Université de Kairouan
ISMAI
Enseignants : Med Essaied Hamrita

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

TD2 : Régression linéaire multiple

Exercice 1

On considère le modèle de régression multiple suivant :

$$\widehat{y}_i = 4 + 0.4x_{1i} + 0.9x_{2i}$$

On dispose aussi des informations suivantes :

$$SCR = 520\,;\; R^2 = 8/60\,;\; (X'X)^{-1} = \frac{1}{3900}\begin{bmatrix} \frac{3900}{29} & 0 & 0 \\ & 80 & -10 \\ & & 50 \end{bmatrix}$$

1) Préciser la taille de l'échantillon.
2) Calculer $\overline{y}$, S_{x_1y} et S_{x_2y}.
3) Estimer la matrice des variances covariances des paramètres et tester la significativité de chaque paramètre.
4) Dresser le tableau d'analyse de la variance et effectuer un test de significativité globale.
5) Tester les hypothèses suivantes :
 i) $H_0 : \beta_1 + \beta_2 = 1$.
 ii) $H_0 : \beta_1 = 0$ et $\beta_2 = 1$.
6) Réestimer les paramètres β_0 et β_2 sous la contrainte $\beta_=0$ et tester la significativité individuelle du modèle contraint.
7) Soit SCR_c la somme des carrés des résidus du modèle contraint ($\beta_1 = 0$). Effectuer le test suivant : $H_0 : SCR_c = SCR$. Le résultat du test est-il prévisible ? Justifier votre réponse.

Exercice 2

La demande d'un bien, notée y_t, est fonction linéaire de son prix unitaire x_{1t} et du prix unitaire d'un bien substituable x_{2t} :

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t,\;\; \varepsilon \overset{iid}{\sim} N(0,\sigma^2) \tag{1}$$

1) Donner une interprétation économique aux paramètres du modèle (1) et donner les signes attendus des paramètres β_1 et β_2.

2) Sur un échantillon de taille $T = 25$, on dispose des information suivantes :

$$\widetilde{X}'\widetilde{X} = \begin{bmatrix} S_{x_1x_1} & S_{x_1x_2} \\ S_{x_2x_1} & S_{x_2x_2} \end{bmatrix} = \begin{bmatrix} 15 & 5 \\ 5 & 10 \end{bmatrix}, \ \widetilde{X}'\widetilde{Y} = \begin{bmatrix} S_{x_1y} \\ S_{x_2y} \end{bmatrix} \begin{bmatrix} 23.25 \\ 52.75 \end{bmatrix}, \ S_{yy} = 16.7.$$

$$\overline{Y} = 5.228, \overline{X}_1 = 0.24, \overline{X}_2 = 0.36$$

Estimer **tous** les paramètres du modèle (1) par MCO ainsi que la matrice des variances covariances de leurs estimateurs.

3) Apprécier la qualité d'ajustement linéaire en calculant le coefficient de détermination ajusté $\overline{R}^2$.

4) Tester, au seuil $\alpha = 5\%$, la significativité individuelle des coefficients.

5) Dresser le tableau d'analyse de la variance et effectuer un test de significativité globale.

6) On cherche maintenant à tester l'hypothèse : $\widehat{\beta}_1 + \widehat{\beta}_2 = 0$.

a) Écrire le modèle contraint.

b) Estimer le modèle contraint et déterminer SCR_c.

c) Quelle est votre décision au risque de 5%.

Exercice 3 (Concours Technologues 2007)

On se propose d'expliquer l'évolution du niveau de production dans le secteur privé, Y, en fonction de deux facteurs K et L, désignant respectivement le stock de capital et l'effectif total employé. Pour cela, on estime le modèle de régression suivant sur un échantillon de 35 observations :

$$y_t = \beta_1 k_t + \beta_2 l_t + \varepsilon_t$$

où $y_t = \ln Y_t - \overline{\ln Y_t}$, $k_t = \ln K_t - \overline{\ln K_t}$, $l_t = \ln L_t - \overline{\ln L_t}$ et ε_t est un terme aléatoire normalement distribué $\varepsilon \overset{iid}{\sim} N(0, \sigma^2)$. On dispose des informations suivantes :

$$X'X = \begin{bmatrix} S_{kk} & S_{kl} \\ & S_{ll} \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ & 6 \end{bmatrix}; X'Y = \begin{bmatrix} S_{ky} \\ S_{ly} \end{bmatrix} = \begin{bmatrix} 3.3 \\ 4.2 \end{bmatrix}; S_{yy} = 3.415.$$

où $S_{xy} = \sum x_t y_t$.

1) Donner la signification économique des coefficients β_1 et β_2.

2) Estimer β_1 et β_2 par la méthode des MCO.

3) Calculer le coefficient de détermination R^2.

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

5) Spécifier l'hypothèse des rendements d'échelle constants et tester statistiquement sa validité empirique au seuil de 5%.

6) Ré-estimer β_1 et β_2 sous l'hypothèse que les rendements d'échelle sont constants.

Exercice 4 (D'aprés Examen 3ième Gestion FSEGT, 1994)

On considère la fonction de production spécifiée selon le modèle suivant :

$$y_t = \beta_0 + \beta_1 L_t + \beta_2 K_t + \varepsilon_t, \qquad t = 1, \ldots, 23 \tag{2}$$

avec ε_t un terme d'erreur vérifiant les hypothèses classiques des moindres carrés ordinaires (MCO), y_t représente le logarithme du niveau de production, L et K indiquent, respectivement, les facteurs de production travail et capital, exprimés également en logarithme.
On sait que le modèle (2) peut se ramener au modèle centré suivant :

$$y_t - \overline{y} = \beta_1 \left(L_t - \overline{L}\right) + \beta_2 \left(K_t - \overline{K}\right) + \varepsilon_t - \overline{\varepsilon}, \;\; t = 1, \ldots, 23 \tag{3}$$

où les variables $\overline{y}$, $\overline{L}$, $\overline{K}$ et $\overline{\varepsilon}$ constituent des valeurs moyennes calculées sur la totalité de l'échantillon. On dispose des informations suivantes :

$$\overline{L} = 10,\ \overline{K} = 5,\ \overline{y} = 12,\ S_{LL} = 12,\ S_{LK} = 8,\ S_{KK} = 12,\ S_{Ly} = 10,\ S_{Ky} = 8,\ S_{yy} = 10$$

1) Mettre le modèle (2) sous la forme matricielle $Y = X\beta + \varepsilon$.
2) Montrer que le modèle centré (3) est obtenu en transformant le modèle (2) par la matrice $M_s = I - S(S'S)^{-1}S'$ où S est le vecteur unité d'ordre approprié.
3) Les erreurs transformées du modèle centré vérifient-elles les hypothèses classiques des MCO. Calculer la matrice des variances-covariances des erreurs ε
4) Estimer les paramètres du modèle centré, β_1, β_2, et σ^2. Peut-on obtenir une estimation pour la constante β_0 ? Interpréter le contenu économique du modèle, et la plausibilité des résultats.
5) Estimer la matrice des variances-covariances du vecteur $\widehat{\beta}$, où $\widehat{\beta}' = (\widehat{\beta}_1 \;\; \widehat{\beta}_2)$.
6) Tester la significativité individuelle des paramètres β_1 et β_2 au seuil de 5%.
7) Tester la significativité globale de la régression au seuil de 5%.
8) Tester simultanément les hypothèses $\beta_1 = 1$ et $\beta_2 = 0$ au seuil de 5%. Interpréter économiquement le résultat de ce test. Procéder selon deux méthodes et vérifier qu'elles fournissent le même résultat.