Université d'Orléans - Master Econométrie et Statistique Appliquée
Econométrie des Variables Qualitatives

Christophe HURLIN

Correction Examen Mai 2008. C. Hurlin

Exercice 1 : Modèle Logit Ordonné

Soit y_i une variable multinomiale associée à l'individu i pouvant prendre 3 modalités, notées respectivement 0,1 et 2. On admet que cette variable admet une représentation telle que :

$$y_i = \begin{cases} 0 & \text{si } y_i^* < \gamma_1 \\ 1 & \text{si } \gamma_1 \leq y_i^* < \gamma_2 \\ 2 & \text{si } y_i^* \geq \gamma_2 \end{cases} \qquad \forall i = 1, .., N \tag{1}$$

où la variable latente y_i^* vérifie :

$$y_i^* = \alpha + \beta_1 x_i + \beta_2 z_i + \varepsilon_i \tag{2}$$

où x_i et z_i désignent deux variables explicatives. On admet que les termes d'erreur ε_i sont *i.i.d.* et sont distribués selon une loi logistique de variance égale à σ_ε^2.

Question 1 *(2 points)* : On note

$$\widetilde{\varepsilon}_i = \frac{\varepsilon_i}{\sigma_\varepsilon} \frac{\pi}{\sqrt{3}} \tag{3}$$

Par construction, $var(\widetilde{\varepsilon}_i) = 1$, dès lors si l'on admet que $\widetilde{\varepsilon}_i$ suit une loi logistique standard, on montre que :

$$\Pr(y_i = 0) = \Lambda \left[\frac{\pi}{\sigma_\varepsilon \sqrt{3}} \left(\gamma_1 - \beta_0 - \beta_1 x_i - \beta_2 z_i \right) \right] \tag{4}$$

$$\begin{aligned} \Pr(y_i = 1) \quad &= \quad \Lambda \left[\frac{\pi}{\sigma_\varepsilon \sqrt{3}} \left(\gamma_2 - \beta_0 - \beta_1 x_i - \beta_2 z_i \right) \right] \\ &\quad - \Lambda \left[\frac{\pi}{\sigma_\varepsilon \sqrt{3}} \left(\gamma_1 - \beta_0 - \beta_1 x_i - \beta_2 z_i \right) \right] \end{aligned} \tag{5}$$

$$\Pr(y_i = 2) = 1 - \Lambda \left[\frac{\pi}{\sigma_\varepsilon \sqrt{3}} \left(\gamma_2 - \beta_0 - \beta_1 x_i - \beta_2 z_i \right) \right] \tag{6}$$

Question 2 *(1 point)* : Par définition, on a:

$$\alpha_j = \frac{\pi}{\sigma_\varepsilon \sqrt{3}} \beta_j \quad j = \{1, 2\} \tag{7}$$

$$c_j = \frac{\pi}{\sigma_\varepsilon \sqrt{3}} \gamma_j \quad j = \{1, 2\} \tag{8}$$

On ne peut à la fois identitifer les seuils et la constante, dès lors on doit imposer une normalisation de type

$$\alpha_0 = 0 \tag{9}$$

Question 3 *(1.5 point)* : On admet que les réalisations des estimateurs sont :

$$\widehat{\alpha}_0 = 0 \quad \widehat{\alpha}_1 = 4 \quad \widehat{\alpha}_2 = -2 \quad \widehat{c}_1 = 0.5 \quad \widehat{c}_2 = 1 \tag{10}$$

Dès lors pour l'indiviu tel que $x_i = 1$ et $z_i = 1.75$, noté l'index $w_i \alpha$, vaut :

$$w_i \widehat{\alpha} = \widehat{\alpha}_0 + \widehat{\alpha}_1 x + \widehat{\alpha}_2 z = 0.5 \tag{11}$$

Par conséquent on obtient :

$$\widehat{\Pr} (y_i = 0) = \Lambda (\widehat{c}_1 - w_i \widehat{\alpha}) = \frac{1}{1 + \exp [- (0.5 - 0.5)]} = 0.5 \tag{12}$$

$$\begin{aligned} \widehat{\Pr} (y_i = 1) &= \Lambda (\widehat{c}_2 - w_i \widehat{\alpha}) - \Lambda (\widehat{c}_1 - w_i \widehat{\alpha}) \\ &= \frac{1}{1 + \exp [- (1 - 0.5)]} - \frac{1}{1 + \exp [- (0.5 - 0.5)]} = 0.1225 \end{aligned} \tag{13}$$

On peut alors vérifier que la somme des trois probabilités est égale à l'unité puisque :

$$\widehat{\Pr} (y_i = 2) = 1 - \Lambda (\widehat{c}_2 - w_i \widehat{\alpha}) = 1 - \frac{1}{1 + \exp [- (1 - 0.5)]} = 0.3775 \tag{14}$$

Question 4 *(2 points)* : On sait que de façon générale on a :

$$L (y_i, \alpha, c_1, c_2) = \prod_{j=0}^{2} [\Lambda (c_{j+1} - w_i \alpha) - \Lambda (c_j - w_i \alpha)]^{y_{ij}} \tag{15}$$

ou encore

$$\begin{aligned} L (y_i, \alpha, c_1, c_2) &= \Lambda (c_1 - w_i \alpha)^{y_{i0}} \\ &\quad \times [\Lambda (c_2 - w_i \alpha) - \Lambda (c_1 - w_i \alpha)]^{y_{i1}} \\ &\quad \times [1 - \Lambda (c_2 - w_i \alpha)]^{y_{i2}} \end{aligned} \tag{16}$$

où la variable dichotomique y_{ij} est définie par :

$$y_{ij} = \begin{cases} 1 & \text{si } y_i = j \\ 0 & \text{sinon} \end{cases} \quad \forall i = 1.., N \ \ \forall j = 0, 1, 2$$

Dès lors la valeur de cette vraisemblance pour un individu pour lequel on observe $y_i = 1$, $x_i = 1$, $z_i = 2$ lorsque l'on suppose que les paramètres sont égaux à leurs valeurs estimés est défini par :

$$\begin{aligned} L (y_i, \alpha, c_1, c_2) &= \Lambda (\widehat{c}_2 - w_i \widehat{\alpha}) - \Lambda (\widehat{c}_1 - w_i \widehat{\alpha}) \\ &= \Lambda (1 - 0.5) - \Lambda (0.5 - 0.5) \\ &= \Pr [y_i = 1] = 0.1225 \end{aligned} \tag{17}$$

Question 5 *(1 point)* : On observe que $\alpha_1 > 0$, dès lors lorsque la variable x_1 augmente d'une unité, la probabilité $\Pr(y_i = 2)$ augmente (cf. cours)

Question 6 *(2 points)* : Quels sont les effets marginaux associés à la hausse d'une unité de la variable x. On effectuera les calculs pour l'individu de référence tel que $x_i = 1$ et $z_i = 2$.

$$\begin{aligned}\frac{\partial \Pr(y_1 = 0)}{\partial x_i} &= -\frac{\exp(w_i\widehat{\alpha} - \widehat{c}_1)}{[1+\exp(w_i\widehat{\alpha} - \widehat{c}_1)]^2} \times \widehat{\alpha}_1 \\ &= -\frac{\exp(0)}{[1+\exp(0)]^2} \times 4 = -2\end{aligned} \tag{18}$$

$$\begin{aligned}\frac{\partial \Pr(y_1 = 1)}{\partial x_i} &= \left\{\frac{\exp(w_i\widehat{\alpha} - \widehat{c}_1)}{[1+\exp(w_i\widehat{\alpha} - \widehat{c}_1)]^2} - \frac{\exp(w_i\widehat{\alpha} - \widehat{c}_2)}{[1+\exp(w_i\widehat{\alpha} - \widehat{c}_2)]^2}\right\} \times \widehat{\alpha}_1 \\ &= \left\{\frac{\exp(0)}{[1+\exp(0)]^2} - \frac{\exp(-0.5)}{[1+\exp(-0.5)]^2}\right\} \times 4 = 0.2264\end{aligned} \tag{19}$$

$$\begin{aligned}\frac{\partial \Pr(y_1 = 2)}{\partial x_i} &= \frac{\exp(w_i\widehat{\alpha} - \widehat{c}_1)}{[1+\exp(w_i\widehat{\alpha} - \widehat{c}_2)]^2} \times \widehat{\alpha}_1 \\ &= \frac{\exp(4)}{[1+\exp(-0.5)]^2} \times 4 = 1.7736\end{aligned} \tag{20}$$

Exercice 2 : Scoring et Défaillances d'Entreprises (8 points)

Question 1 *(1 point)* : cf. cours

Question 2 *(1 point)* : On montre que pour cet individu :

$$\widehat{\Pr}(y_i = 1) = \Phi\left(x_i\widehat{\beta}\right) = 0.3476 \tag{21}$$

La cote est alors égale à :

$$\widehat{c}_i = \frac{\widehat{\Pr}(y_i = 1)}{1 - \widehat{\Pr}(y_i = 1)} = 0.5328 \tag{22}$$

Question 3 *(1 point)* : Calculez l'effet marginal sur la probabilité de défaillance de l'entreprise décrite à la question précédente d'une hausse marginale du ratio dette sur actif total.

$$\frac{\partial \Pr(y_i = 1)}{\partial TDTA_i} = \phi(x_i\beta) \times \beta_{TDTA} \tag{23}$$

Pour l'individu de référence, on a :

$$x_i\widehat{\beta} = -0.3915 \tag{24}$$

et donc par conséquent :

$$\frac{\partial \Pr(y_i = 1)}{\partial TDTA_i} = \phi(-0.3915) \times 2.927 = 1.0819 \tag{25}$$

Commentaire : cf. cours

Question 5 *(3 points)* : On considère les $87^{ème,}$ $88^{ème}$ et $89^{ème}$ firmes de l'échantillon pour lequelles on a obtenu dans le modèle Logit les valeurs suivantes :

i	YD_i	$Index_i$	$\Pr[YD_i = 1]$
87	0	-3.06	0.044
88	1	-0.99	0.269
89	1	-0.66	0.339

On sait que si l'on note $F(z_{88})$ la valeur de la vraisemblance associée à l'individu 88, où $z_{88} = x_{88}\beta$:

$$\widehat{F}(z_{88}) = \frac{\sum_{i=87}^{89} w_i(z_{88}) y_i}{\sum_{i=88}^{89} w_i(z_{88})} \quad \text{avec } w_i(z_{88}) = K\left(\frac{z_i - z_{88}}{h}\right) \tag{26}$$

Etant donné, les valeurs observées de y_i, on a :

$$\widehat{F}(z_{88}) = \frac{w_{88}(z_{88}) + w_{89}(z_{88})}{\left[\sum_{i=88}^{89} w_i(z_{88})\right]} \tag{27}$$

Dans le cas d'une fonction kernel gaussienne $\forall i = 1, .., N$:

$$K\left(\frac{z_i - z_1^0}{h}\right) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z_i - z_{88}}{h}\right)^2\right] \tag{28}$$

On obtient ainsi :

$$w_{87}(z_{88}) = 0.1934 \tag{29}$$

$$w_{88}(z_{88}) = 0.3989 \tag{30}$$

$$w_{89}(z_{88}) = 0.3917 \tag{31}$$

Dès lors :

$$\widehat{F}(z_{88}) = 0.8035 \tag{32}$$