

TD1 : Régression linéaire simple

Exercice 1

On considère le modèle de régression linéaire simple suivant :

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

avec ε_i vérifiant les hypothèses habituelles.

1) Montrer que le résidu $e_i = y_i - \hat{y}_i$ vérifie les relations suivantes :

$$\sum (x_i - \bar{x}) \hat{e}_i = 0 \text{ et } \sum \hat{y}_i e_i = 0$$

2) Soient : $y_i^* = y_i - \bar{y}$ et $x_i^* = x_i - \bar{x}$. Montrer que l'estimateur de la régression de y_i^* sur x_i^* , par la méthode des MCO, que l'on note par $\hat{\beta}_1^*$, est égale à $\hat{\beta}_1$.

3) Estimer, par MCO, le modèle suivant : $y_i = \alpha_1 x_i + u_i$ avec $V(u_i) = \sigma^2$. Montrer que $V(\hat{\alpha}_1) = \frac{\sigma^2}{\sum x_i^2}$.

4) Estimer le modèle suivant : $y_i = \alpha_0 + u_i$ et Calculer la somme des carrés des résidus.

Exercice 2

On considère le tableau suivant :

| Année | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|----------|------|------|------|------|------|------|------|------|------|------|
| <i>Y</i> | 5 | 5.5 | 6 | 7 | 7.2 | 7.7 | 8.4 | 9 | 9.7 | 10 |
| <i>M</i> | 2 | 2.5 | 3.2 | 3.6 | 3.3 | 4 | 4.2 | 4.6 | 4.8 | 5 |

où *Y* désigne le revenu national et *M* la quantité offerte de monnaie (en millions de dinars).

1) Régresser le revenu national *Y* sur la quantité offerte de monnaie *M* et donner les écarts-types estimés.

2) Interpréter les paramètres du modèle.

3) Tester la significativité individuelle au seuil $\alpha = 5\%$.

4) Déterminer l'élasticité au point moyen (\bar{Y}, \bar{M}) et vérifier si l'on peut accepter l'hypothèse d'une élasticité unitaire au seuil $\alpha = 5\%$.

Exercice 3

Le tableau suivant décrit les dépenses de consommation (*Y*) et le revenu hebdomadaire (*X*) pour un échantillon de 10 ménages :

| | | | | | | | | | | |
|----------|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| <i>Y</i> | 70 | 65 | 90 | 95 | 110 | 115 | 120 | 140 | 155 | 150 |
| <i>X</i> | 80 | 100 | 120 | 140 | 160 | 180 | 200 | 220 | 240 | 260 |

On donne : $\sum X_i Y_i = 205500$, $\sum X_i^2 = 332000$, et $\sum Y_i^2 = 132100$.

- 1) Déterminer l'équation d'ajustement linéaire $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ par la méthode des MCO. Interpréter les valeurs trouvées.
- 2) Estimer la variance des erreurs σ^2 . En déduire les variances des estimateurs trouvés.
- 3) Calculer et interpréter le coefficient de détermination R^2 .
- 4) Construire un intervalle de confiance au seuil $\alpha = 5\%$, pour $\hat{\beta}_0$, $\hat{\beta}_1$ et pour $\hat{\sigma}^2$. Interpréter.
- 5) Effectuer le test $H_0 : \beta_1 = 0.3$ vs $H_1 : \beta_1 \neq 0.3$ au niveau de signification de 5%.

Exercice 4 (D'après Concours IFID 2013)

On considère la relation linéaire entre le taux d'endettement public en pourcentage du PIB (noté Y) et le taux d'investissement public en pourcentage du PIB (noté X) sur une période de 40 années :

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t, \quad t = 1, \dots, 40 \quad (1)$$

avec ε_t un terme d'erreur vérifiant les hypothèses classiques des moindres carrés ordinaires (MCO).

On donne $\sum X_t = 320$, $\sum Y_t = -12$, $S_{xx} = 160$, $S_{yy} = 200$ et $S_{xy} = 80$ où $S_{ij} = \sum (i_t - \bar{i})(j_t - \bar{j})$.

- 1) Estimer par la méthode des MCO les paramètres β_0 et β_1 , notés respectivement $\hat{\beta}_0$ et $\hat{\beta}_1$. Commenter.
- 2) Écrire l'équation d'analyse de la variance et calculer ses composantes. En déduire le coefficient de détermination linéaire R^2 .
- 3) Calculer la variance estimée de $\hat{\beta}_1$.
- 4) Effectuer le test de significativité de β_1 au seuil $\alpha = 5\%$. Commenter.

Exercice 5

On considère le modèle suivant : $\hat{y}_t = 1 + 1.8x_t$, $T = 20$ et $\hat{\sigma}_{\beta_1} = 0.05$.

- 1) Tester, au seuil $\alpha = 5\%$:
 - i) $H_0 : \beta_1 = 0$.
 - ii) $H_0 : \beta_1 = 2$.
- 2) Calculer le coefficient de détermination linéaire R^2 .
- 3) Calculer la variation de la somme des carrés résiduelles (ΔSCR) due à $H_0 : \beta_1 = 2$.

Exercice 6

Reprendre les deux exercices (2 et 3) et répondre à ses questions avec le logiciel R (On représentera graphiquement la variable dépendante en fonction de celle explicative).