula con datos reales.

Los algoritmos numéricos, en presencia de esta situación, producen resultados impredictibles, pero siempre indeseables.