Exercice :

Soit le modèle suivant : $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \quad \forall t = 1, \dots, 8$

il repose sur les hypothèses suivantes :

$E(\varepsilon_t) = 0, \quad V(\varepsilon_t) = \sigma^2, \quad Cov(\varepsilon_t, \varepsilon_s) = 0 \ \forall t \neq s$

Le tableau des données relatives aux variables du modèle se présente comme suit :

x_t	2	3	3	2	5	7	4	7
y_t	3	1	3	4	6	5	2	5

On donne par ailleurs $SCR = 12,719$

1/ On demande de procéder aux tests de détectation d'hétéroscédasticité suivants, en expliquant le principe de chaque test :

a/ Test de Goldfeld - Quandt :

Ce test se déroule en 3 étapes :

1er étape : Réordonner les observations de l'échantillon suivant les valeurs croissantes de la variable x_t.

x_t	2	2	3	3	4	5	7	7
y_t	3	4	1	3	2	6	5	5

2ème étape : ommettre $c = \frac{T}{4} = 2$ valeurs centrales et diviser le reste en deux sous échantillons.

ech1 ech2

3ème étape : Effectuer deux regressions

- 1er regression : $\hat{y}_t = 8,614 - 2,548\, x_t$ $\begin{cases} x_t = 2 \rightarrow \hat{y}_t = 3,518 \\ x_t = 2 \rightarrow \hat{y}_t = 3,518 \\ x_t = 3 \rightarrow \hat{y}_t = 0,97 \end{cases} \Rightarrow SCR_1 = 0,502$
- 2ème regression : $\hat{y}_t = 8,589 - 0,514\, x_t$ $\begin{cases} x_t = 5 \rightarrow \hat{y}_t = 6,019 \\ x_t = 7 \rightarrow \hat{y}_t = 4,991 \\ x_t = 7 \rightarrow \hat{y}_t = 4,991 \end{cases} \Rightarrow SCR_2 = 0,000523$

Statistique et sa loi :

sous H_0 vraie, on a $F_c = \frac{SCR_2}{SCR_1} \sim F_\alpha(1,1)$ $\quad 1 = \frac{\overset{=8}{T} - \overset{c=2}{c}}{2} - K \quad K = k+1 = 1+1 = 2$

On a $F_c = 0,001 \ll F_\alpha(1,1) = 161,4$ d'où H_0 est vraie $\Rightarrow$ homoscédasticité d'erreurs

b/ Test de White, sachant que le coeff de termination calculé à partir de la regression du carré des résidus est tel que $R^2 = 0,454$.

On doit régresser le carré des résidus estimés par les MCO dans le modèle initial sur la variable explicative et son carré :

$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 x_t + \beta_1 x_t^2 + \mu_t$

la statistique et sa loi :

Sous H_0 vraie : on a : $\chi^2_c = TR^2 \sim \chi^2_\alpha(2R) = \chi^2_{5\%}(2)$

on a $\chi^2_c = 3,632 < \chi^2_\alpha(2) = 5,991$ alors H_0 est vraie $\Rightarrow$ homoscedasticité des erreurs.

c/ Conclure :

Les deux tests sont concordants et confirment l'hypothèse d'homosced des erreurs.

2/ La méthode des MCO est-elle appropriée pour estimer les paramètres de ce modèle ? Justifier la réponse.

La méthode des MCO est appropriée pour estimer les paramètres de ce modèle puisque les erreurs sont homoscedastiques et non auto-corrélées.

3/ Calculer l'estimateur $\hat{\beta}_1$ ainsi que sa variance estimée.

$$\hat{\beta}_{1MCO} = \frac{\Sigma x_t y_t - T\bar{x}\bar{y}}{\Sigma x_t^2 - T\bar{x}^2} \quad \text{avec } \bar{x} = \frac{1}{T}\Sigma x_t = 4,125 \text{ et } \bar{y} = \frac{1}{T}\Sigma y_t = 3,625$$

$$= \frac{134 - 8\times 4,125\times 3,625}{165 - 8\times(4,125)^2} = 0,498$$

$$\hat{V}(\hat{\beta}_{1MCO}) = \frac{\hat{\sigma}^2}{\Sigma(x_t-\bar{x})^2} \quad \text{avec } \hat{\sigma}^2 = \frac{SCR}{T-2} = \frac{12,719}{6}$$

$$= 0,0734$$

4/ On suppose maintenant que $V(\varepsilon_t) = \sigma^2 x_t^4$. L'estimateur des MCO est-il BLUE ? Justifier la réponse

Estimateur $\hat{\beta}_1$ de β_1 par la méthode des MCO n'est pas BLUE car l'une des hypothèses des MCO à savoir l'homoscedasticité n'est plus vérifiée puisque $V(\varepsilon_t) = \sigma^2 x_t^4$.

5/ Proposer alors une méthode alternative d'estimation en appliquant la procédure de correction.

Sous l'hypothèse d'hétéroscedasticité, la méthode appropriée pour estimer le modèle est la méthode des MCG donnée par la formule suivante : $\hat{\beta}_{MCG} = ({}^tX.V^{-1}X)^{-1}({}^tXV^{-1}Y)$.

Appliquer la MCG revient à appliquer la MCO sur le modèle transformé

...es avoir diviser tous les termes par x_t^2. Em effet : $V(\varepsilon_t) = \sigma^2 . x_t^4$

$\Leftrightarrow \sigma^2 = \frac{V(\varepsilon_t)}{x_t^4} = V\left(\frac{\varepsilon_t}{x_t^2}\right)$

modèle initial : $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$ modèle transformé : $\frac{y_t}{x_t^2} = \frac{\beta_0}{x_t^2} + \beta_1 \frac{x_t}{x_t^2} + \frac{\varepsilon_t}{x_t^2}$

$$V = \begin{pmatrix} x_1^4 & & 0 \\ & x_2^4 \ddots & \\ 0 & & x_8^4 \end{pmatrix}$$

Exercice :

Om considère un modèle de la forme : $y_t = a + bx_t + cz_t + \varepsilon_t, \quad t = 1, \dots, T$

Les termes aléatoires vérifient les hypothèses suivantes : $E(\varepsilon_t) = 0$

$$E(\varepsilon_t \varepsilon_s) = \begin{cases} 0 & t \neq s \\ \sigma^2 z_t & t = s \end{cases}$$ l'hypoth d'autocor est vérifiée

1/ Donner la matrice de variance-covariance des erreurs :

La matrice de variance-covariance des erreurs, notée Ω_ε, est donnée par l'expression suivante : $\Omega_\varepsilon = \sigma^2 V \Leftrightarrow$

$$\Omega_\varepsilon = \sigma^2 \begin{pmatrix} z_1 & & 0 \\ & \ddots & \\ 0 & & z_T \end{pmatrix}$$

2/ Quelle méthode d'estimation doit-on utiliser pour estimer le paramètre b ?

Puisque la variance $V(\varepsilon_t) = \sigma^2 z_t \neq$ constante, les erreurs sont donc hétéroscédastiques. La méthode d'estimation appropriée est la MCG ou encore la MCP.

3/ Procéder à une correction qui permet d'obtenir un modèle ayant des erreurs homoscédastiques.

Puisque $V(\varepsilon_t) = \sigma^2 . z_t \Leftrightarrow \sigma^2 = \frac{V(\varepsilon_t)}{z_t} = V\left(\frac{\varepsilon_t}{\sqrt{z_t}}\right)$. Donc, il faut diviser tous les termes du modèle initial par $(\sqrt{z_t}) \Rightarrow \frac{y_t}{\sqrt{z_t}} = \frac{a}{\sqrt{z_t}} + \frac{b x_t}{\sqrt{z_t}} + \frac{c z_t}{\sqrt{z_t}} + \frac{\varepsilon_t}{\sqrt{z_t}}$

4/ Déterminer l'estimateur des MCG de b et sa variance :

$\hat{b}_{MCG} = \begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{pmatrix} = ({}^tX . V^{-1} . X)^{-1} ({}^tX . V^{-1} . Y)$

vecteur des estimateurs

variance $\hat{\Omega}_{\hat{b}_{MCG}} = ({}^tX . \Omega_\varepsilon^{-1} . X)^{-1}$

$= \sigma^2 ({}^tX . V^{-1} . X)^{-1}$

$\neq$

$\hat{\Omega}_{\hat{b}_{MCO}} = \sigma^2 ({}^tXX)^{-1}$ $V = I$

Exercice :

On considère le modèle suivant : $y_t = a_0 + a_1 x_{1t} + a_2 x_{2t} + u_t$, $\forall t = 1, \dots, 25$

avec :

$$Var(u) = \Omega_u = \begin{pmatrix} 2 & & 0 \\ & 4 & \\ & & \ddots \\ 0 & & & 50 \end{pmatrix}$$

la matrice des variances covariances des erreurs

1/ Les termes d'erreurs sont-ils autocorrélés ?

Non, on n'a pas un problème d'autocorrélation des erreurs car d'après la matrice Ω_u, les termes en dehors de la diagonale principale correspondent à des variances nulles donc les ε_t ne sont pas autocorrélés.

2/ Les termes d'erreurs sont-ils homoscédastiques ? s'ils ne le sont pas quelle transformation doit-être réalisée pour qu'ils deviennent homoscéd ?

Non, les termes d'erreurs ne sont pas homoscédastiques puisque la diagonale principale de Ω_u est représentée par des variances qui ne sont pas constantes.

$$\Omega_u = \begin{pmatrix} 2 & & & 0 \\ & 4 & & \\ & & 6 & \\ 0 & & & \ddots & 50 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 & & 0 \\ & 2 & \\ & & \ddots \\ 0 & & & 25 \end{pmatrix} = 2\sigma$$

on remarque que $t = 1, \dots, 25$

$V(u_t) = 2t = \sigma^2 t$

$\Leftrightarrow \sigma^2 = \frac{V(u_t)}{t} = V\left(\frac{u_t}{\sqrt{t}}\right)$

Pour corriger l'hétéroscédasticité, tous les termes du modèle initial doivent être divisés par $\sqrt{t}$

Exercice :

L'estimation d'un modèle de régression a fourni les résultats suivants :

$\hat{y}_t = -7{,}5 + 3{,}8x_t + 2{,}4 z_t$ $\quad T = 33 \quad \Sigma(y_t - \bar{y})^2 = 25 \quad \Sigma \hat{\varepsilon}_t = 5$

$\Sigma \hat{\varepsilon}_t \hat{\varepsilon}_{t-1} = -3{,}6 \quad \Sigma(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2 = 17{,}2$

1/ Calculer et interpréter le coefficient de détermination R^2

$R^2 = 1 - \frac{SCR}{SCT} = 1 - \frac{5}{25} = 0{,}8 \Rightarrow$ On a une bonne qualité d'ajustement

2/ Tester au seuil de 5% la significativité globale du modèle :

Test de significativité globale : Test de Fisher

$H_0 : \beta_1 = \beta_2 = 0$ contre $H_1 : \exists i$ tq $\beta_i \neq 0$

sous H_0 vraie : on a $F_c = \frac{R^2}{1-R^2} \cdot \frac{T-K}{K-1} \sim F_\alpha(K-1, T-K) = F_{5\%}(2,30)$

Règle de décision :

- Si $F_c \leq F_\alpha$ alors H_0 est vraie
- Sinon, H_1 est vraie

Conclusion : $F_c = \frac{0{,}8}{1-0{,}8} \times \frac{30}{2} = 60 > F_{5\%}(2,30) = 3{,}32$

donc on accepte H_1 : d'où le modèle est globalement significatif au seuil de 5%.

3/ Tester l'hypothèse d'autocorrélation des erreurs d'ordre 1. S'il y a autocorrélation, estimer le coefficient. Quelle méthode d'estimation peut-on employer ?

Test de Durbin et Watson :

Hypothèse : $H_0 : \rho = 0$ contre $H_1 : \rho \neq 0$ avec ρ étant le coeff d'autocorr d'ordre 1 des erreurs

Statistique du test : $DW = \frac{\sum_{t=2}^{T}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{T}\hat{\varepsilon}_t^2} = \frac{17{,}2}{5} = 3{,}44$

Règle de décision : 0 $\rho>0$ d_1 (1,32) d_2 (1,58) $\rho=0$ 2 $\rho=0$ $4-d_2$ (2,42) $4-d_1$ (2,68) $\rho<0$ 4

On remarque que $4 - d_1 < DW = 3{,}44 < 4$, donc on rejette H_0 et on a une autocorrélation négative des erreurs ($\rho < 0$) $\quad \hat{\rho} = 1 - \frac{DW}{2} = -0{,}72$

La méthode des MCO n'est plus valide, on doit appliquer la MCG sur le modèle initial ce qui revient à Appliquer les MCO sur le modèle quasi-diff

Ex6

① Les conditions d'utilisation du test de Durbin et Watson sont :
- le modèle doit contenir une constante.
- le nombre d'observations $T \geq 15$.
- Il s'agit de tester une autocorrélation des erreurs d'ordre 1.

② $$DW = \frac{\sum_{t=2}^{T} (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{T} \hat{\varepsilon}_t^2}$$

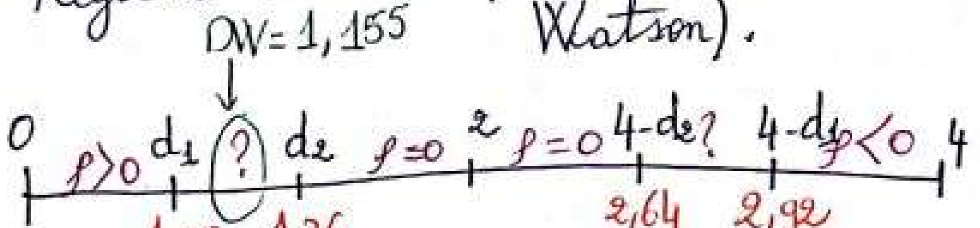
(voir 3ème colonne)

$\Longrightarrow DW = 1,155$

③ Test de Durbin et Watson :

* Hypothèses :
$H_0 : \rho = 0$ contre $H_1 : \rho \neq 0$

* Règle de décision (à travers l'échelle de Durbin et Watson).



D'après la table de Durbin et Watson, pour $T = 15$ et $k = 1 \Rightarrow d_1 = 1,08$ et $d_2 = 1,36$.

⇨ $d_1 < DW = 1,155 < d_2$: il s'agit d'une zone de doute ou d'indécision, mais il serait plus prudent d'accepter H_1 et de supposer l'autocorrélation des erreurs d'ordre 1.

Suite de l'ex6 :

On donne les résultats d'estimation suivants :

$$\hat{\varepsilon}_t = 140{,}868 - 18{,}053\,x_t + 0{,}408\,\hat{\varepsilon}_{t-1} - 1{,}065\,\hat{\varepsilon}_{t-2}$$

avec $R^2 = 0{,}52$.

* Test de Khi-2 :

$$\begin{cases} H_0 : \rho_1 = \rho_2 = 0 \text{ contre} \\ H_1 : \exists \text{ au moins un } \rho_i \neq 0 \ (i = 1,2) \\ \quad \Rightarrow \varphi_1 \neq 0 \text{ ou } \varphi_2 \neq 0 \end{cases}$$

* Statistique utilisée et sa loi :

$$\chi_c^2 = (T-p) \times R^2 \rightsquigarrow \chi_\alpha^2(p) = \chi_{5\%}^2(2)$$

* Règle de décision :

. Si $\chi_c^2 \leqslant \chi_\alpha^2$ alors H_0 est vraie

. Si $\chi_c^2 > \chi_\alpha^2$ alors H_1 est vraie

Conclusion :

$$\chi_c^2 = (15-2) \times 0{,}52 = 6{,}76 > \chi_{5\%}^2(2) = 5{,}991$$

$\Rightarrow$ On accepte H_1

$\Rightarrow$ il existe une autocorrélation d'ordre 2 des erreurs.

* Test de Fisher :

$$\begin{cases} H_0 : \rho_1 = \rho_2 = 0 \text{ contre} \\ H_1 : \rho_1 \neq 0 \text{ ou } \rho_2 \neq 0 \end{cases}$$

* Stat et loi :

$$F_c = \frac{R^2}{1-R^2} \times \frac{T-K-2p}{p} \longrightarrow F_\alpha(p, T-K-2p)$$

* Règle de décision :
  * Si $F_C \leqslant F_\alpha$ alors H_0 est vraie
  * Si $F_C > F_\alpha$ alors H_1 est vraie

Conclusion :

$$F_C = \frac{0,52}{1-0,52} \times \frac{15-2-4}{2} = 4,875$$

$$F_\alpha(p, T-K-2p) = F_{5\%}(2,9) = 4,26.$$

$\Rightarrow F_C > F_\alpha$ alors H_1 est vraie

$\Rightarrow$ Il existe une autocorrélation des erreurs d'ordre 2 selon le test de Fisher.

$\boxed{Ex7}$: $(A'A)^{-1}$ donnée / $A'\hat{\mathcal{E}}$)

$$y_t = \alpha + \beta x_t + \mathcal{E}_t \quad (1)$$

(1) * $\hat{\beta} = \dfrac{\sum x_t y_t - T\bar{x}\bar{y}}{\sum x_t^2 - T\bar{x}^2}$ avec $\begin{cases} \bar{x} = \frac{1}{T}\sum x_t = 0,2 \\ \bar{y} = \frac{1}{T}\sum y_t = 2,4 \end{cases}$

$= 3,429.$

* $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 1,714$

$\Rightarrow \hat{y}_t = 1,714 + 3,429\, x_t$

$\Rightarrow \hat{\mathcal{E}}_t = y_t - \hat{y}_t =$

-8,143	0,857
6,286	7,286
-0,714	
6,857	
-8,285	
1,715	
-6,143	
0,286	

(2) $SCR = \sum_{t=1}^{10} \hat{\mathcal{E}}_t^2 = 316,571$

* $R^2 = 1 - \dfrac{SCR}{SCT}$ avec $SCT = \sum(y_t - \bar{y})^2 = 382,4$

$= \sum y_t^2 - T\bar{y}^2$

$\Rightarrow R^2 = 0,172.$

Suite de l'ex 7 :

(3) On va appliquer le test de Breusch-Godfrey car $T < 15$.

Ce test se base sur l'estimation par les MCO de l'équation intermédiaire suivante :

$$\hat{\varepsilon}_t = a + b x_t + \rho \hat{\varepsilon}_{t-1} + v_t \quad (2)$$

avec : $\hat{\varepsilon}_t$ étant le résidu d'estimation par les MCO du modèle initial $y_t = \alpha + \beta x_t + \varepsilon_t$ (1).

(4) Rappel :

Régression multiple :

Théorème :

$$SCR_{(1)} = Y'Y - \hat{\beta}' X'Y \leftarrow \text{modèle}$$

$$\boxed{SCR_{(2)} = \hat{\varepsilon}'\hat{\varepsilon} - \hat{\beta}' A' \hat{\varepsilon}}$$

$$= \sum \hat{\varepsilon}_t^2 - \hat{\beta}' A' \hat{\varepsilon}$$

$$= SCR_{(1)} - \hat{\beta} A' \hat{\varepsilon}$$

$$* \ \hat{\beta} = \begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{\rho} \end{pmatrix} = (A'A)^{-1} \cdot A'\hat{\varepsilon} = \begin{pmatrix} 0,337 \\ 1,894 \\ -0,550 \end{pmatrix}$$

* Étape préliminaire :

$$R^2_{(2)} = 1 - \frac{SCR_{(2)}}{SCT_{(2)}}$$

et $SCT_{(2)} = \sum \hat{\varepsilon}_t^2 - T \overline{\hat{\varepsilon}}^2$, avec $\overline{\hat{\varepsilon}} = 0$

$$\Rightarrow \boxed{SCT_{(2)} = \sum \hat{\varepsilon}_t^2 = SCR_{(1)} = 316,571}$$

Ainsi $SCR_{(2)} = 316,571 - \begin{pmatrix} 0,554 & 1,894 & -0,550 \end{pmatrix} \begin{pmatrix} 8,145 \\ 8,141 \\ -137,393 \end{pmatrix}$

$$= 316,571 - 93,730.$$

$$\Rightarrow \boxed{SCR_{(2)} = 222,841}$$

* Étape préliminaire :

$$R^2_{(2)} = 1 - \frac{SCR_{(2)}}{SCT_{(2)}}$$

et $SCT_{(2)} = \sum \hat{\varepsilon}_t^2 - T \underbrace{\bar{\hat{\varepsilon}}^2}_{=0}$

$$\Rightarrow \boxed{SCT_{(2)} = \sum \hat{\varepsilon}_t^2 = SCR_{(1)} = 316,571}$$

$$\Rightarrow R^2_{(2)} = 1 - \frac{222,841}{316,571}$$

$$\Rightarrow \boxed{R^2_{(2)} = 0,296}$$

* Hypothèses du test :

$H_0 : \rho = 0$ contre $H_1 : \rho \neq 0$

1ère manière : Test de Fisher :

Statistique et loi :

$$F_c = \frac{R^2}{1 - R^2} \times \frac{T - K - 2p}{p} \rightarrow \underbrace{F_\alpha(p, T - K - 2p)}_{= F_{5\%}(1, 6)}$$

avec : $R^2 = R^2_{(2)}$

* Règle de décision :

. Si $F_C \leqslant F_\alpha$ alors H_0 est vraie.

. Si $F_C > F_\alpha$ alors H_1 est vraie.

* Conclusion :

$$F_C = \frac{0,296}{1-0,296} \times \frac{6}{1} = 2,523.$$

$$F_{5\%}(1,6) = 5,99$$

$\Rightarrow F_C < F_\alpha \Rightarrow$ on accepte H_0

2ème manière : Test de Khi-deux

* Statistique et loi :

$$\chi^2_c = (T-p) \times R^2 \longrightarrow \chi^2_\alpha(p) = \chi^2_{5\%}(1)$$

avec : $R^2 = R^2_{(2)}$

* Règle de décision :

. Si $\chi^2_c \leqslant \chi^2_\alpha$ alors H_0 est vraie

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

* Conclusion :

$$\chi^2_c = (10-1) \times 0,296 = 2,664$$

$$\chi^2_{5\%}(1) = 3,841$$

$\Rightarrow \chi^2_c \leqslant \chi^2_\alpha$ alors on accepte H_0.

Dans les deux cas, on conclut de l'absence d'autocorrélation des erreurs d'ordre 1 au risque de 5%.