

ÉCONOMÉTRIE DES VARIABLES QUALITATIVES

Daniel Szpiro

L'objectif de cette formation est de comprendre et maîtriser les méthodes de régression à mettre en œuvre lorsque la variable dépendante est qualitative (binaire ou catégorielle). Cette situation se rencontre dans différents champs d'application : choix financiers, notation du risque, segmentation de clientèle, marketing, économie du travail, économie de l'environnement, étude des comportements, etc.

À la suite de cette formation, on saura :

- *Quel est le contexte d'utilisation de la régression logistique*
- *Pourquoi la régression linéaire multiple ne peut être mise en œuvre sur des variables endogènes qualitatives*
- *Comment construire un modèle de régression logistique*
- *Quelles sont les conditions d'utilisation de la régression logistique*
- *Comment mesurer l'adéquation du modèle aux données*
- *Comment interpréter les résultats d'une régression logistique (logiciel SAS)*
- *Comment fonctionne la régression logistique ordinale.*
- *Comment fonctionne la régression polytomique*

Après les exposés théoriques, des illustrations seront présentées (commentaires de listings, programmes SAS, etc.). Les TD permettront de maîtriser les raisonnements et démonstrations des méthodes exposées.

Séance 1 :

Chapitre 1 Introduction

section 1 Taxonomie des variables

section 2 Un outil : la dérivation matricielle

section 1 Intuition et exemple le lancer de pièce

section 2 Propriétés asymptotiques : absence de biais, efficacité et normalité

Séance 2 :

Chapitre 2 La qualité d'un estimateur

section 1 Densité, régularité et propriétés dérivées

section 2 Efficacité d'un estimateur et borne de Rao-Cramér

Séance 4 et 5 :

Chapitre 4 Le modèle dichotomique

section 1 Pourquoi ne pas utiliser les m.c.o. ?

1.1 Considérations graphiques

1.2 Problèmes sur les résidus

section 2 Les variables latentes

2.1 Exposé du principe

2.2 Exemples

section 3 L'E.M.V. du Logit

3.1 La vraisemblance de l'échantillon

Séance 3 :

Chapitre 3 L'estimateur du maximum de vraisemblance

- 3.2 (Existence et unicité du maximum)
- 3.3 L'équation de l'estimateur
- 3.4 L'algorithme numérique de Newton-Raphson
- 3.5 (La matrice de variance - covariance asymptotique)

Séance 6 :

- Chapitre 5 Résultats d'estimation et interprétation
- section 1 Validité des résultats
- 1.1 Test d'un coefficient
- 1.2 Test de la nullité d'un ensemble de coefficients
- 1.3 Validité générale du modèle : table de classification et « syndrome du fer à repasser »
- section 2 Interprétation des coefficients
- 2.1 Interprétation du coefficient d'une exogène dichotomique
- 2.2 Interprétation « locale » du coefficient d'une exogène quantitative
- Procédure SAS : Logistic

Séance 7 :

- Chapitre 6 Les variables qualitatives polytomiques
- section 1 Variables à modalités ordonnées
- 1.1 Comment généraliser le cas dichotomique ?
- 1.2 Écriture de la vraisemblance
- section 2 Variables à modalités non ordonnées
- 2.1 Comment généraliser le cas dichotomique ?
- 2.2 Probabilité d'une des occurrences
- 2.3 Vraisemblance du modèle Logit multinomial
- 2.4 Interprétation des coefficients
- 2.5 Modèle Logit conditionnel
- section 3 Variables à choix emboîtés
- 3.1 L'indépendance au choix non retenu
- 3.2 Test de l'hypothèse
- 3.3 Modélisation emboîtée
- Que faire en cas de modalités non exclusives ?
- Procédure SAS : Logistic ou Mdc ;
- Procédure Stata : Mlogit

RÉFÉRENCES BIBLIOGRAPHIQUES

- Stuart, Ord, Arnold (1999) : *Kendall's advanced theory of statistics*, vol 2A - ch 17 (pour l'efficacité et le maximum de vraisemblance)
- Gourieroux C. (1989) : *Économétrie des variables qualitatives*, Economica, 430 p
- Maddala G. S. (1996) : *Limited Dependant and Qualitative Variables in Econometrics*, Econometric Society Monograph, Cambridge UP , 401 p
- Le Blanc D., Lollivier S., Marpsat M., Verger D. (2000), *L'économétrie et l'étude des comportements. Présentation et mise en œuvre des modèles de régression qualitatifs. Les modèles univariés à résidus logistiques ou normaux*, Document n° 0001, Série des documents de travail «Méthodologie Statistique» INSEE.
- Afsa-Essafi C. (2003) : *Les modèles logit polytomiques non ordonnés : théorie et applications*, documents de travail INSEE