Exercice 2.

1) Test de Klein :

- coefficient de corélation entre x_1 et x_2 : $\rho_{x_1 x_2}$

$$\rho_{x_1 x_2} = \frac{cov(x_1, x_2)}{\sqrt{V(x_1)\,V(x_2)}}$$

- $$cov(x_1, x_2) = \frac{1}{T}\sum_{t=1}^{T}(x_{1t} - \bar{x}_1)(x_{2t} - \bar{x}_2)$$

$$= \frac{1}{T}\sum_{t=1}^{T} x_{1t} \cdot x_{2t} - \bar{x}_1 \bar{x}_2$$

avec :

$$\begin{cases} \bar{x}_1 = \frac{1}{T}\sum_{t=1}^{T} x_{1t} = \frac{2}{10} = 0,2 \\ \bar{x}_2 = \frac{1}{T}\sum_{t=1}^{T} x_{2t} = \frac{2}{10} = 0,2 \end{cases}$$

$$\Rightarrow cov(x_1, x_2) = \left[\frac{1}{10} \times (-2)\right] - (0,2 \times 0,2) = -0,24$$

$$V(x_1) = \frac{1}{T}\sum (x_{1t} - \bar{x}_1)^2$$

$$= \frac{1}{T}\sum x_{1t}^2 - \bar{x}_1^2$$

$$= \frac{1}{10} \times 6 - (0,2)^2$$

$$= 0,56.$$

$$V(x_2) = \frac{1}{T}\sum (x_{2t} - \bar{x}_2)^2$$

$$= \frac{1}{T}\sum x_{2t}^2 - \bar{x}_2^2$$

$$= \frac{1}{10} \times 6 - (0,2)^2 = 0,56.$$

$$\Rightarrow \rho_{x_1, x_2} = \frac{-0,24}{0,56} = -0,428$$

$$\Rightarrow \rho^2_{x_1 x_2} = 0,183 < R^2 = 0,93$$

$$\frac{\ldots}{17 - 3 \times (2,334)^2} = -2,548$$

→ il y a absence de multicolinéarité entre x_1 et x_2.

2) Test de Farrar - Glauber :

1ère étape :

Calcul du déterminant de la matrice de corrélation :

$$D = \begin{vmatrix} 1 & \rho_{x_1x_2} \\ \rho_{x_2x_1} & 1 \end{vmatrix} = 1 - \rho^2_{x_1x_2} = 1 - 0,183$$

$\Rightarrow D = 0,817$.

2ème étape :

$$\begin{cases} H_0 : \text{les séries sont orthogonales } (D = 1) \\ H_1 : \text{les séries sont dépendantes } (D < 1) \end{cases}$$

La statistique utilisée et loi de probabilité :

$$\chi^2_c = -\left[T - 1 - \frac{1}{6}(2K+5)\right] \cdot \log D \sim \chi^2_\alpha\left(\frac{1}{2}K(K-1)\right) = \chi^2_{5\%}(3)$$

Règle de décision :

- Si $\chi^2_c \leq \chi^2_\alpha$ alors H_0 est vraie.
- Si $\chi^2_c > \chi^2_\alpha$ alors H_1 est vraie.

- conclusion :

$$\chi^2_c = -\left[10 - 1 - \frac{1}{6}(6+5)\right] \log(0,817)$$

$\Rightarrow \chi^2_c = 1,448$.

- $\chi^2_{5\%}(3) = 7,815$ $\Big\}$ $\Rightarrow \chi^2_c < \chi^2_\alpha$

→ H_0 est vraie.

→ il y a absence de multicolinéarité

Exercice 1 :

1)

Sans constante → centrée → variable à laquelle enlevé sa moyenne

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

$$(X'X)^{-1} = \frac{1}{\det(X'X)} \begin{pmatrix} 113 & -150 \\ -150 & 200 \end{pmatrix}$$

$$\det(X'X) = \begin{vmatrix} 200 & 150 \\ 150 & 113 \end{vmatrix} = 100$$

$$\Rightarrow \hat{\beta} = \frac{1}{100} \begin{pmatrix} 113 & -150 \\ -150 & 200 \end{pmatrix} \begin{pmatrix} 350 \\ 263 \end{pmatrix}$$

$$\rightarrow \hat{\beta} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$$

2ème cas : Suppression d'une observation

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

$$(X'X)^{-1} = \frac{1}{\det(X'X)} \begin{pmatrix} 112 & -149 \\ -149 & 199 \end{pmatrix}$$

$$\rightarrow \det(X'X) = \begin{vmatrix} 199 & 149 \\ 149 & 112 \end{vmatrix} = 87$$

$$\rightarrow \hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \frac{1}{87} \begin{pmatrix} 112 & -149 \\ -149 & 199 \end{pmatrix} \times \begin{pmatrix} 347{,}5 \\ 261{,}5 \end{pmatrix} = \begin{pmatrix} -0{,}5 \\ 3 \end{pmatrix}$$

2) coeff de corrélation entre x_1 et x_2 :

$$\rho_{x_1,x_2} = \frac{cov(x_1,x_2)}{\sqrt{V(x_1)\cdot V(x_2)}} = \frac{\frac{1}{T}\sum(x_{1t}-\bar{x})(x_{2t}-\bar{x}_2)}{\frac{1}{T}\sqrt{\sum(x_{1t}-\bar{x}_1)^2\sum(x_{2t}-\bar{x}_2)^2}}$$

$$X'X = \begin{pmatrix} \sum (x_{1t}-\bar{x}_1)^2 & \sum (x_{1t}-\bar{x}_1)(x_{2t}-\bar{x}_2) \\ \sum (x_{1t}-\bar{x}_1)(x_{2t}-\bar{x}_2) & \sum (x_{2t}-\bar{x}_2)^2 \end{pmatrix}$$

$$= \begin{pmatrix} 200 & 150 \\ 150 & 113 \end{pmatrix}$$

$$\Rightarrow \rho_{x_1 x_2} = \frac{150}{\sqrt{200 \times 113}}$$

$$= 0,998.$$

Remarque : Dans le cas où le modèle est sans constante et les variables sont centrées.

$$X'Y = \begin{pmatrix} \sum (x_{1t}-\bar{x}_1)(y_t-\bar{y}) \\ \sum (x_{2t}-\bar{x}_2)(y_t-\bar{y}) \end{pmatrix}$$

3) Une faible modification du nombre d'observations entraîne une profonde modification des valeurs estimées des coefficients qui sont alors instables.
Ceci est la conséquence directe de la très forte corrélation entre x_1 et x_2.

Le risque de multicolinéarité est donc important afin d'éviter ce problème, il convient de détecter une éventuelle multicolinéarité lors de l'estimation d'un modèle.

Exercice 3 :

1) Interprétation économique

$y_t = a + bx_t + cz_t + u_t \Rightarrow$ modèle log-linéaire

* $b = \frac{\partial y_t}{\partial x_t}$: c'est l'élasticité de la production par rapport au travail
* $c = \frac{\partial y_t}{\partial z_t}$: c'est l'élasticité de la production par rapport au capital
* $a = E(y_t / x_t = z_t = 0)$. c'est la production moyenne autonome (indépendamment des facteurs capital et travail)

2) coefficient de détermination :

$$R^2 = 1 - \frac{SCR}{SCT} = 1 - \frac{SCR}{\sum(y_t - \bar{y})^2}$$

$$\Rightarrow R^2 = 1 - \frac{1,4}{10} = 0,86.$$

$\Rightarrow$ On a une bonne qualité d'ajustement linéaire du modèle.

3) Test de Klein

$$\rho_{x,z} = \frac{cov(x,z)}{\sqrt{V(x) \cdot V(z)}} = \frac{\sum (x_t - \bar{x})(z_t - \bar{z})}{\sqrt{\sum (x_t - \bar{x})^2 \sum (z_t - \bar{z})^2}}$$

$$= \frac{\sum x_t z_t - T\bar{x}\bar{z}}{\sqrt{\left(\sum x_t^2 - T\bar{x}^2\right)\left(\sum z_t^2 - T\bar{z}^2\right)}}$$

Sur la matrice $(X'X)$ on a :

$$X'X = \begin{pmatrix} T & \sum x_t & \sum z_t \\ \cdot & \sum x_t^2 & \sum x_t \cdot z_t \\ \cdot & \cdot & \sum z_t^2 \end{pmatrix}$$ (matrice symétrique)

Avec :

$$\bar{x} = \frac{1}{T}\sum x_t = \frac{230}{23} = 10$$

$$\bar{z} = \frac{1}{T}\sum z_t = \frac{115}{23} = 5$$

$$\Rightarrow \rho_{x,z} = \frac{1158 - 23 \times 10 \times 5}{\sqrt{(2312 - 23 \times 10^2)(587 - 23 \times 5^2)}}$$

$$= \frac{8}{12}$$

$$\Rightarrow \rho_{x,z} = 0,667$$

$\Rightarrow \rho^2_{x,z} = 0,44 < R^2 = 0,86 \Rightarrow$ absence de multicolinéarité.

* Test de Farrar et Glauber

$1^{ère}$ étape :

$$D = \begin{vmatrix} 1 & \rho_{x,z} \\ \rho_{z,x} & 1 \end{vmatrix} = 1 - \rho^2_{x,z} = 1 - 0,44$$

$$\Rightarrow \boxed{D = 0,56}$$

$2^{ème}$ étape : Test de Khi-deux :

- H_0 : les séries sont orthogonales

H_1 : les séries sont dépendantes.

• Statistique et loi :

$$\chi^2_c = -\left[T - 1 - \frac{1}{6}(2K+5)\right] \log D \rightsquigarrow \chi^2_\alpha\left(\frac{1}{2}K(K-1)\right)$$

$$\chi^2_{5\%}(3)$$

• Règle de décision

· Si $\chi^2_c \leq \chi^2_\alpha$ donc H_0 est vraie.

· Si $\chi^2_c > \chi^2_\alpha$ donc H_1 est vraie.

• conclusion

$\chi^2_c = 11,631 > \chi^2_{5\%}(3) = 7,815$

On accepte H_1 : les séries sont colinéaires $\Rightarrow$ présence de multicolinéarité

(Les 2 tests sont contradictoires)

On remarque que les 2 tests conduisent à des résultats différents dans ce cas, on privilégie le test de Farrar et Glauber car son fondement théorique est plus affirmé et il s'agit d'un test plus puissant.

4) Les conséquences d'une éventuelle multicolinéarité sont :

• augmentation de la variance estimée de certains coefficients ce qui implique que les T de student diminuent par conséquent les coefficients deviennent moins précis et non significatifs

• Instabilité des estimateurs des coefficients des moindres carrés en effet, les faibles fluctuations concernant les données entraînent de fortes variations des valeurs estimées des coefficients

• En cas de multicolinéarité parfaite, $\det(X'X) = 0$ donc $(X'X)^{-1}$ n'existe pas, l'estimation des coefficients par les MCO est alors impossible

Exercice 4 :

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

$$\hat{y_t} = 1,471 + 0,822 x_t \quad , R^2 = 0,99 \, , \, cov(\hat{\beta_0}, \hat{\beta_1}) = -0,009$$

1) On a $R^2 = 0,99$, donc on a une bonne qualité d'ajustement linéaire du modèle.

2) Il s'agit d'effectuer un test de significativité individuelle pour le paramètre β_1

* $H_0 : \beta_1 = 0$ contre $H_1 : \beta_1 \neq 0$
* Sous H_0 vraie, $\frac{\hat{\beta_1}}{\hat{\sigma}_{\hat{\beta_1}}} \sim St(T-2) = St(18)$
* Règle de décision :

Si $t_c = \left| \frac{\hat{\beta_1}}{\hat{\sigma}_{\hat{\beta_1}}} \right| \leq t^{T-2}_{\alpha/2}$ alors H_0 est vraie : $\beta_1 = 0$

(β_1 statistiquement non significatif)

* Si $t_c > t^{T-2}_{\alpha/2}$ alors H_1 est vraie : $\beta_1 \neq 0$

(β_1 statistiquement significatif)

* Conclusion :

. On a : $t_c = \left| \frac{0,822}{0,019} \right| = 43,263$

et $t^{T-2}_{\alpha/2} = t^{18}_{0,025} = 2,101$

$\Rightarrow t_c > t^{T-2}_{\alpha/2}$ donc H_1 est vraie : $\beta_1 \neq 0$

$\Rightarrow \beta_1$ est statistiquement significatif

$\Rightarrow$ Oui, la quantité utilisée d'engrais est un facteur déterminant du rendement de blé.

8

T ... 1554
1 5 2 467

3) Dans la 1ère étape de la méthode 2SLS, il s'agit de régresser la variable endogène (x_t) sur l'instrument (z_t). L'équation s'écrit alors comme suit :

$$x_t = \alpha_0 + \alpha_1 z_t + u_t$$

4) Test d'endogénéité d'Hausman :

- H_0 : $cov(x_t, \varepsilon_t) = 0$ (pas de problème d'endogénéité) ⇒ Retenir les MCO

contre

H_1 $cov(x_t, \varepsilon_t) \neq 0$ (x_t est endogène) ⇒ retenir la méthode VI

- Statistique du test et loi :

$$H = (\hat{\beta}_{VI} - \hat{\beta}_{MCO})' (\hat{\Omega}_{\hat{\beta}_{VI}} - \hat{\Omega}_{\hat{\beta}_{MCO}})^{-1} (\hat{\beta}_{VI} - \hat{\beta}_{MCO})$$

$$\rightsquigarrow \chi^2_{\alpha\%}(k) = \chi^2_{5\%}(2)$$

- Règle de décision :
  - Si $H \leq \chi^2_\alpha$ alors H_0 est vraie
  - Si $H > \chi^2_\alpha$ alors H_1 est vraie

- Conclusion :

On a : $\chi^2_{5\%}(2) = 5,991$

$$\hat{\beta}_{VI} = \begin{pmatrix} 2,153 \\ 0,795 \end{pmatrix} ; \hat{\beta}_{MCO} = \begin{pmatrix} 1,471 \\ 0,822 \end{pmatrix}$$

$$\hat{\beta}_{VI} - \hat{\beta}_{MCO} = \begin{pmatrix} 0,682 \\ -0,027 \end{pmatrix}$$

2,334 = 2,667

4) Test d'homogénéité d'Hausman

- H_0 : $cov(x_t, \varepsilon_t) = 0 \Rightarrow$ pas de problème d'endogénéité $\Rightarrow$ Retenir les MCO
- H_1 : $cov(x_t, \varepsilon_t) \neq 0 \Rightarrow$ il y a un problème d'endogénéité $\Rightarrow$ Retenir la méthode des VI

- Statistique et loi :

$$H = (\hat{\beta}_{VI} - \hat{\beta}_{MCO})' (\hat{\Omega}_{\hat{\beta}_{VI}} - \hat{\Omega}_{\hat{\beta}_{MCO}})^{-1} (\hat{\beta}_{VI} - \hat{\beta}_{MCO}) \rightsquigarrow \chi^2_\alpha(k) = \chi^2_{5\%}(2)$$

- Règle de décision :
  - Si $H \leq \chi^2_\alpha$: H_0 est vraie
  - Si $H > \chi^2_\alpha$: H_1 est vraie
- Conclusion :

$$H = 6,522 \quad > \quad \chi^2_{5\%}(2) = 5,981$$

Donc on retient H_1 donc x_t est endogène et par conséquent on doit retenir la méthode des VI.