Partiel TIAD - Master 1 IM
Lundi 10 Mars 2014

Durée 2 h. Calculatrices autorisées. Seul un formulaire sur feuille double est autorisé.

Tran Viet Chi, `chi.tran@univ-lille1.fr`, bureau 316 (bâtiment M3).

Exercice 1 (Equations de vraisemblance pour un modèle Tobit)
On considère le modèle suivant :

$$Y_i = Y_i^* \mathbf{1}_{Y_i^*>0}, \qquad \text{où } Y_i^* = X_i\beta + \varepsilon_i \tag{1}$$

où pour $i \in \{1, \cdots, n\}$, $X_i = (1, X_{i1}, \cdots X_{iK}) \in \mathcal{M}_{1,K+1}(\mathbb{R})$ est le vecteur ligne des variables explicatives caractérisant l'individu i (pour simplifier, on pourra noter $X_{i0} = 1$ dans les calculs). Le vecteur des coefficients est $\beta =^t(\beta_0, \cdots \beta_K)$. On suppose que les ε_i sont des variables i.i.d. de loi $\mathcal{N}(0, \sigma^2)$. La variable $Y_i^* \in \mathbb{R}$ est une variable continue latente, ce sont des Y_i dont on dispose. Comme d'habitude, on notera matriciellement $Y^* = X\beta + \varepsilon$. On notera φ et Φ les densité et fonction de répartition de la loi $\mathcal{N}(0, 1)$, on rappelle bien-sûr que :

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2).$$

On notera $n_1 = \text{Card}\ \{i \,|\, Y_i > 0\}$ et $n_0 = \text{Card}\ \{i \,|\, Y_i = 0\}$.

1. Calculer $\mathbb{P}(Y_i = 0 \,|\, X)$.

2. Ecrire la log-vraisemblance $\log \mathcal{L}(Y_1, \cdots, Y_n; \beta, \sigma^2)$, conditionnellement à X, des expliquées $(Y_1, \cdots Y_n)$.

3. Dériver la log-vraisemblance par rapport à β_k ($k \in \{0, \ldots, K\}$) et σ^2. En déduire que le gradient de la log-vraisemblance par rapport à β est

$$\nabla_\beta \log \mathcal{L} = \sum_{i,\, Y_i=0} \frac{-\varphi(-X_i\beta/\sigma)\, {}^tX_i}{\sigma\Phi(-X_i\beta/\sigma)} + \sum_{i,\, Y_i>0} \frac{(Y_i - X_i\beta)\, {}^tX_i}{\sigma^2} \tag{2}$$

4. Montrer que :

$$\widehat{\sigma}^2 = \frac{1}{n_1} \sum_{i \,|\, Y_i>0} Y_i(Y_i - X_i\widehat{\beta}).$$

5. Montrer que $\widehat{\beta}$ satisfait :

$$\frac{1}{\widehat{\sigma}} \sum_{i \,|\, Y_i>0} (Y_i - X_i\widehat{\beta})\, {}^tX_i = \sum_{i \,|\, Y_i=0} \frac{\varphi(-X_i\widehat{\beta}/\widehat{\sigma})\, {}^tX_i}{\Phi(-X_i\widehat{\beta}/\widehat{\sigma})}. \tag{3}$$

6. On définit par X^1 la matrice obtenue en rassemblant les explicatives pour les individus i tels que $Y_i > 0$ et X^0 celle obtenue pour les individus i tels que $Y_i = 0$. On note également

Y^1 le vecteur des Y_i pour les i tels que $Y_i > 0$. On définit par γ^0 le vecteur de composantes $\varphi(-X_i\beta/\sigma)/\Phi(-X_i\beta/\sigma)$ pour les i tel que $Y_i = 0$. Montrer en réécrivant matriciellement (3) que l'estimateur du maximum de vraisemblance de β est :

$$\widehat{\beta} = ({}^tX^1X^1)^{-1}\,{}^tX^1Y^1 - \widehat{\sigma}({}^tX^1X^1)^{-1}\,{}^tX^0\gamma^0.$$

7. Qu'en déduire quant à la méthode qui consiste à estimer β en faisant une régression MCO sur le sous-échantillon des individus i tels que $Y_i = Y_i^*$? Citer une méthode d'estimation préférable vue en cours.

Exercice 2 (Modèle de marché)

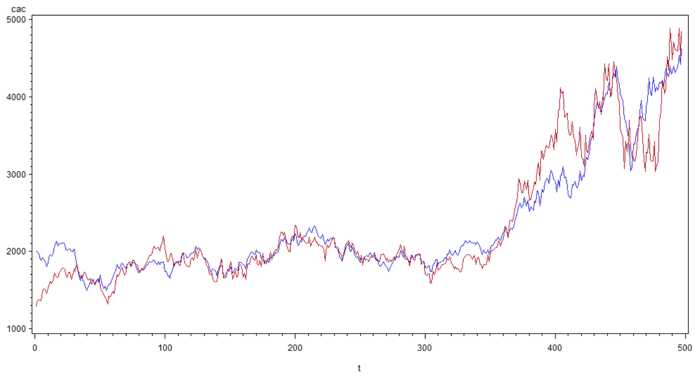

FIGURE 1 – *Evolution hebdomadaire du CAC40 et de l'action d'Elf (renormalisée par le facteur multiplicatif 2325/70) entre janvier 1990 et juillet 1999.*

On considère le modèle simple suivant pour le rendement d'une action :

$$R_t = \alpha + \beta RM_t + \varepsilon_t \tag{4}$$

où R_t est le rendement de l'action au temps t, RM_t est le rendement du portefeuille de marché (un indicateur de santé de l'économie tel que le CAC40 pour la France). Si $\beta > 1$, les variations du portefeuille de marché sont amplifiées par l'action. Elle est considérée comme agressive. Si $\beta < 1$, l'action est défensive. On considère l'action d'Elf. Entre janvier 1990 et juillet 1999, nous avons 497 dates.

1. Dans l'annexe, pour la régression 1, des informations ont été perdues. Recalculer le coefficient de détermination, les statistiques de Student pour le test de significativité des coefficients, la borne supérieure de l'intervalle de confiance pour le coefficient β.

2. La régression est-elle globalement significative ?

3. Faire un test (à 5%) pour dire s'il y a une rupture de tendance à la date 300. Pour ce faire, on utilisera les régressions de l'annexe. On donne (en langage **R**) :

```
qf(0.95,2,493)=3.01401
qf(0.975,2,493)=3.71662
qf(0.025,2,493)=0.02532
qf(0.95,298,195)=1.24296
qf(0.975,298,195)=1.296165
qf(0.025,298,195)=0.776873
```