

APDT TD 4 – LÉCONTE & PLANCHE

Question 1 : interprétation du modèle de choix modal

Les variables explicatives d'un modèle logit sont généralement organisées en trois termes :

- Le coût généralisé, dépendant à la fois de l'offre de transport, du déplacement et du ménage
- La motorisation des ménages, caractéristique propre à ceux-ci
- L'offre de stationnement, caractéristique de l'offre de transport

Pour le fichier calmodpr.xls, fichier de calage du modèle de répartition modale entre TC et VP, l'utilité pour chacun des motifs de déplacements est calculée comme la somme de trois termes :

- Le temps généralisé VP entre zone i et j, divisé par la motorisation de la zone i
- Le temps généralisé TC entre zone i et j, multiplié par la motorisation de la zone i
- La variable densité de la zone j, relative à la population et aux emplois issus de l'EMD

Chacun de ces termes étant affecté d'un coefficient de répartition modale, propre à chaque motif.

On peut donc assimiler le modèle utilisé à un modèle de type logit, où la différence d'utilité peut s'exprimer de la façon suivante :

$$(V_{VP} - V_{TC}) = \log\left(\frac{1 - \%TC}{\%TC}\right), \text{ où } \%TC = \frac{1}{1 + \exp\left(k_m + \frac{ttc_{ij}}{\tau c_m / mot_i} - \frac{tvp_{ij}}{\tau vp_m * mot_i} - \frac{d_j}{\delta_m}\right)}$$

avec ttc_{ij} , le temps généralisé TC entre les zones i et j ;
 tvp_{ij} , le temps généralisé VP entre les zones i et j ;
 mot_i , la motorisation de la zone i ;
 d_j , la densité de la zone j, exprimée en population + emplois à l'ha ;
 k_m , τc_m , τvp_m et δ_m , paramètres de répartition pour le motif m.

Question 2 : Analyse de la précision des données

Les effectifs enquêtés posent un problème de fiabilité statistique des données. Il s'agit donc maintenant de déterminer un nombre de chaînes enquêtées « seuil ». Il faut donc trouver un compromis entre un seuil élevé pour une bonne précision des données et un seuil bas pour retenir un grand nombre d'OD, mais des résultats moins précis.

Analyse de la taille des échantillons pour l'ensemble des OD, pour le motif « autres motifs »

En observant sommairement les colonnes F à H du nombre de chaînes enquêtées, on remarque qu'il y a assez peu d'OD dont l'effectif semble satisfaisant (au moins 30 déplacements). L'introduction d'un seuil sélectionnant les OD significatives apparaît alors pertinente.

- *Seuil de 50 déplacements (par défaut dans le modèle)*

Pour un tel seuil, 55 OD sont retenues pour leur volume de déplacement, ce qui représente 9572 déplacements puis 66.3% des chaînes enquêtées.

Pour un nombre de 50 déplacements, on a 95% de chances que la part de marché TC se situe entre 6.95% et 19.24% pour une valeur réelle de 13.09%. On a donc un intervalle de confiance d'environ 13%.

Le tableau ci-dessous énumère les résultats pour différentes valeurs de seuil. On estime que sur une matrice désagrégée de 625 composantes, il n'est pas pertinent de considérer un seuil supérieur à 50, lequel occasionnerait une perte trop importante de données.

Seuil	10	20	30	40	50
Nombre d'O-D	214	139	96	67	55
Nombre de chaînes	13185	12160	11121	10109	9572
poids de ces O-D	91%	84%	77%	70%	66%
Intervalle de confiance	31%	22%	18%	15%	13%
Evolution poids		7,1%	7,2%	7,0%	3,8%
Evolution confiance		9,8%	4,1%	2,4%	1,6%

Le meilleur compromis entre un intervalle de confiance étroit et un poids de déplacements doit être trouvé. Nous prendrons ici un seuil de 40 déplacements encadré des 2 seuils respectivement bas et hauts de 30 et de 50 déplacements. Nous ne prenons pas un seuil plus bas, car il conduirait à garder la valeur seuil basse de 20 déplacements, dont l'intervalle de confiance ne nous semble pas satisfaisante.

Question 3 : Calage du modèle de répartition sur découpage en 25 zones, avec seuil à 40 OD

Analyse des régressions

Seuil	30	Proba du t	40	Proba du t	50	Proba du t
Coefficient R ²	0.742043459	de Student	0.739054794	de Student	0.760401254	de Student
Valeur critique de F	6.04444E-16		7.44775E-11		1.22712E-09	
Constante	4.277192041	1.29121E-09	3.826993643	4.55097E-05	4.183888666	0.000309615
Variable X 1 (TgTC*mot)	0.041994942	0.005455093	0.063701237	0.015702029	0.055355399	0.095491807
Variable X 2 (TgVP*mot)	-0.062945027	1.04671E-06	-0.058181488	0.000195679	-0.060113633	0.001586002
Variable X 3 (d)	-0.004439399	9.7048E-06	-0.004670253	0.001083061	-0.005739188	0.002440459

Les coefficients de corrélation obtenus sont compris entre 0.73 et 0.76. C'est donc plus de 73% de la variance qui est expliquée par la droite de régression. De plus, la valeur critique du F de Fischer (<5%) nous permet de rejeter l'hypothèse de nullité des coefficients, ceci pour chaque seuil.

La probabilité du t de Student n'est pas satisfaisante dans pour un seuil à 50 (surlignés en jaune). La régression n'est donc pas acceptable pour cette valeur-ci.

D'après la formule énoncée à la question 1, on peut se prononcer sur la cohérence des signes. Pour que la part de TC augmente, il faut que X1 soit de signe opposé à X2 et X3. Même si c'est le cas ici, les coefficients semble être trop faibles pour être significatifs.

Les résultats ne sont pas satisfaisants aux vues des résultats de la régression linéaire. Afin de traiter les O-D à l'effectif trop faible, nous agrégerons les 25 zones en 7, avec pour objectif de réduire la sensibilité au seuil sans trop affecter les O-D de départ.

Question 4: Calage du modèle de répartition sur découpage en 7 zones

La matrice étant ici agrégée, nous nous permettons de tester plusieurs seuils, plus élevés que ce que nous avons pu faire jusqu'à présent.

Seuil	50	60	70	80	90	100	110
Nombre d'O-D	34	33	33	32	29	28	27
Nombre de chaînes	14 076	14 024	14 024	13 949	13 688	13 593	13 492
poids de ces O-D	97,57 %	97,21 %	97,21 %	96,69%	94,88 %	94,22 %	93,52%
Intervalle de confiance	13,49 %	12,30 %	11,37 %	10,63 %	10,01 %	9,49 %	9,05 %
Evolution poids	1 %	0,3 %	0 %	0,5 %	1,8 %	0,66 %	0,7 %
Evolution confiance	1,7 %	1,2 %	0,9 %	0,7 %	0,6 %	0,6 %	0,4 %

Nous constatons tout d'abord que les seuils supérieurs à 80 donnent des nombres d'OD inférieurs à 30. Nous ne travaillerons donc pas avec les seuils 90, 100, 110...

Concernant les seuils restants, les résultats fournis sont bons :

- Les poids des OD sont tous supérieurs à 96 %
- Les intervalles de confiance sont inférieurs à 14 %

Les seuils restants sont donc tous acceptables.

Régressions linéaires :

Seuil	50	Proba du t	60	Proba du t	80	Proba du t
Coefficient R2	0.81440815	de Student	0.813929219	de Student	0.898220614	de Student
Valeur critique de F	3.00685E-07		5.27242E-07		3.96081E-10	
Constante	3.471033612	1.15504E-06	3.549092886	2.73402E-06	4.136993466	7.06937E-10
Variable X 1 (TgTC*mot)	0.034248788	0.007728663	0.033952923	0.009312418	0.028534122	0.003130833
Variable X 2 (TgVP*mot)	-0.038565355	0.000214235	-0.040668178	0.000653483	-0.055222385	2.27346E-07
Variable X 3 (d)	-0.005851689	0.000115043	-0.00576423	0.000202354	-0.004954456	2.64662E-05

Par comparaison des valeurs issues des différentes régressions linéaires, nous pouvons dire que celles correspondant aux seuils 50, 60 et 80 fournissent des bons résultats. Les valeurs sont extrêmement proches d'un seuil à l'autre. (Les résultats sont les mêmes pour les seuils 60 et 70)

Les signes, ici aussi sont cohérents.

Seul le seuil = 80 semble différer des autres valeurs issues des régressions précédentes. Le modèle nous semble stable.

Question 5 : Calage du modèle de répartition sur découpage en 25 zones, sans seuil.

Attention : nous avons décelé une erreur à la dernière minute. L'initialisation a été faite à partir d'une régression linéaire fortement erronée. Nous ne savons pas dans quelle mesure les résultats suivants s'en sont vus affectés.

Nous allons travailler avec une valeur seuil égale à 40. C'est celle qui ne semble la plus pertinente au regard du travail fait dans le cadre de la section 1.

<i>Seuils</i>	30 OD		40 OD		50 OD	
	Résultat moyenne	<i>Résultat écart-type</i>	Résultat moyenne	<i>Résultat écart-type</i>	Résultat moyenne	<i>Résultat écart-type</i>
constante	3,7553	1,3017	3,6748	1,1844	3,9911	1,2798
coeff TC	0,0948	0,0335	0,0890	0,0334	0,0721	0,0375
coeff VP	-0,0646	0,0210	-0,0642	0,0202	-0,0637	0,0233
coeff densité	-0,0059	0,0012	-0,0052	0,0012	-0,0053	0,0013
R2	12,49%	0,03	14,94%	0,07	20,78%	0,15

Cette méthode est-elle moins sensible aux seuils ?

Les résultats sont surprenants dans le sens où les coefficients de détermination sont médiocres : 12, 15 et 21 %.

Pourtant, les coefficients qui forment le tableau ci-dessus sont proches dans l'ensemble. Ce qui laisse penser que la méthode est relativement peu sensible aux seuils. Ainsi, l'erreur provient éventuellement de l'initialisation de la macro. Nous avons fait le choix d'utiliser la configuration 7x7 en seuil = 50 OD. Mais ce choix résulte peut-être de conclusions erronées – notamment en raison de la complexité de l'exercice.