Coumba SY – Master 1 IA

<u>Exercice 1 :</u>

$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \varepsilon_i$, $1 \leq i \leq n$

$Y = X\beta + \varepsilon_i$

$x_{i,j}$:variables exogénes du modèle

ε_i : variables aléatoires indépendantes

ε_i suit une loi norma centrée admettant la même variance σ^2

$$X = \begin{bmatrix} 1 & x1,2 & x1,3 \\ \ldots & \ldots & \ldots \\ 1 & xn,2 & xn,3 \end{bmatrix} \quad \text{et } Y = \begin{bmatrix} y1 \\ \ldots \\ yn \end{bmatrix}$$

$$X'X = \begin{bmatrix} 30 & 20 & 0 \\ 20 & 20 & 0 \\ 0 & 0 & 10 \end{bmatrix}, \quad X'Y = \begin{bmatrix} 15 \\ 20 \\ 10 \end{bmatrix}, \quad Y'Y = 59.5$$

Question 1

- Valeur de n et $\bar{x}_{i,2}$

$$X = \begin{bmatrix} 1 & x1,1 & x1,2 \\ 1 & x2,1 & x2,2 \\ \ldots & \ldots & \ldots \\ 1 & xn,1 & xn,2 \end{bmatrix}$$

$$X' = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ x1,1 & x2,1 & \ldots & xn,1 \\ x1,2 & x2,2 & \ldots & xn,2 \end{bmatrix}$$

$$X'X = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ x1,1 & x2,1 & \ldots & xn,1 \\ x1,2 & x2,2 & \ldots & xn,2 \end{bmatrix} * \begin{bmatrix} 1 & x1,1 & x1,2 \\ 1 & x2,1 & x2,2 \\ \ldots & \ldots & \ldots \\ 1 & xn,1 & xn,2 \end{bmatrix}$$

$$= \begin{pmatrix} n & \sum_{i=1}^{n} xi,1 & \sum_{i=1}^{n} xi,2 \\ \sum_{i=1}^{n} xi,1 & \sum_{i=1}^{n} x^2i,1 & \sum_{i=1}^{n} xi,1xi,2 \\ \sum_{i=1}^{n} xi,2 & \sum_{i=1}^{n} xi,1xi,2 & \sum_{i=1}^{n} x^2i,1 \end{pmatrix}$$

Donc par identification avec X'X = $\begin{bmatrix} 30 & 20 & 0 \\ 20 & 20 & 0 \\ 0 & 0 & 10 \end{bmatrix}$

On obtient : $\boxed{n = 30}$

Et $\boxed{\bar{x}_{i,2} = 0}$

- Coefficient de corrélation des variables $x_{i,1}$ et $x_{i,2}$

$$r = \frac{\sum_{i=1}^{n}(xi,1-\bar{x}i,1)(xi,2-\bar{x}i,2\ )}{\sqrt{\sum_{i=1}^{n}(xi,1-\bar{x}i,1)^{\ 2}}\ \sqrt{\sum_{i=1}^{n}(xi,2-\bar{x}i,2\ )^{2}}}$$

$$= \frac{\sum_{i=1x}^{n}\ xi,2(xi,1-\bar{x}i,1)\ -\sum_{i=1}^{n}\bar{x}i,2(xi,1-\bar{x}i,1)}{\sqrt{\sum_{i=1}^{n}(xi,1-\bar{x}i,1)^{\ 2}}\ \sqrt{\sum_{i=1}^{n}x^2i,2-n\bar{x}^{\ 2}i,2\ -2\sum_{i=1}^{n}\ xi,2\bar{x}i,2}}$$

$$= \frac{\sum_{i=1x}^{n} xi,2xi,1-\sum_{i=1}^{n} xi,2\bar{x}i,1-\ \sum_{i=1}^{n}\ \bar{x}i,2xi,1+\ \sum_{i=1}^{n}\ \bar{x}i,2\ \ \bar{x}i,1}{\sqrt{\sum_{i=1}^{n}(xi,1-\bar{x}i,1)^{\ 2}}\ \sqrt{\sum_{i=1}^{n}x^2i,2}}$$

Or $\sum_{i=1}^{n} xi,2 = 0$; $\bar{x}_{i,2} = 0$

$$= \frac{\sum_{i=1x}^{n} xi,2xi,1}{\sqrt{\sum_{i=1}^{n}(xi,1-\bar{x}i,1)^{\ 2}}\ \sqrt{\sum_{i=1}^{n}x^2i,2}}$$

$\boxed{r = 0}$

Question 2

- **Valeurs de β_0, β_1, β_2**

$$\det(X'X) = \begin{vmatrix} 30 & 20 & 0 \\ 20 & 20 & 0 \\ 0 & 0 & 10 \end{vmatrix}\begin{matrix} 30 & 20 \\ 20 & 20 \\ 0 & 0 \end{matrix}$$

= 30*20*10 – 20*20*10

= 6000 – 4000

det (X'X) = 2000

$$\text{Matrice des mineurs} = \begin{bmatrix} 20 & 0 & 20 & 0 & 20 & 20 \\ 0 & 10 & 0 & 10 & 0 & 0 \\ 20 & 0 & 30 & 0 & 30 & 20 \\ 0 & 10 & 0 & 10 & 0 & 0 \\ 20 & 0 & 30 & 0 & 30 & 20 \\ 20 & 0 & 20 & 0 & 20 & 20 \end{bmatrix}$$

$$\text{Matrice des mineurs} = \begin{bmatrix} 200 & 200 & 0 \\ 200 & 300 & 0 \\ 0 & 0 & 200 \end{bmatrix}$$

$$\text{Com}(X'X) = \begin{bmatrix} 200 & -200 & 0 \\ -200 & 300 & 0 \\ 0 & 0 & 200 \end{bmatrix} = 100 \begin{bmatrix} 2 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$(X'X)^{-1} = \frac{100}{2000} \begin{bmatrix} 2 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0.1 & -0.1 & 0 \\ -0.1 & 0.15 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}$$

$$\Rightarrow \hat{\beta} = \begin{bmatrix} -0.5 \\ 1.5 \\ 1 \end{bmatrix}$$

$$\Rightarrow \boxed{\begin{cases} \hat{\beta}_0 = -0.5 \\ \hat{\beta}_1 = 1.5 \\ \hat{\beta}_2 = 1 \end{cases}}$$

- Valeur de σ^2

$$\widehat{\sigma^2} = \frac{\text{SCR}}{\text{n}-\text{p}-1}$$

$$\text{SCR} = \sum_{i=1}^{n}(yi - \widehat{yi}\ )^2$$

$$= \sum_{i=1}^{n}(yi - \hat{\beta}0 - \hat{\beta}1\text{xi},1 - \hat{\beta}2\text{xi},2)^2$$

$$= \sum_{i=1}^{n} y^2 i + n\widehat{\beta}^{\,2}0 + \hat{\beta}^2 1 \sum_{i=1}^{n} \text{x}^2\text{i},1 + \hat{\beta}^2 2 \sum_{i=1}^{n} \text{x}^2\text{i},2 - 2\,\hat{\beta}^2 0 \sum_{i=1}^{n} \text{yi} - 2\hat{\beta}1 \sum_{i=1}^{n} x\text{i},1\text{yi} - 2\hat{\beta}2 \sum_{i=1}^{n} x\text{i},2\text{yi} + 2\hat{\beta}_0\,\hat{\beta}_1 \sum_{i=1}^{n} x\text{i},1 + 2\hat{\beta}_0\,\hat{\beta}_2 \sum_{i=1}^{n} x\text{i},2 + 2\hat{\beta}_1\hat{\beta}_2 \sum_{i=1}^{n} x\text{i},1\,xi,2$$

$$= 59.5 + 7.5 + 45 + 10 + 15 - 60 - 20 - 30 + 0 + 0$$

$$\boxed{\text{SCR} = 27}$$

$$\widehat{\sigma^2} = \frac{27}{27}$$

$$\boxed{\widehat{\sigma^2} = 1}$$

Question 3

- Calculons β_1 pour un intervalle de confiance à 95 %

$$\frac{\hat{\beta}1 - \beta1}{\widehat{\sigma 1}} = \frac{\hat{\beta}1 - \beta1}{\hat{\sigma}\sqrt{\sum_{i=1}^{n}(\text{X'X})-1(1,1)}} \sim \text{Tn}-\text{p-1} = \text{Tn}-2\text{-1} = \text{Tn}-3$$

$I(\beta 1) = [\hat{\beta}1 \pm t\,27(0.975)\hat{\sigma}\sqrt{(X'X)-1\ (1,1)}\,]$

$I(\beta 1) = [1.5 - 2.052 * \sqrt{0.15}\ ;\ 1.5 + 2.052 * \sqrt{0.15}\,]$

$\boxed{I(\beta 1) = [0.70\ ;\ 2.29\,]}$

- Tester $\beta 2 = 0.8$ à niveau 10%

$I(\beta 2) = [\hat{\beta}2 \pm t\,(n-p-1)(1-\boldsymbol{\alpha/2})\hat{\sigma}\sqrt{(X'X)-1\ (1,1)}\,]$

$I(\beta 2) = [\hat{\beta}2 \pm t\,27(0.95)\hat{\sigma}\sqrt{(X'X)-1\ (1,1)}\,]$

$I(\beta 2) = [1 - 1.703 * \sqrt{0.1}\ ;\ 1 + 1.703 * \sqrt{0.1}\,]$

$\boxed{I(\beta 2) = [0.46\ ;\ 1.54\,]}$

Donc on accepte au niveau 10% l'hypothèse selon laquelle $\beta 2 = 0.8$

Question 4

- Tester $\beta_0 + \beta_1 = 3$ vs $\beta_0 + \beta_1 \neq 3$ au risque 5%

$IC = [(\hat{\beta}0 + \hat{\beta}1) \pm t(n-p-1)\,(1-\boldsymbol{\alpha/2})\hat{\sigma}(\hat{\beta}0 + \hat{\beta}1)]$

Or Var(A + B) = Var(A) + Var(B) + 2cov(A,B)

⇨ $\widehat{\sigma^2}\ (\hat{\beta}0 + \hat{\beta}1) = \widehat{\sigma^2}\ (\hat{\beta}0) + \widehat{\sigma^2}\ (\hat{\beta}1) + 2\,\mathrm{cov}(\hat{\beta}0, \hat{\beta}1)$

$= \widehat{\sigma^2}\,[\,(X'X)^{-1}](1,1) - \widehat{\sigma^2}\,[\,(X'X)^{-1}](2,2) + 2\,\widehat{\sigma^2}\,[\,(X'X)^{-1}](1,2)$

$\hat{\sigma}\,(\hat{\beta}0 + \hat{\beta}1) =$
$\hat{\sigma}\sqrt{[(X'X)-1]\,(1,1) + [(X'X)-1]\,(2,2) + 2[(X'X)-1]\,(1,2)}$

$= \sqrt{0.1 + 0.15 + 2 * (-0.1)}$

$\boxed{\widehat{\sigma^2}\ (\hat{\beta}0 + \hat{\beta}1) = 0.224}$

$IC = [(-0.5 + 1.5) - t_{27}(0.975) * 0.224\ ;\ (-0.5 + 1.5) + t_{27}(0.975) * 0.224\,]$

IC = [0.540 ; 1.460]

Donc on rejette l'hypothèse selon laquelle $\hat{\beta}0 + \hat{\beta}1 = 3$ au risque $\alpha = 5\%$

Question 5

- Valeur de $\bar{y}$

$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \varepsilon_i$, $1 \leq i \leq n$

$\frac{1}{n}\sum yi = \beta_0 + \beta_1\frac{1}{n}\sum xi,1 + \beta_2\sum xi,2 + \bar{\varepsilon}$

$\bar{y} = \beta_0 + \beta_1\frac{1}{n}\sum xi,1 + \beta_2\sum xi,2$; $\bar{\varepsilon} = 0$

$= -0.5 + \frac{1.5}{30} * 20$

$\boxed{\bar{y} = 0.5}$

- Valeur R^2_a

$R^2_a = 1 - \frac{n-1}{n-p-1} * \frac{SCR}{SCT}$

$= 1 - (n-1) * \frac{\widehat{\sigma^2}}{SCT}$

Or $SCT = \sum_{i=1}^{n}(yi - \bar{yi})^2$

$= \sum yi^2 + n\bar{y} - 2\sum yi\,\bar{y}$

Or $y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \varepsilon_i$, $1 \leq i \leq n$

$\sum yi = n*\beta_0 + \sum \beta1\, xi,1 + \beta_2\sum xi,2 + \sum \varepsilon i$, $1 \leq i \leq n$

$= n*\beta_0 + \beta1\sum xi,1 + \beta_2\sum xi,2 + n*\bar{\varepsilon}$

Or $\sum xi,1 = 0$ et $\sum xi,2 = 0$

$\sum yi = 30 * (-0.5) + 1.5 * 20$

$\boxed{\sum yi = 15}$

$SCT = 59.5 + 30 * 0.5^2 - 2 * 0.5 * 15$

$\boxed{SCT = 52}$

$R^2_a = 1 - \frac{29}{52}$

$\boxed{R^2_a = 0.44}$

Question 6

- Intervalle de confiance à 95 % de y_{n+1} si $x_{n+1,1} = 3$ et $x_{n+1,2} = 0.5$

  Soit x'_{n+1} = [1, 3, 0.5 alors la valeur prédite de y_{n+1} est : $\widehat{y(n+1)} = x'_{n+1} * \widehat{\beta}$

  $\widehat{y(n+1)} = \begin{bmatrix} 1, 3, 0.5 \end{bmatrix} * \begin{bmatrix} -0.5 \\ 1.5 \\ 1 \end{bmatrix}$

  $\boxed{\widehat{y(n+1)} = 4.5}$

  $IC = [\widehat{y(n+1)} \pm t27 \ (0.975) \sqrt{1 + x'n(X'X) - 1 \ \ xn + 1}\ ]$

  Or : $1 + x'n(X'X) - 1 \ \ xn + 1 = 1 + \begin{bmatrix} 1, 3, 0.5 \end{bmatrix} * \begin{bmatrix} 0.1 & -0.1 & 0 \\ -0.1 & 0.15 & 0 \\ 0 & 0 & 0.1 \end{bmatrix} * \begin{bmatrix} 1 \\ 3 \\ 0.5 \end{bmatrix}$

  $= 1 + \begin{bmatrix} -0.2, 0.35, 0.05 \end{bmatrix} * \begin{bmatrix} 1 \\ 3 \\ 0.5 \end{bmatrix}$

  $\boxed{1 + x'n(X'X) - 1 \ \ xn + 1 = 1.875}$

  $IC = [4.5 - 2.052 * \sqrt{1.875} \ ; \ 4.5 + 2.052 * \sqrt{1.875}\ ]$

  $\boxed{IC = [1.69 ; 7.31]}$