UMBB/Sciences/Maths/Proba-Stat

Filière: Master MSS semestre3.

Module: **Econométrie.**

Année universitaire 2021-2022

Mars 2022

.

Corrigé EPREUVE de Rattrapage

Exercice 1(8 points)

Pour le Modèle linéaire simple.

$y_i = a_0 + a_1 x_i + \epsilon_i$

$$\begin{bmatrix} i & X = x_i & Y = y_i \\ 1 & -1 & 4 \\ 2 & 0 & 1 \\ 3 & 1 & -1 \\ 4 & 2 & -2 \\ 5 & 3 & -7 \end{bmatrix}$$

On vous donne:

$X = x_i$	$x_i - \bar{X}$	$Y - \bar{Y}$	$Y = y_i$	$(x_i - \bar{X})(y_i - \bar{Y})$	$(x_i - \bar{X})^2$	$(y_i - \bar{Y})^2$
-1	-2	5	4	-10	4	25
0	-1	2	1	-2	1	4
1	0	0	-1	0	0	0
2	1	-1	-2	-1	1	1
3	2	-6	-7	-12	4	36
$\sum = 5$	_	_	-5	$\mathbf{-25}$	**10**	**66**

$\sum_i (x_i - \bar{X})^2 = 10, \quad \sum_i (y_i - \bar{Y})^2 = 66, \quad \sum_i (x_i - \bar{X})(y_i - \bar{Y}) = -25$

(i) Calculer $\hat{a}_1$estimateur de a_1 méthode MCO.

$\hat{a}_1 = \dfrac{cov(X,Y)}{var(X)} = \dfrac{\sum_i (x_i - \bar{X})(y_i - \bar{Y})}{\sum_i (x_i - \bar{X})^2} = \dfrac{-25}{10} = -2.5$

(ii) Evaluer $\rho(X,Y)$ interpréter

$\rho(X,Y) = \dfrac{cov(X,Y)}{\sqrt{var(X)var(Y)}} =$

$= \dfrac{\sum_i (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_i (x_i - \bar{X})^2 \sum_i (y_i - \bar{Y})^2}} = \dfrac{-25}{\sqrt{10 \times 66}} = -0.973\,12$

Interprétation: Trés forte corrélation linéaire négative entre x_i et y_i

(iii) Calculer $\rho(X,Y)\dfrac{\sigma_Y}{\sigma_x}$ Que retrouve t'on? Le résultat est il prévisible? Démonter le résultat en générale.Déduire que $\hat{a}_1$ et$\rho(X,Y)$ ont le meme signe et que la nullité de l'un entraine la nullité de l'autre.

Solution:

$\rho(X,Y)\dfrac{\sigma_Y}{\sigma_x} = \dfrac{-25}{\sqrt{10 \times 66}} \dfrac{\sqrt{66}}{\sqrt{10}} = -2.\,5$

On retrouve $\hat{a}_1$.

le résultat est prévisible.

Justification et démonstyration en générale:

$\hat{a}_1 = \dfrac{\sum_i (x_i - \bar{X})(y_i - \bar{Y})}{\sum_i (x_i - \bar{X})^2} = \dfrac{\sum_i (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_i (x_i - \bar{X})^2}\sqrt{\sum_i (x_i - \bar{X})^2}} \dfrac{\sqrt{\sum_i (y_i - \bar{Y})^2}}{\sqrt{\sum_i (y_i - \bar{Y})^2}}$(on multiplie et on divise par la meme quantité)

$$= \dfrac{\sum_i (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_i (x_i - \bar{X})^2 \sum_i (y_i - \bar{Y})^2}} \dfrac{\sqrt{\sum_i (y_i - \bar{Y})^2}}{\sqrt{\sum_i (x_i - \bar{X})^2}} = \rho(X,Y)\dfrac{\sigma_Y}{\sigma_x}$$

$\hat{a}_1$ et $\rho(X,Y)$ ont le meme signe car $\dfrac{\sigma_Y}{\sigma_x}$ est >0

et $\hat{a}_1 = 0 \Leftrightarrow \rho(X,Y) = 0$ car $\dfrac{\sigma_Y}{\sigma_x} \neq 0$

(iv) calculer $\hat{a}_0$ et Déduire $\hat{A} = \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \end{bmatrix}$

$\hat{a}_0 = \bar{Y} - \hat{a}_1 \bar{X}$

$= -1 + 2.5 = 1.5$

$\hat{A} = \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \end{bmatrix} = \begin{bmatrix} 1.5 \\ -2.5 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 3 \\ -5 \end{bmatrix}$

(v) Ecrire $\hat{Y}$ sous la forme $\hat{Y} = X\hat{A}$ Calculer $\hat{Y} - \bar{Y}$ et Déduire SCResiduelle.

$$\hat{Y} = X\hat{A} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 3 \\ -5 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 8 \\ 3 \\ -2 \\ -7 \\ -12 \end{bmatrix} = \begin{bmatrix} 4 \\ \frac{3}{2} \\ -1 \\ -\frac{7}{2} \\ -6 \end{bmatrix}$$

$$(Y-\hat{Y}) = \begin{bmatrix} 4 \\ 1 \\ -1 \\ -2 \\ -7 \end{bmatrix} - \begin{bmatrix} 4 \\ \frac{3}{2} \\ -1 \\ -\frac{7}{2} \\ -6 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{2} \\ 0 \\ \frac{3}{2} \\ -1 \end{bmatrix}$$

$$SCResi = (Y-\hat{Y})^T\ (Y-\hat{Y}) = \begin{bmatrix} 0 \\ -\frac{1}{2} \\ 0 \\ \frac{3}{2} \\ -1 \end{bmatrix}^T \begin{bmatrix} 0 \\ -\frac{1}{2} \\ 0 \\ \frac{3}{2} \\ -1 \end{bmatrix} = \frac{7}{2}$$

(vi)Calculer le coefficient de Détermination R^2 de deux manières. Interpéter. .

$$R^2 = \frac{SCExpl}{SCT} = \frac{62.5}{66} = 0.946\,97$$
$$R^2 = 1 - \frac{SC\,\mathrm{Re}\,s}{SCT} = 1 - \frac{3.5}{66} = 0.946\,97$$

Interprétation

R^2 est une mesure de la qualité de prédiction de la régression (Dans le MLS il est égale à $\rho^2(X,Y)$)

R^2 est aussi le taux de variation des y expliqués par la ou les variables explicatives(variables indépendantes)

Ainsi dans ce cas:

94,70 est le pourcentage de la variation totales des y expliqués par le modèle de regression, le reste 5,30 est du aux erreurs.

conclusion: Le modèle est trés bon

$$\hat{Y}-\bar{Y}\times ones(5,1)=\begin{bmatrix}4\\ \frac{3}{2}\\ -1\\ -\frac{7}{2}\\ -6\end{bmatrix}-\begin{bmatrix}-1\\ -1\\ -1\\ -1\\ -1\end{bmatrix}=\begin{bmatrix}5\\ \frac{5}{2}\\ 0\\ -\frac{5}{2}\\ -5\end{bmatrix}=\begin{bmatrix}5\\ 2.5\\ 0\\ -2.5\\ -5\end{bmatrix}$$

$$SCExpl=\begin{bmatrix}5\\ 2.5\\ 0\\ -2.5\\ -5\end{bmatrix}^T\begin{bmatrix}5\\ 2.5\\ 0\\ -2.5\\ -5\end{bmatrix}=62.5$$

Calculer $\hat{Y}-\bar{Y}\times$ones$(5,1)$ et déduire SCExpli où le vecteur ones$(5,1)=\begin{bmatrix}1\\ 1\\ 1\\ 1\\ 1\end{bmatrix}$

(vii) comparer $*=\dfrac{\sum_i(\hat{y}_i-\bar{Y})(y_i-\bar{Y})}{\sqrt{\sum(\hat{y}_i-\bar{Y})^2}\sqrt{\sum_i(y_i-\bar{Y})^2}}$ avec $\rho(X,Y)$ déduire une relation avec R^2. Comment appelle t'on *?

$*=\dfrac{\sum_i(\hat{y}_i-\bar{Y})(y_i-\bar{Y})}{\sqrt{\sum(\hat{y}_i-\bar{Y})^2}\sqrt{\sum_i(y_i-\bar{Y})^2}}=\dfrac{62.5}{\sqrt{66}\sqrt{62.5}}=0.9731$

Comparaison: on a

$*=\dfrac{\sum_i(\hat{y}_i-\bar{Y})(y_i-\bar{Y})}{\sqrt{\sum(\hat{y}_i-\bar{Y})^2}\sqrt{\sum_i(y_i-\bar{Y})^2}}=\rho(X,Y)$

Relation avec R^2

$R^2=*^2$

car On a: $R^2=\rho^2(X,Y)=\rho^2_{(Y,\hat{Y})}$

$*=\dfrac{\sum_i(\hat{y}_i-\bar{Y})(y_i-\bar{Y})}{\sqrt{\sum(\hat{y}_i-\bar{Y})^2}\sqrt{\sum_i(y_i-\bar{Y})^2}}$ est appelé corrélation entre Y et $\hat{Y}$ et noté $\rho^2_{(Y,\hat{Y})}$

pour la démonstration de l'égalité $\rho(X,Y)=\rho_{(Y,\hat{Y})}$

voir $\hat{y}_i-\bar{Y}=\hat{a}_1(x_i-\bar{X})$

Exercice 2(12 points)

Soit le modèle de régression linéaire multiple

$Y_i = a_0 + a_1 x_{i1} + a_2 x_{i2} + \epsilon_i$

$$\begin{bmatrix} i & x_{i1} & x_{i2} & y_i \\ 1 & -1 & 0 & 3 \\ 2 & 0 & 2 & 6 \\ 3 & 1 & 1 & 2 \\ 4 & 2 & -1 & 0 \\ 5 & -2 & -2 & -1 \\ 6 & 3 & -3 & -15 \\ 7 & -3 & 0 & 5 \\ 8 & 0 & 3 & -8 \end{bmatrix}$$

(i) Donner l'écriture Matricielle du modèle(Utiliser les notations faites en cours), Ecrire $'XX$ en fonction de n, $\sum_{i=1}^{n} x_{i1}$, $\sum_{i=1}^{n} x_{i2}$, ...etc.

$Y = XA + \epsilon$ où

$$Y = \begin{bmatrix} 3 \\ 6 \\ 2 \\ 0 \\ -1 \\ -15 \\ 5 \\ -8 \end{bmatrix} \qquad \frac{1}{8}\sum_i y_i = -1$$

$$X = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -2 & -2 \\ 1 & 3 & -3 \\ 1 & -3 & 0 \\ 1 & 0 & 3 \end{bmatrix} \quad A = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \quad \text{et} \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \\ \\ \\ \epsilon_8 \end{bmatrix}$$

$$X'X = \begin{bmatrix} n & \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i2} \\ \sum_{i=1}^n x_{i1}0 & \sum_{i=1}^n x_{i1}^2 & \sum_i x_{i1}x_{i2} \\ \sum_{i=1}^n x_{i2} & \sum_i x_{i1}x_{i2} & \sum_{i=1}^n x_{i2}^2 \end{bmatrix}$$

Indication: On vous donne$(X\prime X)^{-1}$= $\begin{bmatrix} \frac{1}{8} & 0 & 0 \\ 0 & \frac{7}{187} & \frac{3}{374} \\ 0 & \frac{3}{374} & \frac{7}{187} \end{bmatrix}$

(ii)Calculer $\hat{A}$=$(\hat{a}_0, \hat{a}_1, \hat{a}_2)'$ On commence par présenter la formule sans démonstration

$\hat{A} = (\hat{a}_0, \hat{a}_1, \hat{a}_2)' = (X'X)^{-1}X'Y$

on a X'Y= $\begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -2 & -2 \\ 1 & 3 & -3 \\ 1 & -3 & 0 \\ 1 & 0 & 3 \end{bmatrix}^T \begin{bmatrix} 3 \\ 6 \\ 2 \\ 0 \\ -1 \\ -15 \\ 5 \\ -8 \end{bmatrix} = \begin{bmatrix} -8 \\ -59 \\ 37 \end{bmatrix}$

$$\hat{A} = (\hat{a}_0, \hat{a}_1, \hat{a}_2)' = (X'X)^{-1}X'Y = \begin{bmatrix} \frac{1}{8} & 0 & 0 \\ 0 & \frac{7}{187} & \frac{3}{374} \\ 0 & \frac{3}{374} & \frac{7}{187} \end{bmatrix} \begin{bmatrix} -8 \\ -59 \\ 37 \end{bmatrix} = \begin{bmatrix} -1 \\ -\frac{65}{34} \\ \frac{31}{34} \end{bmatrix} = \begin{bmatrix} -1.0 \\ -1.911\,8 \\ 0.911\,76 \end{bmatrix}$$

(iii) Compléter le tableau d'ANOVA

$$\hat{Y} = X\hat{A} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -2 & -2 \\ 1 & 3 & -3 \\ 1 & -3 & 0 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ -\frac{65}{34} \\ \frac{31}{34} \end{bmatrix} = \begin{bmatrix} \frac{31}{34} \\ \frac{14}{17} \\ -2 \\ -\frac{195}{34} \\ 1 \\ -\frac{161}{17} \\ \frac{161}{34} \\ \frac{59}{34} \end{bmatrix} = \begin{bmatrix} 0.911\,76 \\ 0.823\,53 \\ -2.0 \\ -5.735\,3 \\ 1.0 \\ -9.470\,6 \\ 4.735\,3 \\ 1.735\,3 \end{bmatrix}$$

$$\epsilon = Y - \hat{Y} = \begin{bmatrix} 3 \\ 6 \\ 2 \\ 0 \\ -1 \\ -15 \\ 5 \\ -8 \end{bmatrix} - \begin{bmatrix} \frac{31}{34} \\ \frac{14}{17} \\ -2 \\ -\frac{195}{34} \\ 1 \\ -\frac{161}{17} \\ \frac{161}{34} \\ \frac{59}{34} \end{bmatrix} = \begin{bmatrix} \frac{71}{34} \\ \frac{88}{17} \\ 4 \\ \frac{195}{34} \\ -2 \\ -\frac{94}{17} \\ \frac{9}{34} \\ -\frac{331}{34} \end{bmatrix} = \begin{bmatrix} 2.088\,2 \\ 5.176\,5 \\ 4.0 \\ 5.735\,3 \\ -2.0 \\ -5.529\,4 \\ 0.264\,71 \\ -9.735\,3 \end{bmatrix}$$

SCResi=$\epsilon' \, \epsilon$

$$\epsilon' \, \epsilon = \begin{bmatrix} \frac{71}{34} \\ \frac{88}{17} \\ 4 \\ \frac{195}{34} \\ -2 \\ -\frac{94}{17} \\ \frac{9}{34} \\ -\frac{331}{34} \end{bmatrix}^T \begin{bmatrix} \frac{71}{34} \\ \frac{88}{17} \\ 4 \\ \frac{195}{34} \\ -2 \\ -\frac{94}{17} \\ \frac{9}{34} \\ -\frac{331}{34} \end{bmatrix} = \frac{3561}{17} = 209.47$$

$$\text{ScRes}=\epsilon'\,\epsilon=\begin{bmatrix}\frac{71}{34}\\ \frac{88}{17}\\ 4\\ \frac{195}{34}\\ -2\\ -\frac{94}{17}\\ \frac{9}{34}\\ -\frac{331}{34}\end{bmatrix}^T\begin{bmatrix}\frac{71}{34}\\ \frac{88}{17}\\ 4\\ \frac{195}{34}\\ -2\\ -\frac{94}{17}\\ \frac{9}{34}\\ -\frac{331}{34}\end{bmatrix}=\frac{3561}{17}=209.47$$

$$\hat{Y}-\bar{Y}=\begin{bmatrix}\frac{31}{34}\\ \frac{14}{17}\\ -2\\ -\frac{195}{34}\\ 1\\ -\frac{161}{17}\\ \frac{161}{34}\\ \frac{59}{34}\end{bmatrix}-\begin{bmatrix}-1\\ -1\\ -1\\ -1\\ -1\\ -1\\ -1\\ -1\end{bmatrix}=\begin{bmatrix}\frac{65}{34}\\ \frac{31}{17}\\ -1\\ -\frac{161}{34}\\ 2\\ -\frac{144}{17}\\ \frac{195}{34}\\ \frac{93}{34}\end{bmatrix}=\begin{bmatrix}1.9118\\ 1.8235\\ -1\\ -4.7353\\ 2\\ -8.4706\\ 5.7353\\ 2.7353\end{bmatrix}$$

$$\text{SCExp}=\ (\hat{Y}-\bar{Y})'(\hat{Y}-\bar{Y})=\begin{bmatrix}\frac{65}{34}\\ \frac{31}{17}\\ -1\\ -\frac{161}{34}\\ 2\\ -\frac{144}{17}\\ \frac{195}{34}\\ \frac{93}{34}\end{bmatrix}^T\begin{bmatrix}\frac{65}{34}\\ \frac{31}{17}\\ -1\\ -\frac{161}{34}\\ 2\\ -\frac{144}{17}\\ \frac{195}{34}\\ \frac{93}{34}\end{bmatrix}\ \frac{2491}{17}=146.53$$

SCTot=(Y-$\bar{\text{Y}}$)'(Y-$\bar{\text{Y}}$)

$$\text{Y-}\bar{\text{Y}}=\begin{bmatrix}3\\6\\2\\0\\-1\\-15\\5\\-8\end{bmatrix}-\begin{bmatrix}-1\\-1\\-1\\-1\\-1\\-1\\-1\\-1\end{bmatrix}=\begin{bmatrix}4\\7\\3\\1\\0\\-14\\6\\-7\end{bmatrix}$$

$$(\text{Y-}\bar{\text{Y}})'(\text{Y-}\bar{\text{Y}})=\begin{bmatrix}4\\7\\3\\1\\0\\-14\\6\\-7\end{bmatrix}^T\begin{bmatrix}4\\7\\3\\1\\0\\-14\\6\\-7\end{bmatrix}=356$$

Source de variation	Somme des carrés	Degrés de liberté	Carrée Moye
Modèle (Expliquée par la Régression)	SCExp $=\frac{2491}{17}=146.53$	p=2	SCEx/p=$\frac{146}{}$
Résiduel(Non Expliquée par la régression)	SCResi=$\frac{3561}{17}=209.47$	n-p-1=5	SCRes/n-p-1
Totale	SCTot=356 $=\frac{2491}{17}+\frac{3561}{17}$	n-1=7	-

(iv) Calculer R^2 de deux façon différentes.Interpréter.

$R^2=\frac{SCExp}{SCTot}=\frac{2491}{17\times356}=\frac{2491}{6052}=0.411\,60$

$R^2=1-\frac{SCResi}{SCTot}=1-\frac{3561}{17\times356}=0.411\,60$

$\hat{V}(\hat{A})$=SCRes/n-p-1$\times$(X'X)$^{-1}$

$$\hat{V}(\hat{A})=\frac{3561}{85}\begin{bmatrix}\frac{1}{8}&0&0\\0&\frac{7}{187}&\frac{3}{374}\\0&\frac{3}{374}&\frac{7}{187}\end{bmatrix}=\begin{bmatrix}\frac{3561}{680}&0&0\\0&\frac{24\,927}{15\,895}&\frac{10\,683}{31\,790}\\0&\frac{10\,683}{31\,790}&\frac{24\,927}{15\,895}\end{bmatrix}=\begin{bmatrix}5.236\,8&0&0\\0&1.568\,2&0.336\,05\\0&0.336\,05&1.568\,2\end{bmatrix}$$

Ainsi $\hat{\sigma}_1^2=1.5682$

$\hat{\sigma}_2^2=1.5682$

Indication: pour tous les tests Prendre $\alpha = 5\%$. Effectuez des test unilatérales

(v) Effectuer le test de Student ($H_0 : a_2 = 0$) contre ($H_1 : a_2 \neq 0$).

la stat Tc=$\dfrac{|\hat{a}_2|}{\hat{\sigma}_2^2}$
on a $\hat{a}_2 = 0.911\,76$
Tc=$\dfrac{0.91176}{\sqrt{1.5682}} = 0.728\,08$
$T_\alpha = q_{0.95}(T(\nu = n - p - 1 = 5) = 2.015$

comme $T_c < T_\alpha$ on accepte H_0

La variable X_2 est mauvaise pour expliquer Y

(vi) Effectuer le test de Fischer et Donner l'interprétation

$F_c = \dfrac{CME}{CMR} = \dfrac{SCExp/p}{Sc\operatorname{Re}s/n-p-1} = \dfrac{73.\,265}{41.\,894} = 1.\,748\,8$
$F_{th} = Fisher(p, n-p-1) = F(2,5) = 5.79$

Conclusion: Comme $F_c < F_{th}$

$1.\,748\,8 < 5.79$

On accepte H_0 au niveau de signification95%

Le modèle est mauvais aucune variables exogènes n'est pertinentes pour expliquer Y

Indication:

On vous donne $d_{Low} = 0.56$ et $d_{upper} = 1.78$

(vii) Effectuer le test de Durbin Watson(D.W) et présenter la conclusion