

SERIE TD N°3
Variables indicatrices, multicolinéarité

Exercice N°1 :

Un modèle de production de service du secteur du tourisme est spécifié de la manière suivante :

$$QPS_t = a_0 + a_1VA_t + a_2POP_t + \varepsilon_t$$

Avec :

QPS_t = production du secteur tourisme pour l'année t ;

VA_t = valeur ajoutée du secteur tourisme pour l'année t ;

POP_t = population pour l'année t.

L'économètre chargé de l'estimation de ce modèle sur 18 ans s'interroge sur la perturbation entraînée par l'effet d'une guerre pour l'année 16. Pour répondre à cette question, il intègre à son modèle de base une variable indicatrice D_t tel que :

$$\begin{cases} D_t = 0 \text{ pour } t = 1 \text{ à } 15 \text{ et } t = 17 \text{ et } 18 \\ D_t = 1 \text{ pour } t = 16 \end{cases}$$

L'estimation du modèle économétrique est la suivante :

$$QPS_t = 2340,4 + 23,5VA_t + 0,3POP_t - 120,56D_t + e_t$$

- L'effet de « guerre » a-t-il une influence significative sur la production du service du secteur du tourisme ?

Données :

$$t_{a_0}^* = 4,5, t_{a_1}^* = 2,2, t_{a_2}^* = 2,9, t_{a_3}^* = 5,8, n = 18, R^2 = 0,65, t_{14}^{0,05} = 2,14$$

Exercice N°2

Afin de déterminer les facteurs explicatifs de la réussite de la licence en sciences économiques, on spécifie le modèle suivant :

$$NL = a_0 + a_1ND + a_2DS + \varepsilon$$

Où :

NL= note moyenne obtenue en licence,

ND= note moyenne obtenue en fin de deuxième année,

DS= variable indicatrice de genre (1 pour les hommes et 0 pour les femmes).

L'estimation à partir d'un échantillon de 60 étudiants conduit aux résultats suivants :

$$NL = 8,5 + 0,3ND - 1,2DS + e$$

- Le fait d'être homme ou femme a-t-il une influence sur la note obtenue en licence ?

Données :

$$t_{a_0}^* = 4,5, t_{a_1}^* = 7,1, t_{a_2}^* = 2,3, n = 60, R^2 = 0,72, t_{57}^{0,05} = 1,96$$

SERIE TD N°3

Variables indicatrices, multicolinéarité

Exercice N° 3 :

Une entreprise cherche à déterminer quelles sont les influences respectives de la publicité (x_1), de la promotion auprès des distributeurs (x_2), et de la promotion auprès des consommateurs (x_3) exprimées en milliers d'Euros sur les ventes en tonnes (y).

Pour ce faire, elle dispose, sur 10 ans, des séries statistiques du tableau 1.

-Pour répondre à cette question, on demande de calculer les coefficients de corrélation partielle du premier ordre (r_{yx_1, x_2}^2) et du deuxième ordre ($r_{yx_3, x_1 x_2}^2$) par les deux méthodes : corrélation entre les résidus et à partir de t de *student*.

Année	Y	X ₁	X ₂	X ₃
1	49	35	53	200
2	40	35	53	212
3	41	38	50	211
4	46	40	64	212
5	52	40	70	203
6	59	42	68	194
7	53	44	59	194
8	61	46	73	188
9	55	50	59	196
10	64	50	71	190

Données :

$$y = 3,5809 + 0,7809x_2 + e_1 \dots\dots\dots [1]$$

$$x_1 = 18,3809 + 0,3809x_2 + e_2 \dots\dots\dots [2]$$

$$y = -9,9621 + 0,7368x_1 + 0,50x_2 + e_1 \dots\dots [3]$$

$$x_3 = 258,932 + 0,952x_1 - 0,304x_2 + e_2 \dots\dots [4]$$

$$y = 134,2904 + 0,2062x_1 + 0,3308x_2 - 0,5573x_3 + e_1 \dots\dots [5]$$

$$t_{a_0}^* = 5,2952, t_{a_1}^* = 1,3213, t_{a_2}^* = 3,7701, t_{a_3}^* = -5,7930 \dots\dots \text{Pour le modèle [5]}$$

Exercice N° 4 :

Un économiste cherche à expliquer la variable (y) à l'aide de quatre séries explicatives x_1, x_2, x_3 et x_4 . Il désire auparavant tester une éventuelle multicolinéarité entre ces quatre séries : pour ce faire, il dispose des données du tableau 2.

-On demande d'appliquer les tests suivants :

- 1) test de Klein
- 2) test de Farrar-Glauber

Y	X ₁	X ₂	X ₃	X ₄
8,4	82,9	17,1	92	94
9,6	88,0	21,3	93	96
10,4	99,9	25,1	96	97
11,4	105,3	29	94	97
12,2	117,7	34	100	100
14,2	131,0	40	101	101
15,8	148,2	44	105	104
17,9	161,8	49	112	109
19,3	174,2	51	112	111
20,8	184,7	53	112	111

Données :

$$r_{x_1 x_2}^2 = 0,976, r_{x_1 x_3}^2 = 0,96, r_{x_1 x_4}^2 = 0,974$$

$$r_{x_2 x_3}^2 = 0,938, r_{x_2 x_4}^2 = 0,938, r_{x_3 x_4}^2 = 0,982$$

$$R^2 = 0,998, Xhi_{10}^2(0,05) = 18,31$$

$$y = -13,53 + 0,096x_1 + 0,015x_2 - 0,199x_3 + 0,34x_4 + e_1 \dots\dots [6]$$