Et donc:

$$I(\beta) = -\frac{\partial^2}{\partial \beta^2} \ell(\beta) = -\frac{\partial}{\partial \beta} S(\beta) = ?$$

On cherche l'élément $(j,k)$ de cette matrice $I(\beta)$:

$$[I(\beta)]_{j,k} = -\frac{\partial^2}{\partial \beta_j \partial \beta_k} \ell(\beta) = -\frac{\partial}{\partial \beta_k}\left(\frac{\partial}{\partial \beta_j} \ell(\beta)\right)$$

$$= -\frac{\partial}{\partial \beta_k}\left([S(\beta)]_j\right)$$

$$= +\frac{1}{\sigma^2} \sum_{i=1}^{n} X_{i,j} \times X_{i,k}$$

$$= \frac{1}{\sigma^2}\left(X_i' \; X_j\right)$$

Ainsi, on déduit que:

$$I(\beta) = \frac{1}{\sigma^2}\left(X'X\right)$$

En appliquant maintenant l'algorithme de Newton-Raphson de (4.4.2), on a donc :

$$\hat{\beta}^{(i+1)} = \hat{\beta}^{(i)} + I^{-1}(\hat{\beta}^{(i)}) S(\hat{\beta}^{(i)})$$

$$\Rightarrow \hat{\beta}^{(i+1)} = \hat{\beta}^{(i)} + \left[\frac{1}{\sigma^2}(X'X)\right]^{-1} \frac{1}{\sigma^2}\left(X'Y - X'X\hat{\beta}^{(i)}\right)$$

$$\Rightarrow \hat{\beta}^{(i+1)} = \hat{\beta}^{(i)} + \sigma^2 (X'X)^{-1} \frac{1}{\sigma^2}\left(X'Y - (X'X)\hat{\beta}^{(i)}\right)$$

$$\Rightarrow \hat{\beta}^{(i+1)} = \hat{\beta}^{(i)} + (X'X)^{-1}X'Y - (X'X)^{-1}(X'X)\hat{\beta}^{(i)}$$

$$\Rightarrow \boxed{\hat{\beta}^{(i+1)} = (X'X)^{-1}X'Y \quad !}$$

Conclusion intéressante :

L'approche GLM basée sur la méthode du maximum de vraisemblance et utilisant l'algorithme de Newton-Raphson donne le même résultat pour $\hat{\beta}$ que l'approche "régression multiple" basée sur la minimisation de la distance quadratique

Remarque sur la validation globale et locale du modèle sous la loi normale:

Tel que vu au point (1) de la section (4.4.3),

$$D(\hat{\beta}) \overset{\text{Loi Normale}}{=} SSE = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \text{Somme de } N(0,1) \text{ au carré!} \sim \chi^2(n-(p+1))$$

Ainsi, la statistique $\chi^2_{OBS}$ est donnée par:

$$\chi^2_{OBS} = D(\beta_{H_0}) - D(\beta_{H_1})$$

$$= SSE(M_0) - SSE(M_1)$$

Or, puisque

$$SSE(M_0) \sim \chi^2(n-(p_{H_0}+1))$$

et

$$SSE(M_1) \sim \chi^2(n-(p_{H_1}+1))$$

... on peut dire que tester si

$$\chi^2_{OBS} \geq \chi^2_{\alpha}\Big(\underbrace{[n-(p_{H_0}+1)]-[n-(p_{H_1}+1)]}_{= p_{H_1}-p_{H_0}}\Big)$$

Différence dans le # de param entre $M_0$ et $M_1$

... reviens à tester si:

$$F^* = \frac{[SSE(M_0)-SSE(M_1)]/[p_{H_1}-p_{H_0}]}{SSE(M_1)/(n-p_{H_1})} \geq F_{\alpha}(p_{H_1}-p_{H_0}; n-p_{H_1})$$

... soit le test de Fisher partiel de (3.2.3).

## 4.6) Modèle de régression logistique:

### Contexte motivant l'utilisation de ce modèle:

Phénomène dichotomique (c-a-d: $Y_i$ peut prendre seulement 2 valeurs ... 0 ou 1 !)

ex: • Est-ce que l'assuré $i$ (avec caract. $X_i$) va avoir des sinistres ?

• Est-ce que ce même assuré va renouveler ou non sa police avec nous?

• Est-ce qu'un individu ayant les caractéristiques $X_i$ va décéder ou non dans la prochaine année?

• etc !

Remarque: De façon pratique, il est d'usage de définir $Y_i \in \{0, 1\}$

Dans ce cas, on a que:

$$(Y_i \mid \underset{\sim}{X_i}) \sim \text{Bernoulli}(\underbrace{\pi_i}_{= E(Y_i \mid \underset{\sim}{X_i})}),$$

avec

$$g(\pi_i) = \underset{\sim}{X_i}\underset{\sim}{\beta}$$

Puisque $\pi_i \in ]0,1[$ ($\pi_i$ est une probabilité) et que $\underset{\sim}{X_i}\underset{\sim}{\beta} \in \mathbb{R}$ il faut choisir $g(\cdot)$ telle que:

$$g : ]0,1[ \rightarrow \mathbb{R}.$$

Bien que plusieurs candidats peuvent satisfaire cette condition, il est d'usage d'utiliser une fonction appelée "logit" dans ce contexte:

$$g(x) = \text{logit}(x) \equiv \ln\left(\frac{x}{1-x}\right).$$

Conséquences:

$$g(\pi_i) = X_i \beta$$

$$\Rightarrow \text{logit}(\pi_i) = X_i \beta$$

$$\rightarrow \ln\left(\frac{\pi_i}{1-\pi_i}\right) = X_i \beta$$

... et donc on peut exprimer $\pi_i$ comme suit:

$$\frac{\pi_i}{1-\pi_i} = e^{X_i \beta}$$

$$\Rightarrow \frac{1-\pi_i}{\pi_i} = \frac{1}{e^{X_i \beta}}$$

$$\Rightarrow \frac{1}{\pi_i} - 1 = \frac{1}{e^{X_i \beta}}$$

$$\Rightarrow \frac{1}{\pi_i} = 1 + \frac{1}{e^{X_i \beta}}$$

$$\Rightarrow \pi_i = \frac{1}{1 + \frac{1}{e^{X_i \beta}}}$$

$$\Rightarrow \boxed{\pi_i = \frac{e^{\underset{\sim}{X}_i \underset{\sim}{\beta}}}{1+e^{X_i \beta}}}$$

Interprétations:

(1) $$\ln\left(\frac{\pi_i}{1-\pi_i}\right) = \underset{\sim}{X}_i \underset{\sim}{\beta}$$

$$\Rightarrow \ln\left(\frac{Pr(Y_i=1 \mid \underset{\sim}{X}_i)}{Pr(Y_i=0 \mid \underset{\sim}{X}_i)}\right) = \underset{\sim}{X}_i \underset{\sim}{\beta}$$

... le "log du rapport de cote" est linéaire

(2) $\beta_0$ : le "log du rapport de cote" lorsque tous les variables explicatives $X_i$ sont $= 0$.

(3) $\beta_j$ : Si $X_{ij}$ augmente de une unité (alors que les autres variables explicatives sont fixés) alors le "log du rapport de cote" augmente de $\beta_j$ unité pour l'individu $i$.

Sous la méthode du maximum de vraisemblance, on a :

$$\ell(\beta) = \sum_{i=1}^{n} \ln\left(\Pr(Y_i = y_i \mid X_i)\right)$$

$$= \sum_{i=1}^{n} \ln\left( \pi_i^{y_i} \times (1-\pi_i)^{1-y_i} \right)$$

$$= \sum_{i=1}^{n} \left( y_i \times \ln \pi_i + (1-y_i)\ln(1-\pi_i) \right)$$

$$= \sum_{i=1}^{n} \left( y_i \ln\left(\frac{e^{X_i\beta}}{1+e^{X_i\beta}}\right) + (1-y_i)\ln\left(\frac{1}{1+e^{X_i\beta}}\right) \right)$$

$$= \sum_{i=1}^{n} \left( y_i \ln(e^{X_i\beta}) - y_i \ln(1+e^{X_i\beta}) - (1-y_i)\ln(1+e^{X_i\beta}) \right)$$

$$= \sum_{i=1}^{n} \left( y_i \times (X_i\beta) - \ln(1+e^{X_i\beta}) \right)$$

Par conséquent,

$$S(\beta) = \frac{\partial}{\partial \beta} \ell(\beta) = ?$$

On cherche le $j^e$ élément de ce vecteur $S(\beta)$:

$$[S(\beta)]_j = \frac{\partial}{\partial \beta_j} \ell(\beta)$$

$$= \sum_{i=1}^{n} X_{ij} Y_i - X_{ij} \times \frac{e^{X_i \beta}}{1+e^{X_i\beta}}$$

$$= \sum_{i=1}^{n} X_{ij} \left( Y_i - \frac{e^{X_i \beta}}{1+e^{X_i\beta}} \right)$$

$$\equiv X_j' \left( Y - M(X_j \beta) \right)$$

; où $M(X_j \beta) = \begin{bmatrix} \frac{e^{X_1\beta}}{1+e^{X_1\beta}} \\ \vdots \\ \frac{e^{X_n\beta}}{1+e^{X_n\beta}} \end{bmatrix}$

Ainsi, on déduit que:

$$\boxed{S(\beta) = X' \left( Y - M(X, \beta) \right)}$$

Et donc que:

$$I(\beta) = -\frac{\partial^2}{\partial \beta^2} \ell(\beta) = -\frac{\partial}{\partial \beta} S(\beta) = ?$$

On cherche l'élément (j,k) de cette matrice $I(\beta)$:

$$\left[I(\beta)\right]_{j,k} = -\frac{\partial^2}{\partial \beta_j \partial \beta_k} \ell(\beta) = -\frac{\partial}{\partial \beta_k}\left(\frac{\partial}{\partial \beta_j}\ell(\beta)\right) = -\frac{\partial}{\partial \beta_k}\left[S(\beta)\right]_j$$

$$= \sum_{i=1}^{n} X_{ij} \times X_{ik} \left(\frac{e^{X_i\beta}\left(1+e^{X_i\beta}\right) - \left(e^{X_i\beta}\right)^2}{\left(1+e^{X_i\beta}\right)^2}\right)$$

$$= \sum_{i=1}^{n} X_{ij} \times X_{ik} \left(\frac{e^{X_i\beta}}{\left(1+e^{X_i\beta}\right)^2}\right)$$

$$= \sum_{i=1}^{n} X_{ij} \times X_{ik} \times w_i \quad ; \text{ où } \quad w_i = \frac{e^{X_i\beta}}{\left(1+e^{X_i\beta}\right)^2}$$

Ainsi, on déduit que

$$I(\beta) = X' W X \quad ; \text{ avec}$$

$$W = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & w_n \end{bmatrix}_{n \times n}$$

26-11-2012