

Solutions Exercices Régression Linéaires Série 3

D1 | $N=30$

Modèle : $Y_i = \alpha + \beta X_i + \epsilon_i, i=1, \dots, 30$

$E[\epsilon_i] = 0, \text{Var}(\epsilon_i) = \sigma_\epsilon^2, R = 0,5, \bar{X} = 7 \text{ et } \bar{Y} = 6$

* Calculons $\hat{\beta}_{MCO}$

On utilise, ici, la formule de R^2 :

$$R^2 = \frac{\sum (y_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$$

On sait que :

$$\begin{cases} \hat{y}_i = \hat{\alpha} + \hat{\beta} x_i \\ \bar{y} = \hat{\alpha} + \hat{\beta} \bar{x} \end{cases}$$

$$\Rightarrow R^2 = \frac{\sum (\hat{\alpha} + \hat{\beta} x_i - \hat{\alpha} - \hat{\beta} \bar{x})^2}{\sum (y_i - \bar{y})^2} = \hat{\beta}^2 \frac{\sum (x_i - \bar{x})^2}{\sum (y_i - \bar{y})^2}$$

$$\Rightarrow R^2 = \hat{\beta}^2 \cdot \frac{\bar{S}_x^2}{\bar{S}_y^2}$$

$$\Rightarrow \hat{\beta}^2 = R^2 \frac{\bar{S}_y^2}{\bar{S}_x^2} \Rightarrow \hat{\beta} = |R| \cdot \frac{\bar{S}_y}{\bar{S}_x} = 0,5 \times \frac{6}{7}$$

$$\Rightarrow \hat{\beta} \approx 0,4$$

Réponse A

De | $N=10$

$$\text{Modèle : } y_i = \alpha + \beta x_i + \varepsilon_i$$

$$\mathbb{E}[\varepsilon_i] = 0 ; \text{Var}(\varepsilon_i) = \sigma^2 \varepsilon_i, R^2 = 0.6, \sum y_i = 30$$

$$\sum y_i^2 = 690, \bar{y} = 3.$$

* Calculons s^2 (variance des erreurs)

On utilise son estimateur :

$$\hat{s}^2 = \frac{\sum (y_i - \hat{y}_i)^2}{n-2}$$

$$\text{On sait : } R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

$$\Rightarrow \sum (y_i - \hat{y}_i)^2 = (1 - R^2) \sum (y_i - \bar{y})^2$$

$$\Rightarrow \hat{s}^2 = (1 - R^2) \frac{\sum (y_i - \bar{y})^2}{n-2}$$

$$= 0.4 \times \left(\frac{\sum y_i^2 - 2\bar{y} \sum y_i + n\bar{y}^2}{n-2} \right)$$

$$= 0.4 \left(\frac{690 - 2 \times 3 \times 30 + 10 \times 9}{8} \right)$$

$$\Rightarrow s^2 = 30$$

Réponse B

D3

x : taux de chômage

y : durée de l'arrêt maladie

$T = 21$ ans ($= 1987 - 1967 + 1$)

$$\sum x_i = 1,5095, \quad \sum y_i = 1599,84, \quad \sum (x_i - \bar{x})^2 = 0,01491$$

$$\sum (y_i - \bar{y})^2 = 761,43, \quad \sum (x_i - \bar{x})(y_i - \bar{y}) = 2,8396$$

* Déterminons le coefficient de corrélation linéaire :

D'après ce qui précède :

$$R^2 = \rho^2 = \frac{\sum (x_i - \bar{x})^2}{\sum (y_i - \bar{y})^2}$$

$$\text{On sait que : } \rho_{(x,y)} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2}}$$

$$\Rightarrow R^2 = \left[\frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2}} \right]^2 \times \left(\frac{\sum (x_i - \bar{x})^2}{\sum (y_i - \bar{y})^2} \right)$$

$$= \frac{[\sum (x_i - \bar{x})(y_i - \bar{y})]^2}{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2}$$

$$\Rightarrow R = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \cdot \sqrt{\sum (y_i - \bar{y})^2}} = \frac{2,8396}{\sqrt{0,01491} \cdot \sqrt{761,43}}$$

$$\Rightarrow R = 0,842 \approx 0,84$$

Réponse E

D5

X: Dow Jones

Y: stock

T = 14 semaines

droite de régression par les MCO:

$$\hat{y}_i = -116,607 + 0,195X_i$$

$$\bar{\alpha} = -116,607 = \hat{\alpha} \quad \text{et} \quad 0,195 = \hat{\beta}$$

Statistiques:

$$\sum (x_i - \bar{x})^2 = 17807,080, \quad \sum (y_i - \bar{y})^2 = 949,388$$

* Déterminons R^2 (% de Y expliqués par $\Delta^p X$)
(pourcentage de variation de Y, expliqués par les variations de X.

$$\text{On sait: } R^2 = \hat{\beta}^2 \frac{\sum (x_i - \bar{x})^2}{\sum (y_i - \bar{y})^2} = (0,195)^2 \cdot \frac{17807,080}{949,388}$$

$$\Rightarrow R^2 = 0,7132127 \approx 0,71$$

Réponse c

D6

- Supposons qu'on a:

$$y_i = \alpha + \beta x_i + \epsilon_i, \quad i=1, \dots, N$$

$$\text{Par les MCO} \Rightarrow \hat{y}_i = \hat{\alpha} + \hat{\beta} x_i \quad (\text{droite de régression})$$

* Si la droite de régression n'est pas significative, cela veut dire que la relation entre X et Y n'est pas significative, ou encore les variations de X n'expliquent pas les variations de Y (4)

- Avec le test de Student, cela veut dire qu'on accepte que $\beta = 0$

- si donc $\beta = 0$, $\hat{\beta} = 0$ il vient:

$$\hat{y}_i = \bar{y} \text{ et } \bar{y} = \bar{d}$$

\Rightarrow la meilleure estimation des observations de y est sa propre moyenne: $\hat{y}_i = \bar{y}$.

C'est vraie

DT

$$X: 6,8 \quad 7,0 \quad 7,1 \quad 7,2 \quad 7,4$$

$$Y: 0,8 \quad 1,2 \quad 0,9 \quad 0,9 \quad 1,5$$

$$\bar{x} = 7,1, \quad \sum (y_i - \bar{y})^2 = 0,332, \quad \bar{y} = 1,06$$

$$\sum (x_i - \bar{x})^2 = 0,2, \quad \sum (x_i - \bar{x})(y_i - \bar{y}) = 0,18.$$

* Déterminons R^2

$$R^2 = \beta^2 \cdot \frac{\sum (x_i - \bar{x})^2}{\sum (y_i - \bar{y})^2} = \left[\frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \right]^2 \cdot \frac{\sum (x_i - \bar{x})^2}{\sum (y_i - \bar{y})^2}$$

$$R^2 = \frac{[\sum (x_i - \bar{x})(y_i - \bar{y})]^2}{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2} = \frac{(0,18)^2}{0,332 \times 0,2} = 0,4879$$

$$R^2 \approx 0,5$$

\Rightarrow Réponse C

D10

X: Revenu

$$\bar{X} = 27,5$$

Y: Epargne

$$\bar{Y} = 30$$

$$N = 6$$

$$\sum X^2 = 87,5, \quad \sum Y^2 = 35, \quad \sum XY = 17$$

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2$$

* Déterminons la proportion de variations de Y autour de sa moyenne, expliqués par le modèle:

$$R^2 = \beta^2 \frac{\sum (x_i - \bar{x})^2}{\sum (y_i - \bar{y})^2} = \frac{[\sum (x_i - \bar{x})(y_i - \bar{y})]^2}{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}$$

$$\Rightarrow R^2 = \frac{[\sum_{XY}^2]^2}{\sum X^2 \cdot \sum Y^2} = \frac{(17)^2}{87,5 \times 35} = 0,94 \approx 0,9$$

Réponse E

D12

Y: 1 2 3 4 5, N = 5 obs

$$\bar{Y} = 3; \quad \sum Y_i = 15$$

$$s^2 =$$

$$\frac{\sum (y_i - \hat{y}_i)^2}{n-2} = 1$$

$$\text{ou } y_i = \alpha + \beta x_i + \epsilon_i$$

* Calculus R²

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{(n-2)}{\sum y_i^2 - 2\bar{y} \sum y_i + N \cdot \bar{y}^2}$$

$$R^2 = 1 - \frac{3}{10} = 0,7$$

Réponse E

(6)

A2

Y : nombre d'accidents auto mortel

X_1 : précipitation en pouce

X_2 : volume du trafic en milliers de voitures

$X_3 := \begin{cases} 1 & \text{si WE de vacances se produit dans le mois} \\ 0 & \text{sinon,} \end{cases}$

Modèle: $Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$

$$E[\varepsilon] = 0_{12}, \quad V[\varepsilon] = \sigma^2$$

Statistiques descriptives:

$$(X'X)^{-1} = \frac{1}{6506} \begin{pmatrix} 257 & -82 & -446 \\ -82 & 254 & -364 \\ -446 & -364 & 2622 \end{pmatrix}, \quad S^2 = 1,45241$$

- Déterminons $\hat{\beta}_1$

On utilise l'estimateur de MV de $\beta' = (\beta_1, \beta_2, \beta_3)$, on sait:

$$\hat{\beta}_{MV} = (X'X)^{-1} X'Y$$

$$\text{ou } X'Y = \begin{pmatrix} \sum x_{1i} Y_i \\ \sum x_{2i} Y_i \\ \sum x_{3i} Y_i \end{pmatrix}$$

- Par simple calculs, on a:

$$\sum x_{1i} Y_i = 57, \quad \sum x_{2i} Y_i = 51 \quad \text{et} \quad \sum x_{3i} Y_i = 20$$

- Il vient,

$$\hat{\beta}_{MV} = \frac{1}{6506} \begin{pmatrix} 257 & -82 & -446 \\ -82 & 254 & -364 \\ -446 & -364 & 2622 \end{pmatrix} \begin{pmatrix} 57 \\ 51 \\ 20 \end{pmatrix} = \begin{pmatrix} 0,2377 \\ 0,1537 \\ 1,299 \end{pmatrix}$$

$$\Rightarrow \hat{\beta}_1 = 0,2377 \approx 0,24, \quad \underline{\text{Réponse C}}$$

(7)

A4 Modèle: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \varepsilon$

$Y : 1 \quad 2 \quad 4 \quad 4 \quad 5$

$X_1 : 4 \quad 5 \quad 5 \quad 6 \quad 7$

$X_2 : 1 \quad 1 \quad 2 \quad 3 \quad 3$

$X_1 X_2 : 4 \quad 5 \quad 10 \quad 18 \quad 21$

$E[\varepsilon] = 0, \quad V[\varepsilon] = \sigma^2$

stats: $(X'X)^{-1} = \begin{pmatrix} 87,9 & -18,7 & -32,8 & 6,8 \\ -18,7 & 4,2 & 6,4 & -1,4 \\ -32,8 & 6,4 & 17,1 & -3,1 \\ 6,8 & -1,4 & -3,1 & 6,6 \end{pmatrix}$

- Déterminons $\hat{\beta}_3$

L'estimateur des maximum de vraisemblance de $\beta' = (\beta_0, \beta_1, \beta_2, \beta_3)$

est: $\hat{\beta}_{ML} = (X'X)^{-1} X'Y$

$X'Y = \begin{pmatrix} \sum Y_i \\ \sum X_{1i} Y_i \\ \sum X_{2i} Y_i \\ \sum X_{1i} X_{2i} Y_i \end{pmatrix} = \begin{pmatrix} 16 \\ 93 \\ 38 \\ 231 \end{pmatrix}$

Il vient, $\hat{\beta}_{ML} = \begin{pmatrix} 87,9 & -18,7 & -32,8 & 6,8 \\ -18,7 & 4,2 & 6,4 & -1,4 \\ -32,8 & 6,4 & 17,1 & -3,1 \\ 6,8 & -1,4 & -3,1 & 6,6 \end{pmatrix} \begin{pmatrix} 16 \\ 93 \\ 38 \\ 231 \end{pmatrix} = \begin{pmatrix} -8,3 \\ 1,9 \\ 4,2 \\ -0,6 \end{pmatrix}$

Donc $\hat{\beta}_3 = -0,6 \Rightarrow$ Réponse D

" Régression Linéaire Multiple " " Série 1, Série 2 "

A7

Modelle: $Y = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i$

$i : 1 \quad 2 \quad 3 \quad 4 \quad 5$

$Y : 1,3 \quad 1,5 \quad 1,8 \quad 1,6 \quad 1,7$

$X_1 : 6 \quad 7 \quad 7 \quad 8 \quad 8$

$X_2 : 4,5 \quad 4,6 \quad 4,5 \quad 4,7 \quad 4,6$

$$(X'X)^{-1} = \begin{pmatrix} 1522,73 & 26,87 & -374,67 \\ 26,87 & 0,93 & -7,33 \\ -374,67 & -7,33 & 93,33 \end{pmatrix}$$

Calculons $\hat{\epsilon}_2$

On sait: $\hat{\epsilon}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{i1} - \hat{\beta}_2 X_{i2}$

Pour $i=2 \Rightarrow \hat{\epsilon}_2 = Y_2 - \hat{\beta}_0 - \hat{\beta}_1 X_{21} - \hat{\beta}_2 X_{22}$

- Trouvons $\hat{\beta}_0$, $\hat{\beta}_1$ et $\hat{\beta}_2$

On utilise la formule matricielle de l'EMV de $\beta = (\beta_0, \beta_1, \beta_2)$

$$\Rightarrow \hat{\beta}_{EMV} = (X'X)^{-1} X'Y \quad \text{ou} \quad X'Y = \begin{pmatrix} \sum Y_i \\ \sum X_{i1} Y_i \\ \sum X_{i2} Y_i \end{pmatrix} = \begin{pmatrix} 7,9 \\ 57,3 \\ 36,19 \end{pmatrix}$$

$$\Rightarrow \hat{\beta}_{EMV} = \begin{pmatrix} 1522,73 & 26,87 & -374,67 \\ 26,87 & 0,93 & -7,33 \\ -374,67 & -7,33 & 93,33 \end{pmatrix} \begin{pmatrix} 7,9 \\ 57,3 \\ 36,19 \end{pmatrix} = \begin{pmatrix} 9,9107 \\ 0,2893 \\ -2,2893 \end{pmatrix}$$

Donc,

$$\hat{\epsilon}_2 = 1,5 - (9,9107 + 0,2893 \times 7 - 2,2893 \times 4,6)$$

$$\Rightarrow \hat{\epsilon}_2 = 0,09498 \approx 0,1$$

Réponse D

(9)

A12 / Modèle: $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i, i=1,2,3,4$

$i: 1 \quad 2 \quad 3 \quad 4$

$X_{2i}: -3 \quad -1 \quad 1 \quad 3$

$X_{3i}: -1 \quad 3 \quad -3 \quad 1$

$$\hat{\beta}_3 = \sum_{i=1}^4 w_i Y_i$$

- Déterminons $w = (w_1, w_2, w_3, w_4)$

On sait: $\hat{\beta}_{OLS} = (X'X)^{-1} X'Y$

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = \begin{pmatrix} n & \sum x_{i2} & \sum x_{i3} \\ \sum x_{i2} & \sum x_{i2}^2 & \sum x_{i2} x_{i3} \\ \sum x_{i3} & \sum x_{i2} x_{i3} & \sum x_{i3}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum Y_i \\ \sum x_{2i} Y_i \\ \sum x_{3i} Y_i \end{pmatrix}$$

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 20 \end{pmatrix}^{-1} \begin{pmatrix} \sum Y_i \\ \sum x_{2i} Y_i \\ \sum x_{3i} Y_i \end{pmatrix}$$

$\det(X'X) = 4 \times 20 \times 20 = 1600 \neq 0$

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = \frac{1}{\det(X'X)} \begin{bmatrix} \begin{vmatrix} 20 & 0 \\ 0 & 20 \end{vmatrix} & - \begin{vmatrix} 0 & 0 \\ 0 & 20 \end{vmatrix} & \begin{vmatrix} 0 & 20 \\ 0 & 0 \end{vmatrix} \\ - \begin{vmatrix} 0 & 0 \\ 0 & 20 \end{vmatrix} & \begin{vmatrix} 4 & 0 \\ 0 & 20 \end{vmatrix} & - \begin{vmatrix} 4 & 0 \\ 0 & 0 \end{vmatrix} \\ \begin{vmatrix} 0 & 0 \\ 20 & 0 \end{vmatrix} & - \begin{vmatrix} 4 & 0 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 4 & 0 \\ 0 & 20 \end{vmatrix} \end{bmatrix}^T \begin{pmatrix} \sum Y_i \\ \sum x_{2i} Y_i \\ \sum x_{3i} Y_i \end{pmatrix}$$

(11)

Response B

$$\begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = \begin{pmatrix} 0.05 \times (-1) \\ 0.05 \times (3) \\ 0.05 \times (-3) \\ 0.05 \times (1) \end{pmatrix} = \begin{pmatrix} -0.05 \\ 0.15 \\ -0.15 \\ 0.05 \end{pmatrix}$$

$$w_1 = 0.05 \times (-1) \quad w_2 = 0.05 \times (3) \quad w_3 = 0.05 \times (-3) \quad w_4 = 0.05 \times (1)$$

Den identifikationsmatrix

$$\begin{aligned} \sum_{i=1}^4 p_i &= \sum_{i=1}^4 w_i^2 y_i^2 = \sum_{i=1}^4 (0.05 \times x_{3i} y_i)^2 \\ &= \sum_{i=1}^4 0.0025 x_{3i}^2 y_i^2 = 0.0025 \sum_{i=1}^4 x_{3i}^2 y_i^2 \end{aligned}$$

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} = \begin{pmatrix} 0.05 & 0 & 0 \\ 0 & 0.05 & 0 \\ 0 & 0 & 0.05 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \sum y_i \\ \sum x_{2i} y_i \\ \sum x_{3i} y_i \end{pmatrix}$$

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \frac{1}{1600} \begin{pmatrix} 400 & 0 & 0 \\ 0 & 80 & 0 \\ 0 & 0 & 80 \end{pmatrix} \begin{pmatrix} \sum y_i \\ \sum x_{2i} y_i \\ \sum x_{3i} y_i \end{pmatrix}$$

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \frac{1}{1600} \begin{pmatrix} 400 & 0 & 0 \\ 0 & 80 & 0 \\ 0 & 0 & 80 \end{pmatrix}^T \begin{pmatrix} \sum y_i \\ \sum x_{2i} y_i \\ \sum x_{3i} y_i \end{pmatrix}$$

B1 | Modèle : $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$

$E[\varepsilon_i] = 0$, $V[\varepsilon_i] = \sigma^2$

$CS^2 = \begin{pmatrix} 900 & -144 & 225 \\ -144 & 1600 & 400 \\ 225 & 400 & 625 \end{pmatrix}$: matrice variance-covariance de $\hat{\beta}_{MV}$.

"En cours" $CS^2 = \sigma^2 (X'X)^{-1}$

$CS^2 = (X'X)^{-1} S^2$

$\Rightarrow \hat{\beta}_2 = 40$

- calculons la statistique de student:

$\begin{cases} H_0 : \beta_2 = 0 \\ H_1 : \beta_2 \neq 0 \end{cases} \sim T_{(n-3, \alpha/2)}$, $n =$ nombre observation

$\Rightarrow t_{cal} = \frac{\hat{\beta}_2}{\hat{\sigma}_{\hat{\beta}_2}}$ où $\hat{\sigma}_{\hat{\beta}_2}$: écart-type estimé de $\hat{\beta}_2$

- Trouvons $\hat{\sigma}_{\hat{\beta}_2}$

On sait $CS^2 = \begin{pmatrix} \hat{\beta}_0 & \hat{\beta}_1 & \hat{\beta}_2 \\ \text{var}(\hat{\beta}_0) & \text{cov}(\hat{\beta}_0, \hat{\beta}_1) & \text{cov}(\hat{\beta}_0, \hat{\beta}_2) \\ \dots & \text{var}(\hat{\beta}_1) & \text{cov}(\hat{\beta}_1, \hat{\beta}_2) \\ \dots & \dots & \text{var}(\hat{\beta}_2) \end{pmatrix}$ (symétrique)

On voit que : $\text{var}(\hat{\beta}_2) = 625$

$\Rightarrow \hat{\sigma}_{\hat{\beta}_2} = \sqrt{625} = 25$

Donc , $t_{cal} = \frac{40}{25} = 1,6$ | Réponse E |

On a de regression obtenu:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} \quad n=12$$

$$\Rightarrow \hat{y}_t = 1,2360 + 0,18663 x_{1t} + 0,0557 x_{2t}$$

$$H_0: \beta_1 = \beta_2 = 0$$

$$H_1: \beta_1 \neq 0 \vee \beta_2 \neq 0$$

Stat	9,19
SCR = 9,19	
SCF = 18,39	
SCF = 9,19	

$$(X'X)^{-1} = \begin{pmatrix} 7,70997 & -0,82314 & -1,05159 \\ -0,82314 & 0,10044 & 0,10480 \\ -1,05159 & 0,10480 & 0,15284 \end{pmatrix}$$

Determinant Δ :

$$IC_{q5/10}(\beta_0) = \hat{\beta}_0 \pm t_{\alpha/2}$$

$$IC_{q5/10}(\beta_2) = 0,18517 \pm t_{\alpha/2}$$

$$\text{On veut: } IC_{q5/10}(\beta_2) = 0,18517 \pm t_{(n-3; \alpha/2)} \times \frac{\hat{\sigma}_{\beta_2}}{\sqrt{a_{22}}}$$

$$\Rightarrow t = T_{(n-3; \alpha/2)} \times \frac{\hat{\sigma}_{\beta_2}}{\sqrt{a_{22}}} \quad \text{On prend } \alpha = 0,05$$

Il suffit de trouver $\hat{\sigma}_{\beta_2}$ et $\sqrt{a_{22}}$ (n-3; 0,025)

$$\hat{\sigma}_{\beta_2}^2 = \frac{SCE}{n-3} = \frac{9}{9}$$

(14)

- Matrice de covariance de β

$$\text{Var}(\beta_j) = \sigma^2 \cdot (X'X)^{-1}$$

Il vient, $\text{Var}(\beta_j) = \frac{\sigma}{\text{SCE}} (X'X)^{-1}$

ou encore $\frac{\sigma}{\text{SCE}} (X'X)^{-1}$

$$\Rightarrow \frac{\sigma}{\text{SCE}} \sqrt{(X'X)^{-1}}$$

Pour $j=2$, $\sqrt{\frac{\sigma}{\text{SCE}}} = \sqrt{(X'X)^{-1}}$

On a : $(X'X)^{-1} = 0,15284$ et $\sqrt{(X'X)^{-1}} = 0,39094757$
 $\sqrt{\frac{\sigma}{\text{SCE}}} = \sqrt{0,2082111} = 0,45630157$

$$\Rightarrow \frac{\sigma}{\text{SCE}} = (0,145630157) \times (0,39094757) = 0,1783$$

On soit : $T(9; 0,025) = 2,26215716$

Il vient, $t = 0,1783 \times 2,26215716 = 0,40354618$

$$t \approx 0,140$$

Réponse :

Modelle:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

SS/DF	21	51.25	= n = 6
SCR = 42	2		
SCE = 12	3		
SCT = 54	5		

- Calculons la statistique de Student:

$$(X'X)^{-1} = \begin{pmatrix} 4/3 & -1/4 & -1/3 \\ -1/4 & 1/6 & 0 \\ -1/3 & 0 & 2/3 \end{pmatrix}$$

$H_0: \beta_3 = 1$
 $H_1: \beta_3 \neq 1$
 $\alpha = 0.05$

$$\Rightarrow |t_{\text{calc}}| = \left| \frac{\hat{\beta}_3 - 1}{\hat{\sigma}_{\hat{\beta}_3}} \right| = \frac{2}{3}$$

On sait: $\text{Var}(\hat{\beta}_3) = \frac{\text{SCE}}{3} (X'X)^{-1}_{33}$

Pour $df = 3$, $\text{Var}(\hat{\beta}_3) = \frac{\text{SCE}}{3} (X'X)^{-1}_{33}$

$$\text{Var}(\hat{\beta}_3) = \frac{12}{3} \times \frac{2}{3} = \frac{8}{3}$$

$$\Rightarrow \hat{\sigma}_{\hat{\beta}_3} = \sqrt{\frac{8}{3}} = 1.632993162$$

$$t_{\text{calc}} = \frac{2}{3} = 1.837$$

Alors

$$t_{(3; 0.05)} = 3.1824463$$

$$|t_{\text{calc}}| = 1.8 < t_{(3; 0.05)} \Rightarrow \text{on rejette } H_0$$

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \epsilon_i$$

$n=30$

$$(X'X)^{-1} = \begin{pmatrix} 2,8195 & -0,0286 & -0,0788 & -0,0263 \\ -0,0286 & 0,0027 & 0,0010 & -0,0014 \\ -0,0788 & 0,0010 & 0,0035 & -0,0010 \\ -0,0263 & -0,0014 & -0,0010 & 0,0032 \end{pmatrix}$$

$$Y'Y = 7,095 \quad X'Y = \begin{pmatrix} 261,5 \\ 404,5 \\ 617,5 \\ 5707,0 \end{pmatrix}$$

$$\beta_{OLS} = \begin{pmatrix} 5,22 \\ 1,62 \\ 0,12 \\ -0,15 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}$$

- Déterminer la longueur de IC 95% (β_1) $\Rightarrow \alpha = 1 - 95\% = 5\%$.

On pose $\left\{ \begin{array}{l} \text{Binf} = \beta_1 - T(n-4; 0,025) \times \hat{\sigma}_{\beta_1} \\ \text{Bsup} = \beta_1 + T(n-4; 0,025) \times \hat{\sigma}_{\beta_1} \end{array} \right.$ Soit L la longueur de cet intervalle

$$L = \text{Bsup} - \text{Binf}$$

- Trouver $\hat{\sigma}_{\beta_1}^2 = \text{Var}(\hat{\beta}_1) = \hat{\sigma}^2 (X'X)^{-1}_{11}$
 On sait

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \epsilon_i$$

2) $N=15$

ii) $(X'X)^{-1} =$

$$\begin{pmatrix} 12.166 & -0.133 & 2.105 & -6.131 \\ -0.133 & 0.103 & 0.111 & 0 \\ 2.105 & 0.111 & 2.114 & -2.152 \\ -6.131 & 0 & -2.152 & 4.132 \end{pmatrix}$$

iii) $SCE = 282.82$

- Calculus of $(\beta_3 - \beta_2)$

On self $\hat{\beta}_2 = \frac{SCE}{(X'X)^{-1}}$

four $k=4$ $\hat{\beta}_2 = \frac{SCE}{(X'X)^{-1}}$

$$SCE = \frac{11}{282.82} = 25.71$$

$$\hat{\beta}_2 = 25.71 \cdot (X'X)^{-1}$$

four $\hat{\beta} = 2$ $\hat{\beta}_2 = 25.71 \times (X'X)^{-1} = 25.71 \times 0.03 = 0.7713$

four $\hat{\beta} = 3$ $\hat{\beta}_2 = 25.71 \times (X'X)^{-1} = 25.71 \times 2.14 = 55.2213$

(17)

On each que: (Repeat)

$$\text{Var}(ax + by) = a^2 \text{Var}(x) + b^2 \text{Var}(y) + 2 \text{Cov}(x, y)$$

Given $X = \beta_3$ at $Y = \beta_2$

$a = 1$ at $b = -1$

On 1:

$$\text{Var}(\beta_3 - \beta_2) = \text{Var}(a\beta_3 + b\beta_2)$$

$$= a^2 \text{Var}(\beta_3) + b^2 \text{Var}(\beta_2) + 2 \text{Cov}(\beta_3, \beta_2)$$

$$= \text{Var}(\beta_3) + \text{Var}(\beta_2) + 2 \text{Cov}(\beta_3, \beta_2)$$

$$= \sigma_{\beta_3}^2 + \sigma_{\beta_2}^2 + 2 \text{Cov}(\beta_3, \beta_2)$$

$$= \sigma_{\beta_3}^2 + \sigma_{\beta_2}^2 + 2 [\text{Cov}(\beta_3, \beta_2) - \text{Cov}(\beta_3, \beta_2)]$$

$$\sigma_{\beta_3 - \beta_2}^2 = \sigma_{\beta_3}^2 + \sigma_{\beta_2}^2 - 2 [\text{Cov}(\beta_3, \beta_2) - \text{Cov}(\beta_3, \beta_2)]$$

C5

$$Y = \alpha + \beta X$$

$n=6$

$$X : 7 \quad 14 \quad 15,5 \quad 25,1 \quad 21,4 \quad 32,8$$

$$Y : 14,2 \quad 13,4 \quad 13,2 \quad 26,7 \quad 17,7 \quad 25,5$$

$$\hat{Y} : 11,3 \quad 15,4 \quad 16,2 \quad 21,8 \quad 19,7 \quad 26,3$$

$$\bar{Y} = 18,45$$

Déterminons la statistique de Fisher

$$\left. \begin{array}{l} H_0 : \beta = 0 \quad (Y = \alpha \text{ : modèle réduit})^{(R)} \\ H_1 : \beta \neq 0 \quad (Y = \alpha + \beta X \text{ : modèle complet})^{(C)} \end{array} \right\} \quad (R) \quad (C)$$

$$F_{\text{cal}} = \frac{(SCE_R - SCE_C) / [(n-1) - (n-2)]}{SCE_C / (n-2)}$$

→ On sait que :

$$* SCE_C = \sum_{i=1}^6 \hat{\epsilon}_i^2 = \sum_{i=1}^6 (y_i - \hat{y}_i)^2$$

$$= (14,2 - 11,3)^2 + (13,4 - 15,4)^2 + (13,2 - 16,2)^2 \\ + (26,7 - 21,8)^2 + (17,7 - 19,7)^2 + (25,5 - 26,3)^2$$

$$SCE_C = 50,06$$

$$* SCE_R = \sum_{i=1}^6 \hat{\epsilon}_i^2 = \sum_{i=1}^6 (y_i - \bar{y})^2 = \sum_{i=1}^6 (y_i - 18,45)^2$$

B9

i) Y : nombre annuel de décès dans un hôpital

ii) X : nombre de litte dans l'hôpital

iii) $D = \begin{cases} 1 & \text{si hôpital = privé} \\ 0 & \text{si hôpital = public} \end{cases}$

iv) $Y = \beta_1 + \beta_2 X + \beta_3 D$

v) $\text{Var}(\hat{\beta}) = \sigma^2 (X'X)^{-1} = \begin{pmatrix} 1,89952 & -0,00364 & -0,82744 \\ -0,00364 & 0,0001 & -0,0041 \\ -0,82744 & -0,0041 & 8,79655 \end{pmatrix}$

\Rightarrow Déterminons $\sigma^2(\hat{\beta}_1 + 600\hat{\beta}_2)$

D'après B8,

$$\sigma^2(\hat{\beta}_1 + 600\hat{\beta}_2) = \sqrt{\sigma_{\hat{\beta}_1}^2 + (600)^2 \sigma_{\hat{\beta}_2}^2 + 2 \text{Cov}(\hat{\beta}_1, \hat{\beta}_2)}$$

(*) $\sigma_{\hat{\beta}_1}^2 = 1,89952$, $\sigma_{\hat{\beta}_2}^2 = 0,0001$

$$\Rightarrow \sigma_{\hat{\beta}_1 + 600\hat{\beta}_2}^2 = \sqrt{1,89952 + 360000 \times 0,0001 + 2 \text{Cov}(\hat{\beta}_1, \hat{\beta}_2)}$$

03

$$* S_{CE} = (14^2 - 18^2) + \dots + (25^2 - 18^2) = 1891455$$

$$\Rightarrow F_{cal} = \frac{1891455}{(1891455 - 50106) / 1} = \frac{50106 / 4}{1113821}$$

$$F_{cal} \approx 11$$

Réponse D

DL
 SQR = 115,11 : nombre de variables explicatives
 SCE = 138,89 : nombre d'observations - nombre de paramètres
 SCT = 1254,00 : nombre d'observations - 1

- Déterminer \bar{R}^2 (R^2 ajusté)

$$\bar{R}^2 = 1 - (1 - R^2) \times \left(\frac{n-1}{n-k}\right)$$

$$R^2 = \frac{SQR}{SCT} = \frac{115,11}{1254,00} = 0,089$$

$$\bar{R}^2 = 1 - (1 - 0,089) \times \left(\frac{n-1}{n-k}\right)$$

Par lecture du tableau d'analyse de la variance

$$DL_{SCE} = n - k = 5$$

$$DL_{SCE} = n - 1 = 7$$

$$R^2 = 1 - (1 - 0,089) \times \frac{5}{7} \approx 0,8449$$

$$\Rightarrow \bar{R}^2 \approx 0,84$$

Réponse A

(20)

Éléments de Corrections série 3

Régression Multiple

[A2]

$$Y_i = \beta \ln x_i + \epsilon_i, \quad i=1,2$$

Par application de MCO:

$$\text{Min}_{\beta} \sum_{i=1}^2 \hat{\epsilon}_i^2 = \text{Min}_{\beta} \sum_{i=1}^2 (Y_i - \beta \ln x_i)^2$$

Conditions du 1^{er} ordre:

$$\frac{\partial \sum_{i=1}^2 \hat{\epsilon}_i^2}{\partial \beta} = 0 \Rightarrow \partial \sum_{i=1}^2 \ln x_i (Y_i - \beta \ln x_i) = 0$$

$$\sum_{i=1}^2 Y_i \ln x_i - \beta \sum_{i=1}^2 (\ln x_i)^2 = 0$$

$$\Rightarrow \hat{\beta} = \frac{\sum_{i=1}^2 Y_i \ln x_i}{\sum_{i=1}^2 (\ln x_i)^2} = \frac{Y_1 \ln x_1 + Y_2 \ln x_2}{(\ln x_1)^2 + (\ln x_2)^2}$$

Si $x_1 = e$ et $x_2 = e^2$, il vient:

$$\hat{\beta} = \frac{Y_1 \ln e + Y_2 \ln e^2}{(\ln e)^2 + (\ln e^2)^2} = \frac{Y_1 + 2Y_2}{1+4} = \frac{Y_1 + 2Y_2}{5}$$

\Rightarrow Réponse c

(2)

A4

$$Y_i = \alpha \cdot \beta^{x_i} u_i$$

- Modèle non-linéaire.

On linéarise ce modèle par passage au logarithme:

$$\ln Y_i = \ln \alpha + x_i \ln \beta + \ln u_i$$

On pose:

$$Y'_i = \ln Y_i$$

$$\alpha' = \ln \alpha$$

$$\beta' = \ln \beta$$

$$u'_i = \ln u_i$$

$$\Rightarrow Y'_i = \alpha' + \beta' x_i + u'_i, \quad u'_i = \text{terme des erreurs.}$$

MCO: $\Leftrightarrow \min_{\beta} \sum u_i^2 = \min_{\beta} \sum (Y'_i - \alpha' - \beta' x_i)^2$

Posons $S = \sum (Y'_i - \alpha' - \beta' x_i)^2$
 $= \sum (\ln Y_i - \ln \alpha - x_i \ln \beta)^2$

C.T.O.: $\frac{\partial S}{\partial \beta} = 0 \Rightarrow \ln \beta = \frac{\sum x_i \ln Y_i - \sum \ln Y_i \cdot \bar{x}}{\sum x_i^2 - n \cdot \bar{x}^2}$

$$\Rightarrow \ln \beta = \frac{\sum x_i \ln Y_i - (\sum \ln Y_i) \bar{x}}{\sum x_i^2 - n \bar{x}^2}$$

Ans

X_2	X_1	Y_i	Z_i
2	1	4	1
3	2	9	2
1	6	36	3
5	8	64	4
6	4	16	5

$$\sqrt{Y_i} = \frac{\sqrt{b_0 + b_1 X_1 + b_2 X_2 + \epsilon_i}}{1}$$

Reference

\Rightarrow On prend l'exponentielle,
 $\ln b = e^{0.1324} \approx e^{0.132}$

$\ln b = \frac{52.96 - 8.61 \times 4.5}{8 \times (4.5)^2} = \frac{13.615}{42} = 0.324$

$\sum_{i=1}^n X_2 = 204$; $\sum_{i=1}^n \ln Y_i = 8.61$

$\bar{X}_2 = \frac{\sum_{i=1}^n X_2}{n} = \frac{204}{8} = 25.5$

$\sum_{i=1}^n X_1 \ln Y_i = 52.96$