ÉCONOMÉTRIE DES VARIABLES QUALITATIVES

UNIVERSITÉ DU MANS (EXAMEN, L3)

EXERCICE I

(1) Soit un échantillon de variables dichotomiques $(y_1, \ldots, y_N)$ identiquement et indépendament distribués selon une loi de Bernouilli de paramètre p. Chaque y vaut 1 avec probabilité p ou 0 avec probabilité $(1-p)$. Calculer l'espérance et la variance de y.

(2) Écrire la fonction de probabilité de y.

(3) Écrire la fonction de log-vraisemblance.

(4) Calculer l'estimateur du maximum de vraisemblance de p.

(5) Calculer la variance de cet estimateur et caractériser le comportement asymptotique de cet estimateur.

EXERCICE II

On s'intéresse aux conséquences de l'omission d'une variable exogène dans un modèle logit. On suppose que les données sont générées par un modèle logit admettant deux variables explicatives, x_1 et x_2. La probabilité que y_i ($i = 1, \ldots, N$) soit égal à 1 est :

$$\mathbb{P}\left(y_i = 1\right) = F\left(x_{1,i}b_1 + x_{2,i}b_2\right)$$

où F est la fonction de répartition d'une loi logistique, b_1 et b_2 les vraies valeurs (inconnues) des paramètres.

(1) Donner une expression de la fonction de répartition d'une loi logistique, $F(z)$.

(2) Donner l'expression de la fonction de densité de probabilité d'une loi logistique, puis montrer que :

$$\frac{\mathrm{d}}{\mathrm{d}z}\log F(z) = 1 - F(z)$$

et

$$\frac{\mathrm{d}}{\mathrm{d}z}\log 1 - F(z) = -F(z)$$

(3) L'économètre estime le modèle logit en omettant la variable x_2 (qu'il n'observe pas). Écrire la fonction de vraisemblance.

(4) Calculer la condition du premier ordre, définissant l'estimateur du maximum de vraisemblance du paramètre b_1. Montrer que cette condition peut s'écrire sous la forme :

$$\sum_{i=1}^{N}\left(y_i - F\left(x_{1,i}\hat{b}_1\right)\right)x_{1,i} = 0$$

Interpréter cette condition, en la comparant à la condition d'identification dans modèle linéaire (avec les MCO par exemple). Est-il possible d'obtenir une expression analytique de $\hat{b}_1$? Pourquoi ? Expliquer comment on peut obtenir $\hat{b}_1$.

(5) Expliquer pourquoi la condition du premier ordre est liée à la condition moment suivante :

$$\mathbb{E}\left[\left(y - F(x_1\hat{b}_1)\right)x_1\right] = 0$$

(6) En utilisant la loi des espérances itérées et le modèle générateur des données, montrer que l'on peut réécrire la condition de moment sous la forme :

$$\mathbb{E}_x\left[\left(F(x_1b_1 + x_2b_2) - F(x_1\hat{b}_1)\right)x_1\right] = 0$$

(7) Sous quelle condition $\hat{b}_1$ est-il un estimateur convergent ? Comparer avec la condition d'absence de biais de variable omise dans un modèle linéaire.