Exercice 1

Almamy Youssouf LY

February 3, 2021

$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i$, $1 \leq$ i $\leq n$.

$Y = X\beta + \epsilon$

x_{ii} : *variablesexogénes*

ϵ_i: variable aléatoires indépendantes suivant une loi normale N(0, σ^2)

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

Y'Y = 59.5

Question 1

Valeur de n et $\bar{x}_{i2}$

X = $\begin{pmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{pmatrix}$

X' = $\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n1} \end{pmatrix}$

X'X = $\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n1} \end{pmatrix}$ $\begin{pmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{pmatrix}$

X'X = $\begin{pmatrix} n & \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i2} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i1}x_{i2} \\ \sum_{i=1}^n x_{i2} & \sum_{i=1}^n x_{i2}x_{i1} & \sum_{i=1}^n x_{i2}^2 \end{pmatrix}$ (1)

Par identification avec X'X = $\begin{pmatrix} 30 & 20 & 0 \\ 20 & 20 & 0 \\ 0 & 0 & 10 \end{pmatrix}$ (2)

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

$\bar{x}_{i2} = \frac{1}{n} \cdot \sum_{i=1}^n x_{i2} \ = \frac{1}{30} \cdot \ 0$

Donc $\boxed{\bar{x} = 0}$ (*)

Coefficient de corrélation des variables x$_{i1} et x_{i2}$

$$\begin{aligned} r &= \frac{\sum_{i=1}^n (x_{i1} - \bar{x}_{i1})(x_{i2} - \bar{x}_{i2})}{\sqrt{\sum_{i=1}^n (x_{i1} - \bar{x}_{i1})^2} \sqrt{\sum_{i=1}^n (x_{i2} - \bar{x}_{i2})^2}} \\ &= \frac{\sum_{i=1}^n x_{i2}(x_{i1} - \bar{x}_{i1}) - \sum_{i=1}^n \bar{x}_{i2}(x_{i1} - \bar{x}_{i1})}{\sqrt{\sum_{i=1}^n (x_{i1} - \bar{x}_{i1})^2} \sqrt{\sum_{i=1}^n x_{i2}^2 + n\bar{x}_{i2}^2 - 2\sum_{i=1}^n x_{i2}\bar{x}_{i2}}} \\ &= \frac{\sum_{i=1}^n x_{i2}x_{i1} - \sum_{i=1}^n x_{i2}\bar{x}_{i1} - \sum_{i=1}^n \bar{x}_{i2}x_{i1} + \sum_{i=1}^n \bar{x}_{i2}\bar{x}_{i1}}{\sqrt{\sum_{i=1}^n (x_{i1} - \bar{x}_{i1})^2} \sqrt{\sum_{i=1}^n x_{i2}^2}} \\ &= \frac{\sum_{i=1}^n x_{i2}x_{i1}}{\sqrt{\sum_{i=1}^n (x_{i1} - \bar{x}_{i1})^2} \sqrt{\sum_{i=1}^n x_{i2}^2}} \end{aligned}$$

Or d'aprés les matrices (1) et (2) $\sum_{i=1}^n x_{i2}x_{i1} = 0$

Donc $\boxed{\text{r = 0}}$

Question 2

Valeur de $\widehat{\beta}_0$; $\widehat{\beta}_1$ *et* $\widehat{\beta}_2$

$$det(X'X) = \begin{vmatrix} 30 & 20 & 0 \\ 20 & 20 & 0 \\ 0 & 0 & 10 \end{vmatrix} \begin{vmatrix} 30 & 20 \\ 20 & 20 \\ 0 & 0 \end{vmatrix} = 30 \times 20 \times 10 - 20 \times 20 \times 10 = 2000$$

$$matrice\ des\ mineurs = \begin{vmatrix} 20 & 0 & 20 & 0 & 20 & 20 \\ 0 & 10 & 0 & 10 & 0 & 0 \\ 20 & 0 & 30 & 0 & 30 & 20 \\ 0 & 10 & 0 & 10 & 0 & 0 \\ 20 & 0 & 30 & 0 & 30 & 20 \\ 20 & 0 & 20 & 0 & 20 & 20 \end{vmatrix} = \begin{vmatrix} 200 & 200 & 0 \\ 200 & 300 & 0 \\ 0 & 0 & 200 \end{vmatrix}$$

$$com(X'X) = \begin{vmatrix} 200 & -200 & 0 \\ -200 & 300 & 0 \\ 0 & 0 & 200 \end{vmatrix} = 100 \begin{vmatrix} 2 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 2 \end{vmatrix}$$

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

$(\text{X'X})^{-1} = \begin{vmatrix} 0.1 & -0.1 & 0 \\ -0.1 & 0.15 & 0 \\ 0 & 0 & 0.1 \end{vmatrix}$

$$\widehat{\beta} = \begin{pmatrix} 0.1 & -0.1 & 0 \\ -0.1 & 0.15 & 0 \\ 0 & 0 & 0.1 \end{pmatrix} \begin{pmatrix} 15 \\ 20 \\ 10 \end{pmatrix}$$

Donc

$$\boxed{\widehat{\beta} = \begin{pmatrix} -0.5 \\ 1.5 \\ 1 \end{pmatrix}}$$

Donc $\widehat{\beta}_0 = -0.5$; $\widehat{\beta}_1 = 1.5$ et $\widehat{\beta}_2 =$

Valeurs de la variance

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

$$Or\ n-p-1 = 30-2-1 = 27$$

$$Et\ SCR = \sum_{i=1}^{n}(y_i - \widehat{y_i})^2 = \sum_{i=1}^{n}(y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_{i1} - \widehat{\beta}_2 x_{i2})^2$$

$= \sum_{i=1}^n y_i^2 + n\widehat{\beta}_0^{\ 2} + \widehat{\beta}_1^{\ 2}\sum_{i=1}^n x_{i1}^2 + \widehat{\beta}_2^{\ 2}\sum_{i=1}^n x_{i2}^2 - 2\widehat{\beta}_0\sum_{i=1}^n y_i - 2\widehat{\beta}_1\sum_{i=1}^n x_{i1}y_i - 2\widehat{\beta}_2\sum_{i=1}^n x_{i2}y_i$

$+ 2\widehat{\beta}_0\widehat{\beta}_1\sum_{i=1}^n x_{i1} + 2\widehat{\beta}_0\widehat{\beta}_2\sum_{i=1}^n x_{i2} + 2\widehat{\beta}_1\widehat{\beta}_2\sum_{i=1}^n x_{i1}x_{i2}$

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

$= 27$

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

Question 3

Intervalle de confiance pour β_1 à 95%

$\frac{\widehat{\beta_1 - \beta_1}}{\widehat{\sigma_1}} = \frac{\widehat{\beta_1 - \beta_1}}{\widehat{\sigma}\sqrt{(XX)^{-1}_{11}}} \sim t_{n-p-1}$

$\mathrm{I}(\beta_1) = [\widehat{\beta}_1 - t_{n-p-1}(1-\frac{\alpha}{2})\widehat{\sigma}\sqrt{(XX)^{-1}_{11}}, \widehat{\beta}_1 + t_{n-p-1}(1-\frac{\alpha}{2})\widehat{\sigma}\sqrt{(XX)^{-1}_{11}}]$
$I(\beta_1) = [1.5 - t_{27}(0.975)\sqrt{0.15}; 1.5 + t_{27}(0.975)\sqrt{0.15}]$
$I(\beta_1) = [1.5 - 2.052\sqrt{0.15}; 1.5 + 2.052\sqrt{0.15}]$
$Donc$ $\boxed{\mathrm{I}(\beta_1) = [0.70; 2.29]}$

Tester $\beta_2 = 0.8$ *à niveau* 10%
$I(\beta_2) = [\widehat{\beta}_2 - t_{n-p-1}(1-\frac{\alpha}{2})\widehat{\sigma}\sqrt{(XX)_{11}^{-1}}, \widehat{\beta}_2 + t_{n-p-1}(1-\frac{\alpha}{2})\widehat{\sigma}\sqrt{(XX)_{11}^{-1}}]$
$I(\beta_2) = [1 - t_{27}(0.95)\sqrt{0.1}; 1 + t_{27}(0.95)\sqrt{0.1}]$
$I(\beta_2) = [1 - 1.703\sqrt{0.1}; 1 + 1.703\sqrt{0.1}]$
$D'où \boxed{\mathrm{I}(\beta_2) = [0.46; 1.54]}$

Donc on accepte au niveau 10% l'hypothése selon laquelle $\beta_2 = 0.8$

Question 4

Tester $\beta_0 + \beta_1$ vs $\beta_0 + \beta_1$ *au risque* 5%

$\mathrm{I}(\beta_0 + \beta_1) = [(\widehat{\beta}_0 + \widehat{\beta}_1) - t_{n-p-1}(1-\frac{\alpha}{2})\widehat{\sigma}(\widehat{\beta}_0 + \widehat{\beta}_1), (\widehat{\beta}_0 + \widehat{\beta}_1) + t_{n-p-1}(1-\frac{\alpha}{2})\widehat{\sigma}(\widehat{\beta}_0 + \widehat{\beta_1})]$

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

D'où $\widehat{\sigma}^2_{(\widehat{\beta_0}+\widehat{\beta_1})} = \widehat{\sigma}^2(\widehat{\beta}_0) + \widehat{\sigma}^2(\widehat{\beta}_1) + 2cov(\widehat{\beta}_0, \widehat{\beta}_1)$

$\widehat{\sigma}^2_{(\widehat{\beta_0}+\widehat{\beta_1})} = \widehat{\sigma}^2[(X'X)^{-1}]_{11} - \widehat{\sigma}^2[(X'X)^{-1}]_{22} + 2\widehat{\sigma}^2[(X'X)^{-1}]_{12}$

$\widehat{\sigma}_{(\widehat{\beta_0}+\widehat{\beta_1})} = \widehat{\sigma}\sqrt{[(X'X)^{-1}]_{11} - [(X'X)^{-1}]_{22} + 2[(X'X)^{-1}]_{12}}$

$\widehat{\sigma}_{(\widehat{\beta_0}+\widehat{\beta_1})} = \sqrt{0.1 + 0.15 + 2 \times (-0.1)}$

Donc $\widehat{\sigma}_{(\widehat{\beta_0}+\widehat{\beta_1})} = 0.224$

D'où: $\mathrm{I}(\beta_0 + \beta_1) = [(-0.5 + 1.5) - t_{27}(0.975)0.224; (-0.5 + 1.5) + t_27(0.975)0.224]$

$\mathrm{I}(\beta_0 + \beta_1) = [0.540; 1.460]$

Donc on rejette l'hypothése selon laquelle $\widehat{\beta}_0 + \widehat{\beta}_1 = 3$ *au risque* $\alpha = 5\%$

Question 5

Valeur de $\bar{y}$

$\mathrm{y}_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i$

$\frac{1}{n}\sum_{i=1}^{n} y_i \;=\; \beta_0 + \beta_1 \frac{1}{n}\sum_{i=1}^{n} x_{i1} + \beta_2 \frac{1}{n}\sum_{i=1}^{n} x_{i2} + \bar{\epsilon}$

Or par définition ϵ *est de moyenne nulle*

D'où $\;=\; \beta_0 + \beta_1 \frac{1}{n}\sum_{i=1}^{n} x_{i1} + \beta_2 \frac{1}{n}\sum_{i=1}^{n} x_{i2}$

$\bar{y} \;=\;$ -0.5 + $\frac{1.5}{30}$ $\times$20

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

Valeur de R_a^2

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

$\mathrm{R}_a^2 = 1 - (n-1) \times \frac{\widehat{\sigma}^2}{SCT}$

Or $\mathrm{y}_i = \beta_0 + \beta_1 x_{i1} + \beta_1 x_{i2} + \epsilon_i, 1 \leq i \leq n$

$\sum_{i=1}^{n} y_i = n\beta_0 + \sum_{i=1}^{n} \beta_1 x_{i1} + \sum_{i=1}^{n} \beta_2 x_{i2} + \sum_{i=1}^{n} \epsilon_i$

$\sum_{i=1}^{n} y_i = n\beta_0 + \beta_1 \sum_{i=1}^{n} x_{i1} + \beta_2 \sum_{i=1}^{n} x_{i2} + n\bar{\epsilon}$

Or $\sum_{i=1}^{n} x_{i1} = 0$ *et* $\sum_{i=1}^{n} x_{i2} = 0$

D'où $\sum_{i=1}^{n} y_i = 30 \times (-0.5) + 1.5 \times 20$

$\boxed{\sum_{i=1}^{n} y_i = 15}$

SCT = $59.5 + 30 \times 0.5^2 - 2 \times 0.5 \times 15$

$\boxed{\text{SCT} = 52}$

Par suite: $R_a^2 = 1 - \frac{29}{52}$

Donc $\boxed{R_a^2 = 0.44}$

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 $x'_{n+1} = [1, 3, 0.5]$, *la valeur prédite pour* y_{n+1} *est* : $\widehat{y}_{n+1} = 4.5$

IC $= [\widehat{y}_{n+1} + t_{27}(0.975)\sqrt{1 + x'_{n+1}(X'X)^{-1}x_{n+1}}, \widehat{y}_{n+1} - t_{27}(0.975)\sqrt{1 + x'_{n+1}(X'X)^{-1}x_{n+1}}]$

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

$1 + x'_{n+1}(X'X) - 1x_{n+1} = 1 + (-0.2, 0.35, 0.05) \begin{pmatrix} 1 \\ 3 \\ 0.5 \end{pmatrix}$

$1 + x'_{n+1}(X'X) - 1x_{n+1} = 1.875$

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

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