Statistique et aide à la décision
Session 3

David Causeur
Agrocampus Ouest
IRMAR CNRS UMR 6625

Plan

Estimation des paramètres de régression

Supposons $Y = \pm 1$ et X quantitative :

$$\text{logit}\, \mathbb{P}(Y = +1 \mid X = x) \quad = \quad \beta_0 + \beta_1 x$$

Ajuster un modèle de régression logistique revient à **estimater** les paramètres β_0 et β_1.

Méthode d'estimation : maximum de vraisemblance.

Vraisemblance d'un modèle

La vraisemblance $\ell_{x,y}(\beta)$, où $\beta = (\beta_0, \beta_1)'$, d'un modèle de régression logistique est la probabilité conjointe (pour tout $i = 1, \ldots, n$) d'observer $Y_i = y_i$ pour le $i$ème individu pour lequel $X_i = x_i$:

$$\ell_{x,y}(\beta) = \mathbb{P}\Big(Y_1 = y_1, \ldots, Y_n = y_n \mid X_1 = x_1, \ldots, X_n = x_n \Big).$$

Les individus étant mutuellement indépendants,

$$\ell_{x,y}(\beta) \quad = \quad \mathbb{P}\Big(Y_1 = y_1 \mid X_1 = x_1 \Big) \ldots \mathbb{P}\Big(Y_n = y_n \mid X_n = x_n \Big).$$

- 

Déviance d'un modèle

La déviance $\mathcal{D}_{x,y}(\beta)$ d'un modèle de régression logistique est définie par $\mathcal{D}_{x,y}(\beta) = -2 \log \ell_{x,y}(\beta)$

On en déduit une expression explicite :

$$\begin{aligned} \mathcal{D}_{x,y}(\beta) &= -2 \sum_{i=1}^{n} \log \frac{1}{1 + \exp\big(-y_i(\beta_0 + \beta_1 x_i)\big)}, \\ &= 2 \sum_{i=1}^{n} \log\Big(1 + \exp\big(-y_i(\beta_0 + \beta_1 x_i)\big)\Big). \end{aligned}$$

Déviance d'un modèle

La déviance $\mathcal{D}_{x,y}(\beta)$ d'un modèle de régression logistique est définie par $\mathcal{D}_{x,y}(\beta) = -2 \log \ell_{x,y}(\beta)$

La déviance s'interprète comme le critère des moindres carrés :

- $\mathcal{D}_{x,y}(\beta) \geq 0$,
- $\mathcal{D}_{x,y}(\beta) = 0$ pour un ajustement parfait,
- Plus $\mathcal{D}_{x,y}(\beta)$ est petit, meilleur est l'ajustement.

- 

Minimisation de la déviance

Soit $\mathcal{S}_{x,y}(\beta)$ le gradient de la déviance

$$\mathcal{S}_{x,y}(\beta) = \begin{pmatrix} \frac{\partial \mathcal{D}_{x,y}}{\partial \beta_0}(\beta) \\ \frac{\partial \mathcal{D}_{x,y}}{\partial \beta_1}(\beta) \end{pmatrix}.$$

L'estimateur $\hat{\beta}$ du maximum de vraisemblance de β vérifie

$$\mathcal{S}_{x,y}(\hat{\beta}) = 0.$$

Plan

Précision de l'estimation

Lorsque n est grand,

$$\begin{pmatrix} \hat{\beta}_0 - \beta_0 \\ \hat{\beta}_1 - \beta_1 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} ; V_{\hat{\beta}} = (X' V X)^{-1} \right)$$

où

$$X = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \; V = \begin{pmatrix} v(x_1) & 0 & \dots & 0 \\ 0 & v(x_2) & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & v(x_n) \end{pmatrix}.$$

et $v(x_i) = \pi(x_i)(1 - \pi(x_i))$.

Précision de l'estimation

Test de Wald de $H_0 : \beta_1 = 0$

$$z_{\beta_1} = \frac{\hat{\beta}_1}{\sqrt{\hat{V}(\hat{\beta}_1)}} \sim_{H_0} \mathcal{N}(0;1), \text{ approximativement si n est grand}$$

Intervalle de confiance de niveau $1-\alpha$ de β_1

$$\left[\hat{\beta}_1 - u_{1-\frac{\alpha}{2}}\sqrt{\hat{V}(\hat{\beta}_1)}; \hat{\beta}_1 + u_{1-\frac{\alpha}{2}}\sqrt{\hat{V}(\hat{\beta}_1)}\right],$$

où $u_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1-\frac{\alpha}{2}$ de la loi $\mathcal{N}(0;1)$.

-