Université de Kairouan
ISMAI
Enseignants : Med Essaied Hamrita

Année Universitaire 2015/2016
1$^{\text{ière}}$ année Mastère IF
Modélisation Statistique

TD3 : Extensions de la régression linéaire

Exercice 1

On dispose d'une série d'observations trimestrielles sur la consommation de boissons allant du premier trimestre 2008 au dernier trimestre 2015. On veut vérifier si ces observations sont entachées de fluctuations saisonnières. Pour cela, on propose d'estimer le modèle suivant :

$$C_t = \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \alpha_4 D_{4t} + \varepsilon_t \tag{1}$$

où D_{it} vaut 1 si l'observation est celle du trimestre i ($i = 1, \dots, 4$) et 0 sinon.
On donne les informations suivantes :

$$\sum_{t=1}^{32} D_{1t}C_t = 8;\ \sum_{t=1}^{32} D_{2t}C_t = 10;\ \sum_{t=1}^{32} D_{3t}C_t = 14;\ \sum_{t=1}^{32} D_{4t}C_t = 16;\ \sum_{t=1}^{32} C_t^2 = 72$$

1) Interpréter les paramètres de ce modèle.
2) Peut-on estimer ce modèle on y introduisant un terme constant en plus des 4 variables indicatrices. Justifier le réponse.
3) Calculer le niveau de consommation moyen sur l'ensemble de la période.
4) Donner une estimation efficace des paramètres.
5) Montrer qu'en dépit de l'absence d'un terme constant dans le modèle, la somme des résidus est nulle.
6) Calculer la matrice des variances covariances des estimateurs.
7) Montrer que ce modèle peut être réécrit sous la forme suivante :

$$C_t = \beta_0 + \beta_1 D_{1t} + \beta_2 D_{2t} + \beta_3 D_{3t} + \varepsilon_t \tag{2}$$

8) Interpréter les paramètres de ce modèle et en donner une estimation efficace.
9) Écrire l'hypothèse d'absence d'effets saisonniers et tester sa validité empirique en utilisant respectivement les modèles (1) et (2).
10) Calculer la variabilité totale de la consommation de boissons corrigée des fluctuations saisonniers.

Exercice 2

On considère le modèle suivant : $S_i = \beta_0 + \beta_1 sex_i + \varepsilon_i$ où S_i désigne le taux de salaire horaire de l'employé i exprimé en dinars, sex_i une variable indicatrice qui vaut 1 si l'employé i est une femme et 0 si i est un homme et ε_i le terme d'erreur vérifiant : $\varepsilon_i \overset{iid}{\sim} N(0,\sigma^2)$.

1) Interpréter les coefficients du modèle.

2) L'estimation de ce modèle par la technique des MCO sur un échantillon de taille 30 fournit les résultats suivants :

$$\widehat{S}_i = \underset{(.5)}{10} - \underset{(.75)}{3}\, sex_i$$

où les chiffres entre parenthèses indiquent les écart-types estimés des paramètres.

a) Peut-on conclure à une différence significative du taux de salaire selon le sexe.

b) Calculer la part de la variabilité totale des taux de salaire expliquée par la variable sex.

3) On considère maintenant le modèle suivant :

$$S_i = \beta_0 + \beta_1 sex_i + \beta_2 A_i + \beta_3 A_i^2 + \varepsilon_i$$

où A_i indique l'age de l'employé i. Les résultats d'estimations sont :

$$\widehat{S}_i = -4 - 2sex_i + 0.6A_i - 0.006A_i^2; \;\; R^2 = 0.6,\, SCR = 5.4$$

a) Préciser l'intérêt d'introduire la variable A_i^2.

b) Tester, au seuil de 5%, la significativité globale du modèle.

c) Tester, au seuil de 5%, l'hypothèse nulle $H_0 :\; \beta_2 = \beta_3 = 0$.

d) Calculer le salaire moyen d'une employée âgée de 40 ans. Le comparer à celui d'un employé ayant le même âge.

e) Donner l'équation de l'effet marginal de la variable A_i sur le taux de salaire, toutes choses égales par ailleurs.

f) En déduire le niveau d'âge pour lequel le taux de salaire est maximum.

Exercice 3 (Concours IFID 2013)

On considère le modèle suivant : $Y_t = \alpha + \beta X_t + \theta X_{t-1} + \lambda Y_{t-1} + \varepsilon_t, \quad t = 1,\ldots,40.$

1) Écrire en fonction des paramètres les effets de court terme et de long terme d'une variation de Y_t sur la variation de X_t.

L'estimation du modèle a fournit les résultats suivants :

$$\widehat{Y}_t = \underset{(1.5)}{-6.05} + \underset{(0.35)}{1.16}\, X_t - \underset{(0.18)}{0.44}\, X_{t-1} - \underset{(0.12)}{0.36}\, Y_{t-1}$$

Les chiffres entre parenthèses correspondent aux écart-types estimés des paramètres.

2) Calculer les estimateurs de ECT et ELT.

3) Donner une estimation du retard moyen.

4) Tester au seuil de 5%, l'hypothèse $H_0 :\; ELT = 0.5$ sachant que la matrice des variances covariances estimée est :

$$\widehat{V}\left(\widehat{\beta}\right) = 10^{-4}\begin{pmatrix} 5 & -1.5 & 2 \\ -1.5 & 3 & 1 \\ 2 & 1 & 2 \end{pmatrix}$$