The logistic regression

Jean-Michel Marin

University of Montpellier
Faculty of Sciences

HAX912X - 2023/2024

Definition

We study the binary logistic model

$Y \in \{0, 1\}$ the response
X explanatory variables possibly containing factors associated with contrasts

We observe $(X_1, Y_1), \ldots, (X_n, Y_n)$

We suppose that

- $Y_i | X_i = x_i \sim \mathscr{B}\left(\frac{\exp(x_i^T \beta)}{1+\exp(x_i^T \beta)}\right)$
- knowing $X_1 = x_1, \ldots, X_n = x_n$ the variables Y_i are independent

Definition

We have

$$\mathbb{P}(Y = 1|X = x) = \frac{\exp(x^T\beta)}{1 + \exp(x^T\beta)} = \mathbb{E}(Y|X = x)$$

$$\mathbb{V}(Y|X = x) = \mathbb{E}(Y|X = x)(1 - \mathbb{E}(Y|X = x)) = \frac{\exp(x^T\beta)}{(1 + \exp(x^T\beta))^2}$$

Definition

We have directly modeled the distribution of $Y|X = x$ via its expectation

This is a generalized linear model associated with the logistic link function

$$\log\left(\frac{\mathbb{E}(Y|X=x)}{1-\mathbb{E}(Y|X=x)}\right) = x^T\beta$$

Estimation de β

The likelihood function is

$$V(\beta) = \prod_{i=1}^{n} \left(\frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)} \right)^{y_i} \left(\frac{1}{1 + \exp(x_i^T \beta)} \right)^{1-y_i}$$

The log-likelihood function is

$$L(\beta) = \sum_{i=1}^{n} \left\{ y_i x_i^T \beta - \log(1 + \exp(x_i^T \beta)) \right\}$$

Estimation de β

$$\frac{\delta L}{\delta \beta}(\beta^*) = 0_d$$

$$\iff \sum_{i=1}^{n} \left\{ y_i x_i - \frac{x_i \exp(x_i^T \beta^*)}{1 + \exp(x_i^T \beta^*)} \right\} = 0_d$$

There is no analytical solution to this system

We use an iterative algorithm derived from the Newton-Raphson procedure

Estimation de β

The hessian matrix of $L(\cdot)$ is given by

$$\frac{\delta L^2}{(\delta\beta)^2}(\beta) = -\sum_{i=1}^{n} \frac{\exp(x_i^T\beta)}{(1+\exp(x_i^T\beta))^2} x_i x_i^T$$

Estimation de β

The Newton-Raphson iterative procedure consists of calculating a sequence of $\beta^{(0)}, \dots, \beta^{(t)}, \dots$ telle que

$$\beta^{(t)} = \beta^{(t-1)} - \left[\frac{\delta L^2}{(\delta\beta)^2} (\beta^{(t-1)}) \right]^{-1} \frac{\delta L}{\delta\beta} (\beta^{(t-1)})$$

This sequence converges to a value β^* such that $\frac{\delta L}{\delta\beta}(\beta^*) = 0_d$

Variability of estimates

Since the maximum likelihood estimator is used, it is possible to construct asymptotic confidence intervals

$$\mathbb{V}(\hat{\beta}) \approx \left[\sum_{i=1}^{n} \frac{\exp(x_i^T \beta)}{(1 + \exp(x_i^T \beta))^2} x_i x_i^T \right]^{-1}$$