Corrigé de la Série de TD N°1

Corrigé de l'Exercice 1

a) Estimation des paramètres du modèle de régression linéaire par la méthode des MCO.

La moyenne de X est notée M_X
La moyenne de Y est notée M_Y

	X	Y	X-M_X	Y-M_Y	(X-M_X)(Y-M_Y)	(X-M_X)(X-M_X)
1,00	4,00	2,00	-3,50	-3,00	10,5	12,25
2,00	6,00	3,00	-1,50	-2,00	3	2,25
3,00	8,00	5,00	0,50	0,00	0	0,25
4,00	12,00	10,00	4,50	5,00	22,5	20,25
Somme	30,00	20,00	0,00	0,00	**36**	**35,00**
Moyenne	7,50	5,00				

$\hat{a}_1 = \sum (X-M_X)(Y-M_Y) / \sum (X-M_X)^2 = \mathbf{36/35} = \mathbf{1{,}02}$

$\hat{a}_0 = M_Y - \hat{a}_1 * M_X = 5 - 1{,}02\,(7{,}5) = 5 - 7{,}56 = \mathbf{-2{,}65}$

a) Le coefficient $\hat{a}_1$ est –il significativement différent de 0 ?
Y_t estimé = $\hat{a}_0 + \hat{a}_1 X_t$ = **- 2,65** +1,02 X_t ; $e_t = Y_t - Y_t$ estimé

Exemple : Y_1 estimé = **- 2,65** +1,02 (4)= **1,43**

e_1= Y_1- Y_1estimé = 2-**1,43** = **0,57**.

Tableau – Calcul du résidus d'estimation

Y estimé	e	e*e
1,43	0,57	0,32
3,41	-0,41	0,17
5,53	-0,53	0,28
9,65	0,35	0,12
28,19	0,00	0,90

b) Tester de au seuil de 5% la significativité du coefficient associé à x_t

H0 = $\hat{a}_1 = 0$; H1= $\hat{a}_1 \neq 0$

Sous l'hypothèse H0, cette relation devient : $\frac{\hat{a}_1 - 0}{\hat{\sigma}_{\hat{a}_1}} = \frac{\hat{a}_1}{\hat{\sigma}_{\hat{a}_1}} = t^*_{\hat{a}_1} \rightarrow$ loi de Student à (n– 2) degrés de liberté.

$\widehat{\sigma}^2{}_{\hat{a}_1} = \frac{\widehat{\sigma}^2{}_{\varepsilon}}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$; Nous savons que $\widehat{\sigma}^2{}_{\varepsilon} = \frac{\sum_t e_t^2}{(n-2)}$

$\widehat{\sigma}^2{}_{\varepsilon} = \frac{\sum_t e_t^2}{(n-2)} = \frac{0{,}9}{2} = 0{,}45$

$\widehat{\sigma}^2{}_{\hat{a}_1} = \frac{0{,}45}{35} = 0{,}01$; $\sigma_{\hat{a}_1} = \sqrt{\widehat{\sigma}^2{}_{\hat{a}_1}} = \sqrt{0{,}01} = 0{,}1$

$t^* = 1{,}02/0{,}1 = \mathbf{10{,}2}$ $t^* = \mathbf{10{,}2} > T_{2.8}^{0.05} = \mathbf{4{,}30}$; Alors on rejette H_0 et on accepte H_1 ; Donc le coefficient $\hat{a}_1$ **est significativement différent de 0** ; la variable explicative X **contribue significativement** dans l'explication de la variable Y

Exercice 2 :

Nous avons $\delta(\varepsilon) = \sqrt{SCR/(n-2)} =$ 10,66 ⇒ SCR = (10,66)2 × 18 = 2 045,44

Nous pouvons calculer SCE et SCT à l'aide du coefficient de détermination.

R2 = 0,23 = 1 − SC R/SCT

SC T = SC R/(1 − R2) = 2045,44/(1 − 0,23) = 2 656,42

Or SCT = SCE + SCR ⇒ SCE = 610,98

Nous pouvons calculer maintenant $F^* = \frac{R^2}{(1-R^2)(n-2)} = \frac{SCE}{SCR/(n-2)} = 5{,}40$

Dans le cas d'un modèle de régression simple , $t^{*2} = F^*$; $t^* = \sqrt{F^*} = 2{,}32$

Nous avons $t^*{}_{a_1} = \frac{\hat{a}_1}{\sigma_{(\hat{a}_1)}}$, L'écart type du coefficient : $\sigma_{(\hat{a}_1)} = \frac{\hat{a}_1}{t^*{}_{a_1}} = \left|\frac{-32{,}95}{2{,}32}\right| = 14{,}20$

On pose le test d'hypothèses : H0 : $a_1 = 0$ contre l'hypothèse H_1 : $a_1 = \neq 0$

Sous H0, nous pouvons écrire $t^*{}_{a_1} = \frac{\hat{a}_1}{\sigma_{(\hat{a}_1)}} = \left|\frac{-32{,}95}{14{,}20}\right| = 2{,}32 > t^{5\%}{}_8 = 2{,}306$

Nous sommes donc dans la zone de refus de H_0, le coefficient a_1 est significativement différent de 0.

Exercice 3

1) **La propension marginale à consommer est-elle significativement différente de 0 ?**

$\hat{a}_1 = \frac{\sum_{t=1}^{n}(x_t-\bar{x})(y_t-\bar{y})}{\sum_{t=1}^{n}(x_t-\bar{x})^2} = \frac{50104729}{61156000} = 0{,}78$

$\hat{a}_0 = \bar{y} - \hat{a}_1\,\bar{x}$ = 9985,57- 0,78 (11280) = 1176, 08

Nous savons que : $\frac{\hat{a}_1 - a_1}{\widehat{\sigma}_{\hat{a}_1}}$ suit donc une loi de Student à (n– 2) degrés de liberté

Sous l'hypothèse H0, cette relation devient : $\frac{\hat{a}_1 - 0}{\hat{\sigma}_{\hat{a}_1}} = \frac{\hat{a}_1}{\hat{\sigma}_{\hat{a}_1}} = t^*_{\hat{a}_1}$ → loi de Student à (n– 2) degrés de liberté.

Nous savons que $\hat{\sigma}^2_{\hat{a}_1} = \frac{\hat{\sigma}^2_\varepsilon}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$, nous connaissons $\sum_{t=1}^{n}(x_t - \bar{x})^2 = 64156000$

On doit déterminer la valeur de $\hat{\sigma}^2_\varepsilon$. L'estimateur de la variance de l'erreur nous est donné comme suit : $\hat{\sigma}^2_\varepsilon = \frac{1}{(n-2)} \sum_t e_t^2$

Nous devons caliculer la valeur de e_t^2. Pour cela il faut déterminer la série ajustée $\hat{y}_t$

Calcul de $\hat{y}_t$ e_t

La série ajustée $\hat{y}_t$ est calculée par application des estimations $\hat{a}_0$ et $\hat{a}_1$

$\hat{y}_t = \hat{a}_0 + \hat{a}_1 x_t$, exemple : $\hat{y}_1$ = 1 176,08 + 0,78 × 8 000 = 7 423,95 ;

$e_t = y_t - \hat{y}_t$, exemple : $e_1 = y_1 - \hat{y}_{t1}$ = 7 389,99 – 7 423,95 = –33,96

Les résultats sont illustrés dans le tableau suivant :

Tableau – Calcul du résidu d'estimation

$\hat{y}_t$	e_t	e_t^2
7 423,95	– 33,96	1 153,38
8 204,93	– 35,28	1 244,98
8 595,43	236,28	55 830,26
8 595,43	57,41	3 296,40
8 829,72	– 41,64	1 733,93
9 766,90	– 150,69	22 707,42
10 547,88	45,57	2 076,39
11 328,87	– 142,76	20 379,08
12 890,83	– 132,74	17 620,12
13 671,81	197,81	39 127,38
Somme	0,00	165 169,3
Moyenne	0,00	16 516,93

Source : Régis Bourbonnais « *économétrie : cours et exercices corrigés*», 9ème édition Dunod, Paris, 2015

b) Calcul de l'estimation de la variance de l'erreur et de l'écart type du coefficient de régression.

L'estimation de la variance de l'erreur est donc égale à

$$\hat{\sigma}^2_\varepsilon = \frac{\sum_t e_t^2}{(n-2)} = \frac{165\,169,3}{8} = 20\,646,16$$

Ce qui nous permet de calculer la variance estimée de $\hat{a}_1$

$$\hat{\sigma}^2_{\hat{a}_1} = \frac{\hat{\sigma}^2_\varepsilon}{\sum_{t=1}^{n}(x_t - \bar{x})^2} = \frac{20\,646,16}{64\,156\,000} = 0,000\,321\,8 \text{ Donc } \sigma_{\hat{a}_1} = 0,017\,9$$

c) Calcul du ratio de *Student* et règle de décision.

Ce test permet de tester la pertinence d'une variable explicative qui figure dans un modèle et sa contribution à l'explication du phénomène que l'on cherche à modéliser. Dans notre exemple, nous calculons le ratio de Student :

Nous avons $t^*_{a_1} = \frac{\hat{a}_1}{\delta(\hat{a}_1)} = \frac{0,78}{0,0179} = 43,7 : \ t^*_{a_1} > t^{5/2\%}_8 = 2,306$

La propension marginale à consommer est donc significativement différente de 0, la variable

Revenu est bien explicative de la variable.

Détermination d'un intervalle de confiance, au seuil de 95 %, pour la propension marginale à consommer.

Nous savons que : $\frac{\hat{a}_1 - a_1}{\hat{\sigma}_{\hat{a}_1}}$ suit donc une loi de Student à (n– 2) degrés de liberté

L'intervalle de confiance est donné par : $\frac{\hat{a}_1 - a_1}{\hat{\sigma}_{\hat{a}_1}} \pm \ t^{\alpha/2}{}_{n-2} \rightarrow a_1 = \hat{a}_1 + \hat{\sigma}_{\hat{a}_1} \cdot t^{\alpha/2}{}_{n-2}$

Donc : $a_1 = 0{,}78 \pm 2{,}306\ (0{,}0179)$