CHAPITRE 2 : LE MODELE LOGIT

L'estimation du modèle PROBIT peut s'avérer assez compliqué à calculer (à cause de la loi normale). On peut le remplacer par le modèle LOGIT qui fait référence à la loi logistique. La fonction de densité de probabilité est :

$$\boldsymbol{\lambda(l) = \frac{e^{-l}}{(1 + e^{-l})^2}}$$

$\boldsymbol{l} \in [-\infty, \infty]$

La fonction logit :

$$\boldsymbol{\Lambda(l) = P[L \leq l] = \frac{1}{1 + e^{-l}}}$$

Qui est plus simple que (1.3).

La probabilité p s'écrit de fait quand y prend la valeur 1 :

$$\boldsymbol{p = P[L \leq cste + bx] = \Lambda(cste + bx) = \frac{1}{1 + e^{-(cste+bx)}}}$$

$$\boldsymbol{= \frac{\exp(cste + bx)}{1 + \exp(cste + bx)}}$$

On en déduit la probabilité quand y=0 (du fait de la contrainte P(y_i=0) + P(y_i=1) = 1) :

$$1 - p = \frac{1}{\mathbf{1} + \mathbf{exp}(\boldsymbol{cste} + \boldsymbol{bx})}$$

Un seul vecteur est nécessaire pour calculer les deux probabilités.

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Autre méthode pour calculer les probabilités p et 1-p sans déduction, en calculant 2 vecteurs :

$$\boldsymbol{p_1} = \frac{\mathbf{exp}(\boldsymbol{cste_0} + \boldsymbol{b_0x})}{\mathbf{exp}(\boldsymbol{cste_0} + \boldsymbol{b_0x}) + \mathbf{exp}(\boldsymbol{cste_1} + \boldsymbol{b_1x}}$$

$$\boldsymbol{p_2} = \frac{\mathbf{exp}(\boldsymbol{cste_1} + \boldsymbol{b_1x})}{\mathbf{exp}(\boldsymbol{cste_0} + \boldsymbol{b_0x}) + \mathbf{exp}(\boldsymbol{cste_1} + \boldsymbol{b_1x}}$$

Si on pose $cste_0$ = 0 et b_0 = 0 on retrouve bien les mêmes résultats car $\mathbf{exp}(\mathbf{0} + \mathbf{0}\boldsymbol{x}) = \mathbf{1}$

NB : de ce fait le modèle logit pourra être étendu aux cas avec choix supérieur à 2.

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Pour le calcul des effets marginaux, probabilité estimée, ML…on change seulement :

$$\emptyset \; par \; \lambda$$

$$\varphi \; par \; \Lambda$$

SECTION 1 : UN EXEMPLE (suite)

Le modèle logit que nous voulons estimer est :

$$\boldsymbol{P(auto = 1) = \Lambda(cste + bX)}$$

Les estimations par ML avec SAS 9.4 sont :

Model Fit Summary	
Number of Endogenous Variables	1
Endogenous Variable	Y
Number of Observations	21
Log Likelihood	-6.16604
Maximum Absolute Gradient	2.51183E-7
Number of Iterations	7
Optimization Method	Quasi-Newton
AIC	16.33208
Schwarz Criterion	18.42113

Mesures du critère qualificatif de lissage		
Mesure	Valeur	Formule
Likelihood Ratio (R)	16.732	2 * (LogL - LogL0)
Upper Bound of R (U)	29.065	- 2 * LogL0
Aldrich-Nelson	0.4435	R / (R+N)
Cragg-Uhler 1	0.5492	1 - exp(-R/N)
Cragg-Uhler 2	0.7329	(1-exp(-R/N)) / (1-exp(-U/N))
Estrella	0.6947	1 - (1-R/U)^(U/N)
Adjusted Estrella	0.5497	1 - ((LogL-K)/LogL0)^(-2/N*LogL0)
McFadden's LRI	0.5757	R / U
Veall-Zimmermann	0.7639	(R * (U+N)) / (U * (R+N))
McKelvey-Zavoina	0.8969	
N = # d'observations, K = # de régresseurs		

Algorithm converged.

Résultats estimés des paramètres					
Paramètre	DDL	Valeur estimée	Erreur type	Valeur du test t	Approx. de Pr > |t|
Intercept	1	-0.237575	0.750477	-0.32	0.7516
X	1	0.531098	0.206425	2.57	0.0101

On en déduit les éléments suivants :

$$\frac{dp}{dx} = \lambda(-.237575 + .531098 * aug) * .531098$$

Calculs des effets marginaux (logit)

Aug	$\lambda(-.237575 + .531098 * \text{aug})$	$\frac{dp}{dx}$
5 mn (.5)	0,249951097	0,132748528
10 mn (1)	0,244691653	0,129955248
15 mn (1.5)	0,231439264	0,12291693
20 mn (2)	0,211890569	0,112534658
25 mn (2.5)	0,188290618	0,10000077
30 mn (3)	0,162937801	0,08653594
35 mn (3.5)	0,137795147	0,073182727

A titre de comparaison :

Comparaison des effets marginaux (logit & probit)

Aug	$logit \frac{dp}{dx}$	$probit \frac{dp}{dx}$
5 mn (.5)	0,132748528	0,119245007
10 mn (1)	0,129955248	0,11640672
15 mn (1.5)	0,12291693	0,111107729
20 mn (2)	0,112534658	0,103690475
25 mn (2.5)	0,10000077	0,094615401
30 mn (3)	0,08653594	0,084413746
35 mn (3.5)	0,073182727	0,073636458

On remarque que même si les estimations par ML sont différentes les probabilités marginales sont à peu près identiques. Il en va de même pour le **EMM = 0.046154**.

On peut aussi prévoir le comportement d'un individu qui doit choisir entre ces deux modes de transport :

$$\hat{\boldsymbol{p}} = \boldsymbol{\Lambda}\left(\widehat{\boldsymbol{cste}} + \hat{\boldsymbol{b}} * \boldsymbol{aug}\right) = \frac{\mathbf{1}}{\mathbf{1} + \boldsymbol{e}^{-(\boldsymbol{cste}+\boldsymbol{b}*\boldsymbol{aug})}}$$

Calcul et comparaison des probabilités estimées (logit & probit)

Aug	$\hat{p}$ logit	$\hat{p}$ probit
5 mn (.5)	0,506993044	.5341
10 mn (1)	0,572858402	.5931
15 mn (1.5)	0,636237792	.6501
20 mn (2)	0,695216369	.7038
25 mn (2.5)	0,748413732	.7535
30 mn (3)	0,795063042	.7983
35 mn (3.5)	0,834969928	.8372

On retrouve quasiment le même résultat.

Lequel choisir ? Pas de réponse pour l'instant.
On va se focaliser sur les mesures qualitatives des estimations. Comment mesurer la qualité de l'estimation ?

NB1 : à la place de la loi normale ou logistique on aurait pu aussi choisir :

1. le modèle de Weibull
2. le modèle log log complémentaire
3. le modèle de Gompertz
4. le modèle de Burr…

NB2 : Amemiya (1981) a proposé les approximations suivantes entre OLS-Probit-Logit dans l'intervalle de probabilités de 30% à 70% pour les pentes :

$$\boldsymbol{\widehat{\beta}_{ols} \approx .4\beta_{probit}}$$

$$\boldsymbol{\widehat{\beta}_{ols} \approx .25\beta_{logit}}$$

$$\boldsymbol{\beta_{logit} \approx 1.6\beta_{probit}}$$

SECTION 2 : TESTS D'HYPOTHESE DE WALD

On retrouve les tests standards économétriques, à savoir :

- le test du t-ratio,
- les tests de combinaisons linéaires…

SECTION 3 : TESTS D'HYPOTHESE ML

Nous ne pouvons pas calculer le R^2, le test de Fisher…car nous sommes en non linéaire. Toutefois la fiabilité des résultats peut être appréhendée par de multiples éléments (dans le désordre) :

- LR (ratio de vraisemblance)

On veut tester :

$$\begin{cases} \boldsymbol{Ho} & \boldsymbol{\beta_{k-1} = 0} \\ \boldsymbol{H1} & \boldsymbol{au\ moins\ 1\ \beta \neq 0} \end{cases} \sim \boldsymbol{\chi^2_{k-1}}$$

H0 veut dire : tous les coefficients nuls sauf la constante.

Si LR > χ^2_{k-1} alors on accepte H1

Remplace le test de Fisher usuel.

- Indice du ratio de vraisemblance de McFadden

$$R^2_{MF} = 1 - \frac{Ln\ L}{Ln\ L_0} \in [0, 1]$$

Avec :

Ln L la log-vraisemblance du modèle non contraint
Ln l_0 la log-vraisemblance du modèle contraint (tous les β =0 sauf la constante)

- Estrella ou Scaled R²

$$R^2_{E1} = 1 - \left(\frac{Ln\ L}{Ln\ L_0}\right)^{-\left(\frac{2}{n}\right)Ln\ L_0} \in [0, 1]$$

- Estrella ajusté

$$R^2_{E2} = 1 - \left(\frac{Ln\ L - k}{Ln\ L_0}\right)^{-\left(\frac{2}{n}\right)Ln\ L_0} \in [0, 1]$$

- Autres statistiques

$$R^2_{CU1} = 1 - \left(\frac{L_0}{L}\right)^{\frac{2}{N}} \quad \text{(Cragg-Uhler 1)}$$

$$R^2_{CU2} = \frac{1-(L_0/L)^{\frac{2}{N}}}{1-L_0^{\frac{2}{N}}} \quad \text{(Cragg-Uhler 2)}$$

$$R^2_A = \frac{2(\ln L - \ln L_0)}{2(\ln L - \ln L_0)+N} \quad \text{(Aldrich-Nelson)}$$

$$R^2_{VZ} = R^2_A \frac{2\ln L_0 - N}{2\ln L_0} \quad \text{(Veall-Zimmermann)}$$

Référence :

https://books.google.fr/books?id=OE0UfAhit4kC&pg=PA941&lpg=PA941&dq=adjusted+estrella&source=bl&ots=pgKxBqfg3s&sig=opinApOL54kXthwNFriEYY_b7G8&hl=fr&sa=X&ved=0ahUKEwjzm-2w64DVAhVESBQKHRpxClEQ6AEIVzAK#v=onepage&q=adjusted%20estrella&f=false