Université d’Orléans
Faculté de Droit d’Economie et de Gestion

Master 1 ESA
Econométrie et Statistique Appliquée

TD SERIES TEMPORELLES

(Polycopié d’exercices)

Sessi TOKPAVI

Année Universitaire 2006/07

SEANCE 1 & 2

EXERCICE 1 : Processus Moving Average (MA), *Théorème de Wold, stationnarité et inversibilité*

1. Enoncez le théorème de Wold et rappelez brièvement son intérêt pour la modélisation des processus linéaires.
2. Donnez l'ordre des différents processus MA suivants et précisez s'ils sont stationnaires ou non (justifiez la réponse à la dernière question). Le processus u_t est un bruit blanc et L est l'opérateur de retard

$$\text{(a)} \quad x_t = (1 - 0.8L)u_t$$

$$\text{(b)} \quad x_t = (1 - 0.4L + 1.2L^2)u_t$$

$$\text{(c)} \quad x_t = \left(\sum_{i=0}^{\infty} (-0.5)^i L^i \right) u_t$$

$$\text{(d)} \quad x_t = \left(\sum_{i=0}^{\infty} (1.8)^i L^i \right) u_t$$

3. En déduire une conclusion générale quant à la stationnarité des processus MA.
4. Expliquez pourquoi l'hypothèse d'inversibilité est souvent requise dans l'étude des processus linéaires stochastiques. Identifiez parmi les quatre processus précédents, ceux pour lesquels, la vérification de cette hypothèse n'a pas de sens. Pour les autres, précisez s'ils sont inversibles ou non.

EXERCICE 2 : *Etude d'un Processus MA(1)*

Soit le processus MA(1) suivant, où u_t est un bruit blanc de variance notée σ_u^2

$$x_t = (1 + 0.7L)u_t$$

1. Calculez l'espérance et la variance du processus x_t. Le processus est-il stationnaire ? Au vu de la conclusion tirée à la question 3) de l'exercice 1, le calcul des deux moments est-il nécessaire pour répondre à la question précédente ?
2. Le processus est-il inversible ?

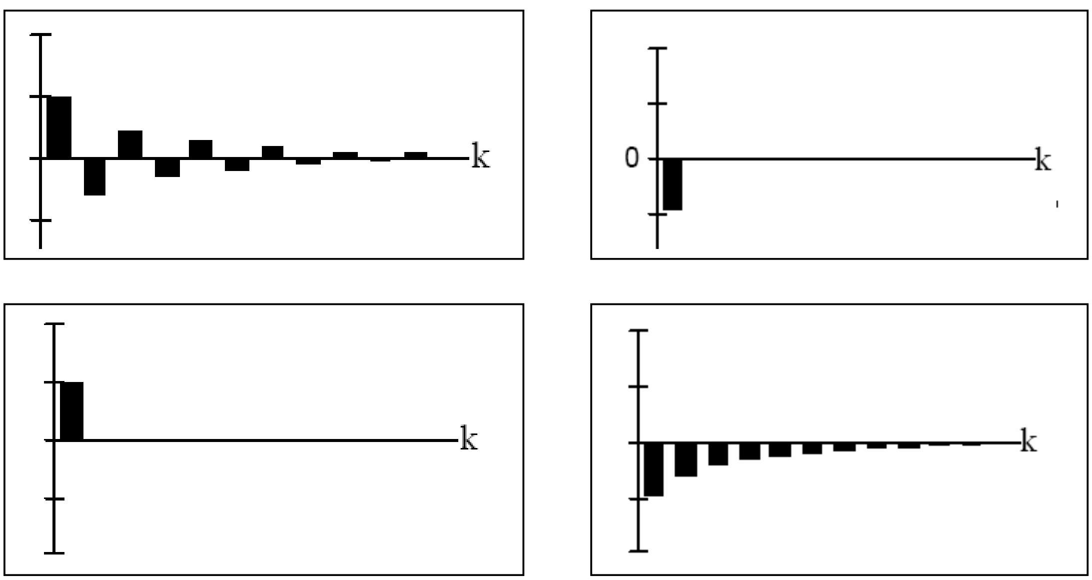
3. Calculez γ_k la fonction d'autocovariance de x_t et en déduire la fonction d'autocorrélation totale.
4. En utilisant les équations de Yule-Walker, donnez l'expression de la fonction d'autocorrélation partielle.

EXERCICE 3: *Etude d'un Processus MA(1) – suite.*

On considère à présent le processus MA(1) suivant : $x_t = (1 - 0.5L)u_t$

1. Reprendre les questions 3) et 4) de l'exercice précédent pour ce processus
2. Ci-dessous, sont représentées les fonctions d'autocorrélation totale et partielle des deux processus de l'exercice 2 et 3. Sans se préoccuper de l'ordre de grandeur, associez chaque graphique à celle de la fonction d'autocorrélation totale (et partielle) respective des processus.

(a) (b)



(c) (d)

EXERCICE 4 : *Moments Conditionnels*

La variable x_t est générée par le processus MA(1) d'écriture $x_t = u_t + \theta u_{t-1}$ où u_t est un processus en bruit blanc de variance $\sigma_u{}^2$.

1. Donnez les expressions des prédicteurs formulés en t pour t+1 et t+2. Précisez l'espérance et la variance conditionnelles (à l'information disponible en t) des erreurs

commises à l'horizon d'une période. Quelle est l'espérance non conditionnelle des erreurs à une période.

2. Calculez la MSE des prédicteurs à une période issus du processus x_t. Comparez-la à celle du processus MA(2) suivant $x_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}$, $\theta_1 \neq 0$ et $\theta_2 \neq 0$. A quelle condition les MSE des deux processus sont-elles égales.

EXERCICE 5 : *Etude d'un Processus MA(2)*

Soit le processus MA(2) suivant, où u_t est un bruit blanc de variance notée σ_u^2

$$x_{t+1} = (1 - 0.7L + 0.1L^2)u_{t+1}$$

1. Calculez l'espérance et la variance du processus x_{t+1}. Le processus est-il stationnaire ?
2. Le processus est-il inversible ? justifier.
3. On note $\widetilde{x}_t(1)$ la prévision à une période de x_{t+1} et $e_t(1)$ l'erreur de prévision correspondante. Calculez ces deux quantités, ainsi que l'espérance conditionnelle (en t) de $e_t(1)$. Quelle est l'implication de ce résultat en terme de propriété de $\widetilde{x}_t(1)$ en tant qu'estimateur de x_{t+1} .
4. Démontrez formellement que cette propriété demeure valable pour un processus $MA(\infty)$.
5. Calculez l'erreur de prévision à deux périodes $e_t(2)$ et établir la formule de la corrélation entre $e_t(1)$ et $e_t(2)$.
6. On dispose d'un échantillon de réalisations particulières de x_t, de taille T. On confie à un étudiant le soin de calculer pour une date t donnée, l'erreur de prévision moyenne à une période et à deux périodes. L'étudiant renvoie respectivement pour les valeurs moyennes de $e_t(1)$ et $e_t(2)$ -0.05 et -0.02. Que pensez-vous de ces résultats, même si vous ne disposez pas de l'échantillon en question.
7. Calculez γ_k la fonction d'autocovariance de x_t et en déduire la fonction d'autocorrélation totale. Quelle est la mémoire du processus ?
8. En utilisant les équations de Yule-Walker, donnez l'expression de la fonction d'autocorrélation partielle. Caractériser son évolution en fonction de k.

EXERCICE 6 :

Soit un processus stochastique x_t satisfaisant à la relation suivante, avec ε_t un bruit blanc :

$$x_t = 0.4x_{t-1} + \varepsilon_t$$

Une information supplémentaire vous est donnée, à savoir qu'il s'agit d'un processus MA(1), soit :

$$x_t = u_t - \theta u_{t-1}$$

1. Calculez θ.
2. Les valeurs de θ trouvées sont-elles toutes admissibles ? sinon pourquoi ?
3. Calculez ϕ_{22}, ϕ_{33} et représenter la fonction d'autocorrélation partielle pour k=0, 1, 2, 3

EXERCICE 7 :

1. Etudiez la stationnarité et l'inversibilité des deux processus suivants :

$$(1) \quad x_t = \Delta y_t \quad \text{avec} \quad y_t = at + b + \varepsilon_t$$

$$x_t = \Delta^2 y_t \quad \text{avec} \quad y_t = at^2 + bt + c + \varepsilon_t$$

ε_t est un bruit blanc de variance σ^2.

2. Identifiez les deux processus.

SEANCE 3 & 4

EXERCICE 1 :

5. Expliquez brièvement pourquoi les processus AR(p) sont toujours inversibles et énoncez la (les) condition(s) de stationnarité.
6. Expliquez pourquoi les autocorrélations partielles d'ordres supérieurs à p, sont nulles pour un processus AR(p).
7. On considère à présent le processus (1) où ε_t est un bruit blanc de variance σ^2. On suppose que le processus a débuté à la période 0 telle que y_0 est la condition initiale connue.

$$(1) \quad y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$$

   a. En utilisant la méthode d'itération Backward, exprimez y_t en fonction de la séquence $\{\varepsilon_t\}$, de y_0 et des paramètres du modèle (1).
   b. Calculez l'espérance de y_t en utilisant l'expression trouvée en i). Sans postuler des conditions supplémentaires pour la dynamique de y_t, peut-on affirmer qu'il est stationnaire ? Si non, donnez les deux conditions supplémentaires qui assurent la stationnarité de y_t. Interprétez.
   c. En supposant les deux conditions vérifiées, calculer la variance du processus y_t.
8. Calculez les deux moments précédents (moyenne et variance de y_t) en utilisant directement l'expression (1).

EXERCICE 2 :

Soit le processus AR(p) suivant, supposé stationnaire, avec ε_t un bruit blanc de variance σ^2 :

$$y_t = \sum_{i=1}^{p} a_i y_{t-i} + \varepsilon_t$$

1. Donnez la formule de γ_k , l'autocovariance d'ordre k. En déduire celle de l'autocorrélation totale ρ_k .
2. Vérifiez si les deux processus suivants sont stationnaires. Utilisez les résultats de la question 1 et calculez les autocorrélations totales d'ordre 1, 2, et 3 pour ces processus :

   i. $y_t = 0.7 y_{t-1} - 0.49 y_{t-2} + \varepsilon_t$

   ii. $y_t = 0.8 y_{t-1} + \varepsilon_t$
3. Calculez pour ces deux processus, la fonction d'autocorrélation partielle.

EXERCICE 3 :

Soit $y_t = 1.1y_{t-1} - 0.3y_{t-2} + \varepsilon_t - 0.2\varepsilon_{t-1} - 0.15\varepsilon_{t-2}$

1. Le processus est-il stationnaire ? inversible ? Justifiez.
2. Calculer les coefficients d'autocorrélation ρ_1 et ρ_2 et le coefficient d'autocorrélation partielle ϕ_{22}.
3. Calculez les prévisions $y_t(l)$, l=1, 2, 3, 4.

EXERCICE 4 :

Soit y_t une série temporelle donnée par :

$$y_t = 0.2y_{t-1} + 0.15y_{t-2} + \varepsilon_t + 0.3\varepsilon_{t-2}$$

1. Vérifiez si y_t est stationnaire et expliquez brièvement les implications de cette condition
2. Donner la représentation $MA(\infty)$ de y_t et trouver sa fonction d'autocorrélation.
3. Montrez qu'il s'agit d'un processus sur-paramétré, c'est-à-dire qu'il peut être représenté par un processus plus simple.
4. Retrouvez alors les valeurs trouvées pour la fonction d'autocorrélation.

EXERCICE 5 : (*Exercice complémentaire*)

On considère le processus autorégressif suivant avec $|a_2| < 1$. ε_t est un bruit blanc de variance σ^2 :

$$y_t = a_0 + a_2 y_{t-2} + \varepsilon_t$$

1. Calculez :
   i. $E_{t-2}(y_t)$ ii. $E_{t-1}(y_t)$ iii. $E_t(y_{t+2})$ iv. $Cov(y_t, y_{t-1})$ v. $Cov(y_t, y_{t-2})$
   vi. les autocorrélations partielles d'ordre 1 et 2.
2. Déterminez l'expression de la prévision en t de y_{t+l} et celle de l'erreur de prévision correspondante $e_t(l)$. Calculez le corrélogramme de la séquence $\{e_t(l)\}$, c'est-à-dire $E_t(e_t(l))$, $Var(e_t(l))$ et $E(e_t(l), e_t(l-j))$ $\quad j = 0,\ldots,l$

NB : Les séances 5 et 6 seront consacrées à l'identification et à l'estimation des processus ARMA. Elles se dérouleront en salle informatique sous le logiciel SAS. Relisez le cours d'Introduction à SAS de M. Sébastien RINGUEDE.

SEANCE 5&6

Identification et Estimation des processus ARMA

EXERCICE :

A)

1. Rappelez sans faire de calcul l'évolution suivant k des fonctions d'autocorrélation totale (ρ_k) et partielle (ϕ_{kk}) d'un processus AR(1)
2. Simulez sous SAS (pour 100 périodes) le processus AR(1) gaussien suivant : $y_t = 0.2 + 0.7y_{t-1} + u_t$, avec u_t tirée d'une normale de moyenne 0 et de variance unitaire et $y_0 = 0.67$
3. Quel est selon vous, l'intérêt de fixer la valeur initiale à 0.67 (identifier 0.67 à l'un des moments du processus à simuler). Dans le cas où on poserait $y_0 = 1$ par exemple, quelle précaution faut-il prendre lors de la simulation ?
4. Graphez pour le processus les fonctions d'autocorrélations totales et partielles empiriques. [***Commande Identify de la proc ARIMA sous SAS***]. Comparez les valeurs trouvées aux valeurs théoriques espérées. Que constatez-vous ?
5. En utilisant la distribution asymptotique des corrélations partielles, vérifiez que $\phi_{11} \neq 0$, $\phi_{22} \neq 0$ et $\phi_{33} \neq 0$. Vérifiez de même que $\rho_1 \neq 0$, $\rho_2 \neq 0$ et $\rho_3 \neq 0$. [***en réalité SAS vous permet de répondre graphiquement à ces questions pour n'importe quelle valeur de k***]
6. On veut tester l'hypothèse jointe suivante : $H0 : \rho_1 = \rho_2 = ... = \rho_k$ pour k=6, 12, 18. A quoi correspond l'hypothèse nulle ? Quelle(s) statistique(s) peut-on utiliser pour répondre à la question posée par le test ?
7. Supposez maintenant que l'échantillon en face est issu d'un processus inconnu. Estimer les paramètres en supposons qu'il s'agit d'un processus AR(1). [***Commande Estimate <plot> de la proc ARIMA sous SAS***]. Quelle propriété doit vérifier les résidus ?
8. Le processus AR(1) simulé étant inversible, il admet une représentation MA(∞). En réalité un processus fini MA(q) suffit. Estimer les paramètres en supposant respectivement que les données en présence suivent respectivement un processus MA(1) et MA(2). Les résidus obtenus sont-ils respectivement des bruits blancs ?
9. Dans le cas où pour l'un ou l'autre des deux processus précédents, les coefficients sont tous significatifs avec l'hypothèse de bruit blanc pour les résidus, quel critère utilisez pour choisir le meilleur modèle parmi les trois [AR(1), MA(1) et MA(2)]
10. Conclusion : Rappeler la méthodologie de Box-Jenkins pour l'identification et l'estimation des processus ARMA.

B)

Cas pratique : récupérez les données qui vous sont fournies et utilisez la méthodologie de Box-Jenkins pour identifiez le processus suivi par le processus générateur.

SEANCE 7

Exercice1 : On considère le processus $(1-\phi L^4)y_t = c + u_t$, où c et ϕ sont des constantes et u_t un processus en bruits blancs de variance $\sigma_u^{\ 2}$.

1- A quelle(s) condition(s) ce processus est inversible ?
2- A quelle(s) condition(s) est-il stationnaire ?
3- Calculer $E_{t-2}[y_t]$, $E_{t-4}[y_t]$, $E_t[y_{t+4}]$, $E_{t-2}[y_{t+4}]$.
4- Calculez $Cov(y_t, y_{t-1})$, $Cov(y_t, y_{t-4})$, ainsi que les
5-
6- autocorrélations partielles ϕ_{11} et ϕ_{44}

Exercice 2 : Soit $x_t = (1-0.6L)(1-0.4L^4)u_t$, u_t est un bruit blanc de variance égale à 2.0. Quels sont les 6 premiers coefficients d'autocorrélations totales de x_t ? Sans la calculer, pouvez-vous décrire sa fonction d'autocorrélation partielle ?

Exercice 3 : On considère le processus MA saisonnier suivant sur données trimestrielles : $x_t = (1-0.6L)(1-0.3L^4)u_t$ u_t est un bruit blanc de variance égale à 1.0.

1- Quelle est la valeur des coefficients de corrélation $\rho_x(k) = corr(x_t, x_{t-k})$, $k = 1,...,10$?
2- Quelle est la mémoire de ce processus ?
3- Est-il stationnaire ?
4- Donnez la valeur des quantités suivantes : $E[x_t]$, $V[x_t]$, $E_t[x_{t+12}]$, $V_t[x_{t+12}]$, $E_t[x_{t+1}]$ et $V_t[x_{t+1}]$?

Exercice 4 : Sur une série constituée de 200 observations, on a calculé les 12 premiers coefficients d'autocorrélations totales et partielles. Ils vous sont indiqués ci-après. Quel processus sélectionneriez-vous ?

Autocorrelations

1 :	-0.3197042	-0.2070592	0.0246393	0.0067892	0.0152394	0.0169965
7 :	-0.0577921	0.0029184	-0.0234394	0.0140073	-0.0047081	0.0967007

Partial Autocorrelations

1 :	-0.3197042	-0.3444795	-0.2205285	-0.1889114	-0.1295193	-0.0787624
7 :	-0.1269198	-0.1114012	-0.1618641	-0.1509490	-0.1781872	-0.0355190

Exercice 5 : Le théorème de Wold affirme que toute variable stationnaire est la somme de variables indépendantes identiquement distribuées d'espérance nulle et de variance σ^2. En conséquence, il est impossible de prévoir l'évolution future d'une variable stationnaire. Que pensez-vous de cette dernière affirmation. Construisez un exemple simple illustrant votre propos.

SEANCE 8

(*Racines Unitaires, tests DF et ADF*)

Exercice1 : racine unitaire

Soient les processus suivants en y, x, v et z, avec ε^y, ε^x, ε^v et ε^z des bruits blancs :

$$y_t = 0.2y_{t-1} + 0.35y_{t-2} + \varepsilon_t^{\,y} \tag{1}$$

$$x_t = 0.7x_{t-1} + 0.35x_{t-2} + \varepsilon_t^{\,x} \tag{2}$$

$$v_t = v_{t-1} + 2v_{t-2} + \varepsilon_t^{\,v} \tag{3}$$

$$z_t = 0.1y_t + 0.05v_{t-1} + \varepsilon_t^{\,z} \tag{4}$$

Ces quatre processus sont-ils stationnaires ? Quel est leur ordre d'intégration ?

Exercice2 : Test de Dickey-Fuller

On recherche la présence de racine unitaire dans une variable quelconque. Pour ce faire, un test de Dickey-Fuller est effectué sur différents sous échantillons et sous différentes conditions.

1. Rappelez la logique de ce test (y compris équation estimée sous sa forme générale et hypothèses testées).

Déterminez à chaque fois si la variable x peut être considérée comme stationnaire aux seuils de risque exigés (***bien entendu, vous indiquerez clairement les valeurs critiques des tests, avec 4 chiffres après la virgule***).

2. 115 observations, avec constante, au seuil de risque de 5%. La statistique calculée est égale à -2.3714 ;
3. 61 observations, sans constante (ni trend), au seuil de risque de 1%. La statistique calculée est égale à -0.0047 ;
4. 37 observations, avec constante et trend, au seuil de risque de 10%. La statistique calculée est égale à -3.1987;

Exercice3 : Test de Dickey-Fuller Augmenté

On procède à un test de Dickey-Fuller Augmenté sur la série de la masse monétaire (*M3*) en France pour la période 1978 :4-2000 :2 (en fréquence trimestrielle). On utilise le critère ***pmax*** (ou ***kmax***).

1. Expliquez les différences du test de Dickey-Fuller Augmenté par rapport au Dickey-Fuller standard. Posez la régression réalisée dans le cadre du test de Dickey-Fuller Augmenté si la série *M3* suit un AR2.
2. Expliquez **clairement** le principe du critère *pmax.*
3. Expliquez **clairement** le principe du test de Ljung-Box (objectif, hypothèses, statistique calculée, loi).

Le tableau qui suit reproduit les résultats obtenus. Apparaissent, dans l'ordre, le nombre de retard, le SL associé au test de Ljung-Box, le SL associé au test de student du dernier M3 retardé, et la statistique calculée de DF.

4. Quel est le nombre de retard retenu pour tester la stationnarité de la variable M3 dans le cadre du test ADF (*dans les questions 4 et 5, vous considérez les cas avec et sans trend*) ?

5. La masse monétaire est-elle stationnaire à 5% ?

Critère Pmax, variable M3

Avec constante				Avec constante et trend			
lag	**SL(Qstat)**	**SL(t_lag)**	**StatADF**	**lag**	**SL(Qstat)**	**SL(t_lag)**	**StatADF**
12	0.81993	0.20398	1.92446	12	0.81579	0.21241	0.87886
11	0.44524	0.33500	1.64869	11	0.45435	0.36084	0.90213
10	0.61509	0.07044	1.47757	10	0.62589	0.09123	0.98871
9	0.65896	0.19548	1.15702	9	0.66486	0.25193	1.16431
8	0.69968	0.46649	0.98069	8	0.69253	0.40418	1.27215
7	0.69824	0.97446	1.10861	7	0.68610	0.93556	1.24601
6	0.69593	0.59093	1.12702	6	0.69143	0.60654	1.25302
5	0.68742	0.34794	1.07427	5	0.68605	0.37039	1.24183
4	0.56154	0.00056	0.98576	4	0.55041	0.00054	1.23908
3	0.34256	0.00145	1.13675	3	0.54678	0.00056	1.27865
2	0.24577	0.00675	1.27658	2	0.49278	0.00034	1.42786
1	0.25677	0.00098	1.12786	1	0.44526	0.00027	1.56754
0	0.20986	0.00087	1.09876	0	0.27897	0.00100	1.21080

NB : Les séances 9 et 10 auront lieu en salle informatique. On programmera sous SAS, le critère Pmax.

Master 1 ESA Année Universitaire 2005-2006

Contrôle – Séries temporelles

Exercice1 :

1- La fonction d'autocorrélation totale d'une série temporelle est donnée comme suit :

k	1	2	3	4	5	6
$\hat{\rho}_k$	-.36457	XXXXX	XXXXX	XXXXX	XXXXX	-.13273
$\hat{\sigma}_{\hat{\rho}_k}$	0.089443	0.100631	0.114326	0.116154	0.116156	0.116514

k	7	8	9	10	11	12
$\hat{\rho}_k$	0.12924	-.16093	0.22162	-.22858	0.15280	-.16012
$\hat{\sigma}_{\hat{\rho}_k}$	0.117717	0.118847	0.120578	0.123794	0.127125	0.128586

a- Déterminez un intervalle de confiance à 95% pour ρ_k, avec k=2,3,4 et 5.

b- On vous précise qu'il s'agit d'un processus MA(2). Quelles sont les autocorrélations totales (manquantes) qui devraient se situer hors de l'intervalle de confiance associé. Justifiez votre réponse.

2- On considère deux processus indépendants : $x_t \to MA(2)$ et $y_t \to MA(1)$ (les bruits blancs respectifs, notés u_t et v_t sont orthogonaux).

a- Quelle est la mémoire de x_t et y_t ?

b- On définit $z_t = x_t + y_t$. Quelle est la mémoire de z_t ? (calculez les corrélations $(z_t, z_{t-1})...(z_t, z_{t-j})$)

c- En déduire le processus suivi par z_t.

Exercice2 : Sur une série temporelle de longueur 125, quatre estimations ont été réalisées, avec les résultats suivants (Entre parenthèses, les écart-types des paramètres estimés. Q(.) est la statistique de Ljung-Box pour les résidus issus des estimations)

$$\text{(a)}: \Delta y_t = \underset{(0.35273)}{2.26492} - \underset{(0.00379)}{0.01683} y_{t-1} \qquad Q(24) = 232.99$$

$$\text{(b)}: \Delta y_t = \underset{(0.27333)}{0.83774} - \underset{(0.00274)}{0.00708} y_{t-1} + \underset{(0.05989)}{0.72513} \Delta y_{t-1} \qquad Q(24) = 29.17$$

$$\text{(c)}: \Delta y_t = \underset{(0.28549)}{0.86443} - \underset{(0.00283)}{0.00721} y_{t-1} + \underset{(0.09126)}{0.77765} \Delta y_{t-1} - \underset{(0.08982)}{0.06987} y_{t-2} \qquad Q(24) = 24.62$$

1- A quel type de problème, ces différentes estimations apportent-elles une réponse ? Commentez-les résultats de l'estimation (a) et concluez (si possible) sur la nature du processus suivi par y_t.

2- On peut opposer les deux dernières régressions à la première. Quel est le test qui leur est associées. Rappelez le principe du test (Equations, hypothèses testées, statistique et seuil théorique).

3- Commentez les résultats des régressions (b) et (c) et conclure quant à la nature du processus suivi par y_t.

Exercice3 : Pour modéliser le logarithme de la différence première de l'indice de la production industrielle observé sur une période de 122 points, Enders (Applied Econometrics Time Series, 1995, pp. 109) sélectionne quatre représentations possibles. Les résultats sont donnés dans le tableau qui suit :

	p=1 q=0	**p=2 q=0**	**p=1 q=1**	**p=1 q=1, 4**	**p=1 q=2**
ϕ_0	0.011 (4.14)	0.011 (3.31)	0.012 (2.63)	0.011 (2.76)	0.012 (2.62)
ϕ_1	0.618 (8.54)	0.456 (5.11)	0.887 (14.9)	0.791 (9.21)	0.887 (13.2)
ϕ_2		0.258 (2.89)			
θ_1			-0.484 (-4.22)	-0.409 (-3.62)	-0.483 (-4.19)
θ_2					-0.002 (-0.019)
θ_4				0.315 (3.36)	
SSR	0.0156	0.0145	0.0141	0.0134	0.0141
AIC	-503.3	-506.1	-513.1	-518.2	-511.1
SBC	-497.7	-497.7	-504.7	-507	-499.9
Q(12)	23.6 **(0.008)**	11.7 **(0.302)**	11.7 **(0.301)**	4.8 **(0.898)**	11.7 **(0.301)**
Q(24)	28.6 **(0.157)**	15.6 **(0.833)**	15.4 **(0.842)**	9.3 **(0.991)**	15.3 **(0.841)**
Q(30)	40.1 **(0.082)**	22.8 **(0.742)**	22.7 **(0.749)**	14.8 **(0.972)**	22.6 **(0.749)**

(.) T-stat pour la nullité des coefficients

Q(.) Stat de Ljung-Box pour l'analyse de l'autocorrélation dans la série des résidus

(.) P-value correspondant à la Stat de Ljung-Box.

Quel modèle allez-vous retenir (expliquez votre démarche, en particulier quels tests, quels degrés de liberté, quels seuils de risque,…) ?

CORRIGE CONTROLE SERIE TEMPORELLE 1

EXERCICE 1 :

1-

a- Intervalle de confiance à 95% pour $\hat{\rho}_k$, avec k=2,3,4 et 5.

Loi asymptotique de $\hat{\rho}_k$: $\hat{\rho}_k \sim N(0, \hat{\sigma}_{\hat{\rho}_k}{}^2)$, avec $\hat{\sigma}_{\hat{\rho}_k}{}^2$ donnée par la formule de Bartlett. On en déduit successivement :

$$\frac{\hat{\rho}_k}{\hat{\sigma}_{\hat{\rho}_k}} \sim N(0,1)$$

$$1-\alpha = \Pr ob[Z_{\alpha/2} \leq \frac{\hat{\rho}_k}{\hat{\sigma}_{\hat{\rho}_k}} \leq Z_{1-\alpha/2}]$$

$$1-\alpha = \Pr ob[Z_{\alpha/2}\hat{\sigma}_{\hat{\rho}_k} \leq \hat{\rho}_k \leq Z_{1-\alpha/2}\hat{\sigma}_{\hat{\rho}_k}] \text{ avec } Z_{1-\alpha/2} = -Z_{\alpha/2} = 1.96 \approx 2 \text{ pour } \alpha = 5\%$$

On en déduit l'IC pour $\hat{\rho}_k$: $IC_{95\%}(\hat{\rho}_k) = \left[-2\hat{\sigma}_{\hat{\rho}_k};2\hat{\sigma}_{\hat{\rho}_k}\right]$

Les valeurs de $\hat{\sigma}_{\hat{\rho}_k}$ étant données, on trouve directement les IC pour les valeurs de k données :

$$IC_{95\%}(\hat{\rho}_2) = \left[-2\hat{\sigma}_{\hat{\rho}_2};2\hat{\sigma}_{\hat{\rho}_2}\right] = [-0.201262;0.201262]$$

$$IC_{95\%}(\hat{\rho}_3) = \left[-2\hat{\sigma}_{\hat{\rho}_3};2\hat{\sigma}_{\hat{\rho}_3}\right] = [-0.228652;0.228652]$$

$$IC_{95\%}(\hat{\rho}_4) = \left[-2\hat{\sigma}_{\hat{\rho}_4};2\hat{\sigma}_{\hat{\rho}_4}\right] = [-0.232308;0.232308]$$

$$IC_{95\%}(\hat{\rho}_5) = \left[-2\hat{\sigma}_{\hat{\rho}_5};2\hat{\sigma}_{\hat{\rho}_5}\right] = [-0.232312;0.232312]$$

b- tester la nullité de ρ_k revient à vérifier si $\hat{\rho}_k$ se situe ou non à l'intérieur des intervalles de confiances ci-dessus calculés. Si le processus est un MA(2) les seules autocorrélations non nulles sont ρ_1 et ρ_2. Par conséquent $\hat{\rho}_1$ (resp. $\hat{\rho}_2$) doit se situer hors de l'intervalle $IC_{95\%}(\hat{\rho}_1)$ (resp. $IC_{95\%}(\hat{\rho}_2)$). Les valeurs de $\hat{\rho}_k$ pour k différents de 1 et 2 doivent prendre des valeurs à l'intérieure de l'IC respectif associé.

2-

a- Mémoire de x_t et y_t

Par définition, un processus MA pur d'ordre q à une mémoire égale à q (l'ordre de la dernière autocorrélation non nulle) On en déduit que la mémoire de x_t est égale à 2 et celle de y_t égale à 1. (point besoin ici de faire des calculs).

b- Mémoire de $z_t = x_t + y_t$

Calcul de γ_k pour z_t. $x_t \to MA(2) \Rightarrow x_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}$ $\quad y_t \to MA(1) \Rightarrow y_t = v_t + \alpha_1 v_{t-1}$

$$z_t = x_t + y_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + v_t + \alpha_1 v_{t-1}$$

$$\gamma_k = E(z_t z_{t-k}) \text{ car } E(z_t) = 0 \text{ parce que } E(u_t) = 0 \text{ et } E(v_t) = 0$$

$$\gamma_k = E[(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + v_t + \alpha_1 v_{t-1})(u_{t-k} + \theta_1 u_{t-k-1} + \theta_2 u_{t-k-2} + v_{t-k} + \alpha_1 v_{t-k-1})]$$

$$\gamma_k = E[(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + v_t + \alpha_1 v_{t-1})(u_{t-k} + \theta_1 u_{t-k-1} + \theta_2 u_{t-k-2} + v_{t-k} + \alpha_1 v_{t-k-1})] \quad (2)$$

Il est aisé de remarquer que pour k=1,

$$\gamma_1 = E\left[\theta_1 u_{t-1}{}^2 + \theta_1 \theta_2 u_{t-2}{}^2 + \alpha_1 v_{t-1}{}^2\right]$$

car les bruits sont orthogonaux $E(u_t v_{t-k}) = E(u_{t-k} v_t) = 0$

$$\gamma_1 = E\left[\theta_1 u_{t-1}{}^2 + \theta_1 \theta_2 u_{t-2}{}^2 + \alpha_1 v_{t-1}{}^2\right]$$

$$\gamma_1 = \theta_1 (1 + \theta_2) \sigma_u{}^2 + \alpha_1 \sigma_v{}^2$$

pour k=2 :

$$\gamma_2 = E\left[\theta_2 u_{t-2}{}^2\right]$$

$$\gamma_2 = \theta_2 \sigma_u{}^2$$

pour k>2, on remarque qu'il n'y a aucun terme commun dans les deux facteurs de l'égalité 2, d'où γ_k est nulle pour k>2. Les seules autocorrélations non nulles sont donc ρ_1 et ρ_2. $z_t = x_t + y_t$ est alors de mémoire 2.

NB : les résultats sont identiques à un signe près, si vous écrivez les processus comme suit :

$x_t \to MA(2) \Rightarrow x_t = u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2}$ et $y_t \to MA(1) \Rightarrow y_t = v_t - \alpha_1 v_{t-1}$

c- Processus suivi par $z_t = x_t + y_t$. Il s'agit d'un processus MA(2).

EXERCICE 2 :

1- Telles que présentées, les trois estimations ont des spécifications qui sont celles utiles pour effectuer des tests de stationnarité de type DF -(a)- ou ADF -(b) et (c).

Commentaire du (a) :

C'est une estimation autorégressive d'ordre 1 en différence, afférente au test DF. Toute conclusion quant à la stationnarité ou non du processus y_t doit passer avant tout par l'analyse des résidus. Si c'est un bruit blanc, on peut conclure quant à la stationnarité ou non. Dans le cas contraire aucune conclusion robuste ne peut être faite. La statistique de Ljung-Box pour le jeu d'hypothèse :

$$H0 : \rho_1 = \rho_2 = ... = \rho_{24} = 0 \text{ contre } H1 : \text{il existe au moins un } \rho_i \neq 0$$

conduit à la Stat Q(24)=232.99 qui sous H0 suit un $\chi^2(24-2) = \chi^2(22)$. On a :

$Q(24)=232.99>\chi_{95\%}{}^{2}(22)=33.92$, on rejette H0, donc la série de résidus n'est pas un bruit blanc. Par conséquent, on peut pas conclure quant à la stationnarité ou non de y_t.

2- Il s'agit du test ADF. Principe (Voir TD).

3- Commentaires des résultats de la régression (b) et (c).

Régression (b) :

Test de Ljung-Box sur la série des résidus :

$$H0:\rho_1=\rho_2=...=\rho_{24}=0 \text{ contre } H1: \text{il existe au moins un } \rho_i\neq 0$$

$Q(24)=29.17<\chi_{95\%}{}^{2}(21)=32.67$, on accepte H0, donc la série de résidus est un bruit blanc. On peut conclure quant à la stationnarité ou non de y_t.

Test de Stationnarité :

$$H0:\rho=0 \text{ contre } H1:\rho<0$$

$t_{\hat{\rho}}=\dfrac{-0.00708}{0.00274}=-2.5839>\hat{C}(5\%,125)=-2.8845\Rightarrow$ on accepte H0, la série n'est pas stationnaire.

La significativité des autres coefficients se mesure grâce à la stat traditionnelle de Student. Ainsi, la constante et le coefficient du terme autorégressive d'ordre 1 sont statistiquement différents de zéro (les calculs sont évidents !!!)

Régression (c) :

Test de Ljung-Box sur la série des résidus :

$$H0:\rho_1=\rho_2=...=\rho_{24}=0 \text{ contre } H1: \text{il existe au moins un } \rho_i\neq 0$$

$Q(24)=24.62<\chi_{95\%}{}^{2}(20)=31.41$, on accepte H0, donc la série de résidus est un bruit blanc. On peut conclure quant à la stationnarité ou non de y_t.

Test de Stationnarité :

$$H0:\rho=0 \text{ contre } H1:\rho<0$$

$t_{\hat{\rho}}=\dfrac{-0.00721}{0.00283}=-2.5477>\hat{C}(5\%,125)=-2.8845\Rightarrow$ on accepte H0, la série n'est pas stationnaire.

Quant aux autres coefficients, la constante et le coefficient du terme autorégressif d'ordre 1 sont statistiquement différents de zéro, alors que le coefficient du terme autorégressive d'ordre 2 est non significatif.

Nature du processus suivi par y_t : Tout ce qu'on peut dire, c'est que c'est un processus intégré au moins d'ordre 1. Si le test de stationnarité sur la série en différence conduit à la stationnarité de cette dernière, il s'agirait alors d'un ARI(2) , c'est-à-dire processus intégré d'ordre 1 avec erreur autorégressive d'ordre 2.

EXERCICE 3:

Démarche à suivre pour sélectionner le meilleur modèle :

Etape 1 : Analyse des résidus ; la série des résidus de chaque modèle est-elle un bruit blanc ? On écarte les modèles pour lesquels cette propriété n'est pas valable. Aucun calcul à faire, car les p-value correspondant au test de Ljung-Box sont données. Au risque de 5% , on écarte le modèle 1 car la p-value correspondant au test :

$$H0 : \rho_1 = \rho_2 = ... = \rho_{12} = 0 \text{ contre } H1 : \text{il existe au moins un } \rho_i \neq 0$$

est égale à 0.8%<5%. La série des résidus de ce modèle n'est donc pas un bruit blanc.

Etape 2 : Significativité des coefficients estimés.

On construit la statistique $t_{\hat{\phi}_i} = \left|\frac{\hat{\phi}_i}{\hat{\sigma}_{\hat{\phi}_i}}\right|$ ou $t_{\hat{\theta}_i} = \left|\frac{\hat{\theta}_i}{\hat{\sigma}_{\hat{\theta}_i}}\right|$ qu'on compare (si on utilise l'approximation de la Student par la normale, vu le nombre de points) à 1.96 ou 2.

Si $t_{\hat{\phi}_i} = \left|\frac{\hat{\phi}_i}{\hat{\sigma}_{\hat{\phi}_i}}\right| > 2$ le coefficient ϕ_i est significatif. Dans le cas contraire, il ne l'est pas.

Le modèle 5 est écarté à cette étape, car $t_{\hat{\theta}_2} = \left|\frac{\hat{\theta}_2}{\hat{\sigma}_{\hat{\theta}_2}}\right| < 2$. Il'y a donc une variable superflue dans ce modèle (le terme moyenne mobile d'ordre 2). Pour les autres modèles tous les coefficients sont significatifs (les calculs sont faciles à faire !!!). On a à la fin de cette étape, trois modèles concurrents : modèles 2, 3 et 4.

Etape 2 : utilisation critère AIC, BIC pour les discriminer.

AIC(modèle 4)=-518.2< AIC(modèle 3)=-513.1< AIC(modèle 2)=-506.1

BIC(modèle 4)=-507< AIC(modèle 3)=-504.7< AIC(modèle 2)=-497.7

Aussi bien pour le critère AIC, que pour le critère BIC, le meilleur modèle est le modèle 4

Conclusion : le modèle (parmi la classe des modèles retenus) qui apparaît le mieux approprié pour filtrer la série est le modèle 4. Il y a donc un effet saisonnier dans la série, capté par la variable moyenne mobile d'ordre 4.