

# Université d'Orléans - Maitrise Econométrie

## Econométrie des Variables Qualitatives

Examen Mai 2005. C. Hurlin

### Exercice 1 (10 points) : Modele Logit Conditionnel

On considère une enquête sur les modes de transport reliant trois villes qui a été menée en Australie en 1986 auprès de 210 personnes (Bradley, 1993; Hensher et Greene, 2002). Celles-ci pouvaient choisir entre l'avion, le train, le bus ou la voiture. L'enquête leur demandait pour chaque mode de transport d'indiquer le temps de transfert entre leur domicile et le lieu de départ<sup>1</sup> (variable `tps_transf`), le temps de trajet (variable `tps_trajet`) et le coût du voyage (variable `cout`). L'enquête demandait aussi le revenu du ménage en milleurs de dollars australiens (variable `revenu`) et le nombre de personnes devant voyager (variable `nbre`). Le mode de transport vaut 1 pour l'avion, 2 pour le train, 3 pour le bus et 4 pour la voiture.

**Question 1** (1 point) : Quel modèle proposeriez vous pour modéliser dans ce cas les choix de moyen des transports des individus? Quels sont les principaux avantages de ce modèle?

**Question 2** (1 point) : En supposant qu'il n'existe pas de données manquantes, combien le fichier de données comporte-t-il d'observations au total (nombre de lignes) ? *Vous justifierez précisément votre réponse.*

**Question 3** (1 point) Commentez économiquement les résultats du modèle logit multinomial (cf. figure 1) avec les 3 variables `cout`, `tps_trans` et `t_ps_trajet`. Ces résultats vous paraissent-ils logiques?

**Question 4** (1 point) On considère à présent le même modèle en introduisant des constantes spécifiques à chaque mode de transport (cf. figure 2). Commentez la spécification et les résultats d'estimation.

**Question 5** (2 points) A partir de ses caractéristiques (cf. figure 3) et en utilisant le modèle de la question 4, (i) calculez la probabilité que l'individu 2 choisisse les différents modes de transport. (ii) Quelle prévision donne alors le modèle? Y a-t-il dans ce cas une erreur de prévision?

**Question 6** (2 points) Calculez pour l'individu 2 l'effet marginal d'une hausse de 1 unité du coût de transport associé à la modalité voiture sur la probabilité qu'il prenne la voiture et la probabilité qu'il prenne le bus. Commentez.

**Question 7** (1 point) On veut enfin introduire les variables individuelles (revenu et nombre) dans le modèle. Comment proposez vous d'introduire ces variables dans la spécification sachant qu'elles ne varient pas avec les modalités?

**Question 8** (1 point) Commentez les résultats d'estimation du modèle avec variables individuelles (cf. figure 4).

---

<sup>1</sup>Ce temps est nul dans le cas de la voiture.

## Exercice 2 (6 points) : Modele Logit Indépendant

On cherche à expliquer la catégorie professionnelle CS en fonction l'âge (variable age), de l'âge de fin d'études en années révolues (afinet), du sexe (variable binaire fem qui vaut 1 pour femme et 0 pour homme) et de la nationalité (variable etr, variable binaire qui vaut 1 si l'individu est de nationalité étrangère). La CS est codée de la façon suivante: cs=1 si le salarié est cadre, cd=2 s'il exerce une profession intermédiaire, cs=3 s'il est employé, cs=4 s'il est ouvrier.

**Question 1** (2 points) : Commentez précisément tous les éléments le tableau de résultats d'estimation (figure 1)

**Question 2** (1.5 point) : Calculez l'odds ratio (rapport de risques relatifs) associé au fait d'être une femme lorsqu'il s'agit d'occuper un emploi de cadre plutôt qu'un emploi d'ouvrier.

**Question 3** (2.5 points) : Calculez les effets marginaux de l'âge au point moyen pour les hommes de nationalité française sachant que l'âge moyen par CS est respectivement égal à 37.4; 36.6, 38.8 et 37.6 ans et que l'âge de fin d'études moyens est par CS égal à 22.5; 20.5; 18 et 16.9 ans.

## Exercice 3 (4 points) : Modele Tobit

On dispose d'un échantillon de taille  $N = 10$  et l'on considère un modèle Tobit tel que

$$y_i^* = 0.1 + 0.5x_i + \varepsilon_i \quad (1)$$
$$\varepsilon_i \text{ N.i.d } (0, 4)$$
$$y_i = \begin{cases} y_i^* & \text{si } y_i^* > 0 \\ 0 & \text{sinon} \end{cases} \quad \forall i = 1, \dots, N$$

On admet que la valeur moyenne de la variable  $x_i$  est égale à  $\bar{x}_i = (1/N) \sum x_i = 2$ .

**Question 1** (1 point) (i) Calculez l'effet marginal d'une variation unitaire de variable explicative  $x_i$ , sur la prévision de la variable latente  $y_i^*$ . (ii) Exprimez cet effet sous la forme d'une élasticité.

**Question 2** (1.5 points) On admet que pour un individu la valeur de  $x_i$  est égale à 2. Calculez pour cet individu l'effet marginal d'une variation unitaire de la variable explicative sur la prévision de la variable dépendante  $y_i$ .

**Question 3** (1.5 points) Décomposez l'effet marginal précédent en deux composantes selon la décomposition de McDonald et Moffit (1980). Interpretez.

Figure 1

The MDC Procedure						
Conditional Logit Estimates						
Model Fit Summary						
Dependent Variable	choix					
Number of Observations	210					
Number of Cases	840					
Log Likelihood	-246.85867					
Maximum Absolute Gradient	7.09508E-9					
Number of Iterations	4					
Optimization Method	Newton-Raphson					
AIC	499.71734					
Schwarz Criterion	509.75866					
Discrete Response Profile						
Index	mode	Frequency	Percent			
0	1	58	27.62			
1	2	63	30.00			
2	3	30	14.29			
3	4	59	28.10			
Goodness-of-Fit Measures for Discrete Choice Models						
Measure	Value	Formula				
Likelihood Ratio (R)	88.526	$2 * (\text{LogL} - \text{LogL0})$				
Upper Bound of R (U)	582.24	$- 2 * \text{LogL0}$				
Aldrich-Nelson	0.2965	$R / (R+N)$				
Cragg-Uhler 1	0.3440	$1 - \exp(-R/N)$				
Cragg-Uhler 2	0.3669	$(1 - \exp(-R/N)) / (1 - \exp(-U/N))$				
Estrella	0.3670	$1 - (1 - R/U)^{(U/N)}$				
Adjusted Estrella	0.3454	$1 - ((\text{LogL} - K) / \text{LogL0})^{(-2/N * \text{LogL0})}$				
McFadden's LRI	0.1520	$R / U$				
Veall-Zimmermann	0.4035	$(R * (U+N)) / (U * (R+N))$				
N = # of observations, K = # of regressors						
Parameter Estimates						
Parameter	DDL	Estimation	Erreur standard	t Value	Approx Pr >  t	Gradient
tps_transf	1	-0.0340	0.004643	-7.32	<.0001	-978E-12
tps_trajet	1	-0.002193	0.000458	-4.79	<.0001	-7.1E-9
cout	1	0.008891	0.004877	1.82	0.0683	5.31E-10

Source : Les modèles logit polytomiques non ordonnés : théorie et applications, Cédric AFSA ESSAFI, INSEE

Figure 2  
*Résultats de l'estimation*

Parameter Estimates						
Parameter	DDL	Estimation	Erreur standard	t Value	Approx Pr >  t	Gradient
avion	1	4.7399	0.8675	5.46	<.0001	3.43E-10
train	1	3.9532	0.4686	8.44	<.0001	2.8E-10
bus	1	3.3062	0.4583	7.21	<.0001	-609E-12
tps_transf	1	-0.0969	0.0103	-9.37	<.0001	-1.88E-8
tps_trajet	1	-0.003995	0.000849	-4.70	<.0001	-2.08E-7
cout	1	-0.0139	0.006651	-2.09	0.0365	1.643E-8

Source : Les modèles logit polytomiques non ordonnés : théorie et applications, Cédric AFSA ESSAFI, INSEE

Figure 3

*Les 8 premières observations de la table en entrée*

ident	mode	choix	tps_transf	tps_trajet	cout	revenu	nbre
1	1	0	69	100	59	35	1
1	2	0	34	372	31	35	1
1	3	0	35	417	25	35	1
1	4	1	0	180	10	35	1
2	1	0	64	68	58	30	2
2	2	0	44	354	31	30	2
2	3	0	53	399	25	30	2
2	4	1	0	255	11	30	2

Source : Les modèles logit polytomiques non ordonnés : théorie et applications, Cédric AFSA ESSAFI, INSEE

Figure 4

*Résultats de l'estimation*

Parameter Estimates						
Parameter	DDL	Estimation	Erreur standard	t Value	Approx Pr >  t	Gradient
avion	1	6.0352	1.1382	5.30	<.0001	3.73E-8
train	1	5.5735	0.7113	7.84	<.0001	2.6E-8
bus	1	4.5047	0.7958	5.66	<.0001	-8.25E-8
tps_transf	1	-0.1012	0.0111	-9.08	<.0001	-3.38E-6
tps_trajet	1	-0.004131	0.000893	-4.63	<.0001	-0.00003
cout	1	-0.008670	0.007876	-1.10	0.2710	1.516E-6
revenu_avion	1	0.007481	0.0132	0.57	0.5710	2.201E-6
revenu_train	1	-0.0592	0.0149	-3.98	<.0001	-3.49E-8
revenu_bus	1	-0.0209	0.0164	-1.28	0.2012	-3.22E-6
nbre_avion	1	-0.9224	0.2585	-3.57	0.0004	7.452E-8
nbre_train	1	0.2163	0.2336	0.93	0.3546	7.809E-8
nbre_bus	1	-0.1479	0.3428	-0.43	0.6661	-2.61E-7

Source : Les modèles logit polytomiques non ordonnés : théorie et applications, Cédric AFSA ESSAFI, INSEE