

Analyse des séries temporelles

Applications à l'économie et à la gestion

Manuel et exercices corrigés

Régis Bourbonnais
Michel Terraza



3^e édition

DUNOD

Analyse des séries temporelles

Applications à l'économie et à la gestion

Régis Bourbonnais
Michel Terraza

3^e édition

DUNOD

Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.

Le Code de la propriété intellectuelle du 1^{er} juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements

d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour

les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée.

Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du

Centre français d'exploitation du droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).



© Dunod, Paris, 2010

ISBN 978-2-10-056111-7

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

Table des matières

Introduction	1
Partie I L'analyse classique des séries chronologiques	5
1. L'analyse de la saisonnalité	9
I. La détection de la saisonnalité	9
A. La représentation graphique et le tableau de Buys-Ballot	9
B. Analyse de la variance et test de Fisher	11
C. La fonction d'autocorrélation	15
D. Le spectre	19
II. La sélection du schéma	20
A. La procédure de la bande	20
B. Le test de Buys-Ballot	21
III. Les méthodes de désaisonnalisation	22
A. Le principe de la conservation des aires	22
B. Cas d'une saisonnalité rigide	23
C. Cas d'une saisonnalité souple	39
2. Prévision d'une série chronologique	45
I. Prévision d'une chronique non saisonnière	45
A. Tests de détection d'une tendance	46
B. Analyse par régression	48
C. Le lissage exponentiel	49
II. Prévision d'une chronique saisonnière	67
A. Analyse par régression	68
B. Utilisation des coefficients saisonniers	68
C. Prévision par lissage exponentiel de Holt-Winters	73

Partie II	Traitement des séries temporelles, réalisations de processus aléatoires	79
3.	Processus aléatoires stationnaires et processus ARMA	83
I.	Définition d'un processus stochastique	83
II.	Les processus stationnaires	84
	A. Définition d'un processus stationnaire au sens strict : la stationnarité forte	84
	B. La stationnarité d'ordre deux des processus : la stationnarité faible	85
	C. Le processus Bruit Blanc (<i>White Noise</i>)	85
	D. L'ergodicité	86
III.	La fonction d'autocorrélation et la fonction d'autocorrélation partielle	87
	A. La fonction d'autocorrélation	87
	B. La fonction d'autocorrélation partielle	89
	C. Analyse des fonctions d'autocorrélation	91
IV.	La classe des processus aléatoires ARMA linéaires et stationnaires	96
	A. Le théorème de décomposition de Wold	96
	B. Propriétés de l'opérateur retard	98
	C. Définition des processus ARMA	99
	D. La stationnarité et l'inversibilité des processus	102
	E. Les processus ARMA saisonniers	107
	F. Les processus ARMA non saisonniers et saisonniers à la fois	108
4.	Les processus aléatoires dans le domaine des fréquences	117
I.	Filtrage linéaire d'un processus aléatoire	117
	A. Définitions	117
	B. La fonction de réponse impulsionnelle et la fonction de réponse en fréquence du filtre	118
	C. Fonction de transfert, fonction de gain et fonction de phase du filtre	121
	D. Exemples de filtres linéaires	123
II.	Le spectre d'un processus aléatoire	132
	A. Les théorèmes de représentation	132
	B. Le spectre d'une série temporelle filtrée	134
	C. Le spectre d'une chronique ou l'estimateur spectral	135
	D. La lecture d'un spectre	139
	E. Le spectre d'un processus ARMA	143

5. Les processus aléatoires non stationnaires	153
I. Description des processus TS et DS	154
A. Les processus TS	154
B. Les processus DS	154
C. Conséquences d'une mauvaise stationnarisation du processus	157
II. Tests de racines unitaires non saisonnières	162
A. Les tests de Dickey-Fuller simples	163
B. Les tests de Dickey et Fuller augmentés	169
C. Le test de Phillips et Perron	178
D. Le test de Dickey et Pantula	179
E. Le test KPSS	180
F. Le test de Elliot, Rothenberg et Stock (1996)	181
G. Le test Ng-Perron (2001)	182
III. Tests de racines unitaires saisonnières	191
A. Les modèles de base	192
B. Le test de Hylleberg, Engle, Granger et Yoo (HEGY)	193
C. Le test de Franses	194
IV. Les processus ARIMA	198
A. Les processus ARIMA non saisonniers	198
B. Les processus ARIMA purement saisonniers (modèles SARIMA)	199
C. Les processus ARIMA non saisonniers et saisonniers à la fois	200
6. L'identification des processus ARMA	205
II. Les caractéristiques des processus $AR(p)$	206
A. Caractéristiques de la FAC d'un $AR(p)$	206
B. Caractéristique de la FAP d'un $AR(p)$	207
C. Exemple d'application	209
II. Les caractéristiques des processus $MA(q)$	211
A. Caractéristiques de la FAC d'un $MA(q)$	212
B. Caractéristiques de la FAP d'un $MA(q)$	213
III. Les caractéristiques des processus $ARMA(p, q)$	215
A. Caractéristiques de la FAC d'un $ARMA(p, q)$	216
B. Caractéristiques de la FAP d'un $ARMA(p, q)$	217
C. Synthèse	219
IV. Simulations et exercices	220
A. Limite à l'utilisation des fonctions d'autocorrélation	220
B. Exercices	221
V. La pratique de l'identification des processus	227
A. La fonction d'autocorrélation inverse et la fonction d'autocorrélation partielle inverse	227

B. La fonction d'autocorrélation étendue	231
C. Les autres méthodes d'identification	237
7. L'estimation, les tests de validation et la prévision des processus ARMA	239
I. Le problème de l'estimation	239
A. Détermination et estimation de la vraisemblance des processus ARMA	240
B. Les méthodes d'estimation	241
II. Les tests de validation	243
A. L'analyse des racines	243
B. Le test de Student des paramètres	243
C. Le coefficient de détermination	244
D. Les tests sur les résidus de bruit blanc normal	244
E. Les critères de comparaison de modèles	252
III. La prévision	255
A. Les transformations de départ	255
B. Calcul du prédicteur	256
IV. Synthèse de la procédure et exercices d'application	260
8. Processus à mémoires longues et processus non linéaires	281
I. Processus ARFIMA et processus chaotiques	281
A. Les processus ARFIMA	282
B. Les processus chaotiques	293
II. Les modèles ARCH : présentation générale	299
A. Modèle de régression de type ARCH	301
B. Test d'un modèle de type ARCH	304
C. Procédure d'estimation et prévision	305
D. Processus de type GARCH	310
E. Autres processus : variantes des processus ARCH	314
Liste des exercices	321
Tables statistiques	325
Bibliographie	329
Index	337

Introduction

En attribuant conjointement, en 2003, le prix Nobel d'économie à l'Américain Robert F. Engle et au Britannique Clive W. J. Granger, l'Académie des sciences de Suède a récompensé deux économistes dont les travaux sur les méthodes économétriques des séries temporelles ont contribué à améliorer la fiabilité des prévisions économiques et financières.

L'Académie, dans un communiqué, écrit : « Les lauréats de cette année ont développé au cours des années 1980 de nouvelles méthodes statistiques qui portent sur deux propriétés caractéristiques de beaucoup de séries temporelles : la volatilité saisonnière et la non-stationnarité. »

Plus précisément, c'est la *volatilité dynamique* qui a valu son prix à Robert F. Engle, pour désigner des fluctuations susceptibles de varier fortement dans le temps. Il a mis en évidence le concept de modèle à erreurs ARCH (modèle autorégressif à hétéroscédasticité conditionnelle) particulièrement adapté à l'étude de la volatilité des prix des actifs sur les marchés financiers, avec des périodes calmes et des variations de faible amplitude suivies par des épisodes agités et d'importantes fluctuations.

Parallèlement, la distinction du Britannique Clive Granger est venue couronner ses travaux sur les méthodes d'analyses temporelles économiques avec une tendance commune : la co-intégration. C.W.J Granger a montré que l'étude de données non stationnaires avec des méthodes classiques d'analyse des séries temporelles stationnaires peut donner des régressions fallacieuses (*spurious regression*) et donc des conclusions erronées.

En récompensant ces deux économistes, l'Académie royale des sciences de Suède a souligné l'importance croissante de l'analyse statistique pour la modélisation des phénomènes économiques, mais aussi l'indispensable amélioration des instruments d'analyse et de prévision dans un environnement économique de plus en plus incertain.

Ce manuel, s'inspirant des travaux de ces auteurs, traite, en effet, de l'économétrie des séries temporelles et des méthodes de la prévision à court terme réalisée à partir de séries chronologiques. Ces prévisions peuvent être élaborées à partir d'une seule série chronologique, c'est le cas univarié ou bien être le résultat de plusieurs séries chronologiques que l'on met en relation, c'est le cas multivarié.

Cette troisième édition est enrichie par des nouveaux exercices et intègre les développements les plus récents de l'analyse univariée des séries temporelles.

Nous ne considérons dans cet ouvrage que le cas univarié et deux grandes parties sont abordées :

La partie I traite des méthodes standards de traitement des séries temporelles (méthodes de désaisonnalisation et lissage exponentiel).

Dans la partie II la chronique est considérée comme un échantillon d'un processus aléatoire univarié ; dans cette partie sont présentés les modèles ARMA stationnaires, l'analyse spectrale, puis les problèmes de la non stationnarité des chroniques (tests de racine unitaire). L'estimation et la validation des processus ARIMA, qui fondent l'algorithme de Box et Jenkins, sont ensuite abordées. Enfin, les modèles non linéaires (ARFIMA, ARCH...) sont traités dans un dernier chapitre.

Ce livre fait appel à des notions d'économétrie¹ et de statistiques d'un niveau de maîtrise en sciences économiques. Les exposés théoriques sont illustrés par des exemples et des exercices qui permettent au lecteur de se familiariser de manière concrète à la pratique du traitement des séries chronologiques.

Nous avons voulu, par une alternance systématique de cours et d'exercices, répondre à un besoin pédagogique qui est de mettre rapidement en pratique les connaissances théoriques et ainsi, d'utiliser de manière opérationnelle les acquis du cours. De surcroît, le recours à des logiciels, lors de la résolution des exercices, permet une découverte de ces outils et donne une dimension pratique que recherchent l'étudiant et le praticien. Ce manuel constitue un livre d'apprentissage facilement accessible ; c'est pourquoi les démonstrations les plus complexes font l'objet de renvois à une bibliographie plus spécialisée.

Afin que le lecteur puisse lui-même refaire les exercices, les données utilisées (sous format Excel et Eviews) ainsi que les programmes de traitement « *Batch* » de Eviews sont disponibles par téléchargement sur le serveur web :

<http://regisbourbonnais.dauphine.fr>

Dans ce livre, nous voulons donner aux lecteurs (étudiants, chercheurs, économistes d'entreprise) tous les éléments tant théoriques (sans démonstration et

1. Nous recommandons au lecteur souhaitant se familiariser avec l'économétrie et désirant faire quelques exercices d'applications la lecture de Bourbonnais R., *Économétrie : cours et exercices corrigés*, Dunod, 7^e éd., 2009.

exposé superflus) que pratiques lui permettant de résoudre les problèmes auxquels il est confronté dans la manipulation des séries temporelles. Ce livre s'adresse en premier lieu aux étudiants (sciences économiques, gestion, écoles de commerce et d'ingénieurs...) dont la formation requiert une connaissance dans le domaine de l'économétrie des séries temporelles. Gageons qu'il sera un support de cours indispensable et un allié précieux pour préparer les séances de travaux dirigés.

N'oublions pas cependant le praticien de l'économétrie (économiste d'entreprise, chercheurs...) qui, confronté à des problèmes d'estimation statistique, trouvera dans ce livre les réponses pratiques aux différentes questions qu'il peut se poser.

Enfin, nous exprimons toute notre gratitude à toutes les personnes – collègues et étudiants – qui ont eu la gentillesse de nous faire des commentaires et dont les conseils et les suggestions contribuent à la qualité pédagogique de ce livre. Nous restons, bien entendu, les seuls responsables des erreurs qui subsisteraient¹.

1. Les lecteurs souhaitant faire des remarques et des commentaires peuvent nous contacter.
Régis Bourbonnais : regis.bourbonnais@dauphine.fr
Michel Terraza : terraza@lameton.univ-montp1.fr

Partie I

L'analyse classique des séries chronologiques

Une série temporelle ou encore chronique est une succession d'observations au cours du temps représentant un phénomène économique (prix, ventes...) ; par hypothèse, le pas du temps des observations est considéré constant : l'heure, le jour, le mois, le trimestre, l'année. Nous supposons également que la chronique ne contient ni observations manquantes, ni valeurs aberrantes ou accidentelles sur toute la période d'observation. Si tel est le cas une correction des données est réalisée en utilisant la méthode de l'extrapolation linéaire, la prévision de la (ou des) valeurs en cause ou encore l'intuition du modélisateur. La valeur courante en t de la chronique est notée x_t , où t le temps est compris entre 1 et n avec n le nombre total d'observations de la chronique. On appelle h le nombre de points ou de valeurs à prévoir de la chronique. La prévision de la série temporelle – de $(n + 1)$ à $(n + h)$ connaissant l'historique de x_1 à x_n – porte le nom d'horizon de la prévision.

Il est toujours utile, en première analyse, de représenter l'évolution temporelle d'un phénomène à l'aide d'un graphique ayant en ordonnée la valeur du phénomène économique x_t et en abscisse le temps t . Comme le temps est discret, le graphique obtenu est un diagramme en bâtons. Par tradition, on retient le polygone des fréquences de la représentation nommé : profil temporel de la chronique.

Les techniques traditionnelles de traitement des chroniques procèdent par décomposition puis reconstitution de la chronique pour effectuer la prévision. Cette approche suppose que la structure de la chronique peut être décomposée en éléments simples (modélisables), et donc plus facilement prévisibles, pour ensuite être reconstituée pour donner la prévision de la chronique.

Les premières études sur les chroniques ont amené à considérer de façon standard quatre grandes composantes :

- la tendance ou « *trend* » notée T_t , censée décrire le mouvement de long terme, de fond ou encore structurel du phénomène. Ce mouvement est traditionnellement représenté par des formes analytiques simples : polynomiales, logarithmiques, exponentielles, cycliques, logistiques. C'est ainsi qu'en économie la tendance contient des cycles longs de Kondratieff (cycle aperiodique de 40 à 60 ans), de Kuznets (20 ans), des cycles de Juglar (cycle de crise de 10 ans)... ;
- la composante cyclique notée C_t . En conjoncture, elle est représentée par le cycle de Kitchin d'une période de 3 à 5 ans. Dans la plupart des travaux sur les séries temporelles, la tendance et le cycle sont regroupés en une seule composante appelée l'extra-saisonnier E_t ;
- la composante saisonnière notée S_t : composante cyclique relativement régulière de période intra-annuelle et qui correspond souvent à des phénomènes de mode, de coutume, de climat... ;
- la composante résiduelle notée R_t . Elle rassemble tout ce que les autres composantes n'ont pu expliquer du phénomène observé. Elle contient donc de nombreuses fluctuations, en particulier accidentelles, dont le caractère est exceptionnel et imprévisible (catastrophes naturelles, grèves, guerres...). Comme par hypothèse ce type d'événement est censé être corrigé, le résidu présente – en général – une allure aléatoire plus ou moins stable autour de sa moyenne.

Remarquons que ces différentes composantes s'entendent pour des séries économiques, le plus souvent, mensuelles ou trimestrielles liées à la conjoncture. Dans le domaine de l'entreprise, les composantes sont conservées mais les périodicités sont parfois différentes (hebdomadaire par exemple).

La technique de décomposition- reconstitution repose, bien évidemment, sur un modèle qui l'autorise. Ce modèle porte le nom de schéma de décomposition. Il en existe essentiellement trois grands types :

- le schéma additif qui suppose l'orthogonalité (indépendance) des différentes composantes. Il s'écrit : $x_t = E_t + S_t + R_t$. Dans ce schéma la saisonnalité est rigide en amplitude et en période ;
- le schéma multiplicatif : $x_t = E_t \times S_t + R_t$, dans lequel la composante saisonnière est liée à l'extra-saisonnier (saisonnalité souple avec variation de l'amplitude au cours du temps) ;

- le schéma multiplicatif complet : $x_t = E_t \times S_t \times R_t$ (interaction générale des trois composantes). Il est actuellement le plus utilisé en économie. Il est commode puisque le logarithme de la chronique conduit au schéma additif.

En définitive, dans ces méthodes traditionnelles, deux questions sont importantes : l'existence d'une saisonnalité et le type de schéma à retenir ; elles constituent le chapitre 1 de cette partie. Nous examinons par la suite (chapitre 2) les techniques de prévision.

1. L'analyse de la saisonnalité

L'étude de la saisonnalité est un préalable au traitement d'une série chronologique. En effet, lorsque cette composante existe, il convient de l'isoler afin de pouvoir analyser les autres caractéristiques. Une désaisonnalisation systématique, sans tester l'existence de cette composante, peut créer un « bruit » parasite nuisible à l'analyse de la chronique et donc dégrader la qualité de la prévision. Dans ce chapitre, nous allons, par conséquent, présenter les techniques permettant de tester l'existence d'une composante saisonnière, puis nous examinons les méthodes de désaisonnalisation.

I. La détection de la saisonnalité

A. La représentation graphique et le tableau de Buys-Ballot

L'analyse graphique d'une chronique suffit, parfois, pour mettre en évidence une saisonnalité. Néanmoins, si cet examen n'est pas révélateur ou en cas de doute, le tableau de Buys-Ballot permet d'analyser plus finement l'historique. La figure 1.1 des ventes trimestrielles d'un produit festif indique une saisonnalité marquée au quatrième trimestre, ce que nous pouvons confirmer à l'aide du tableau de Buys-Ballot.

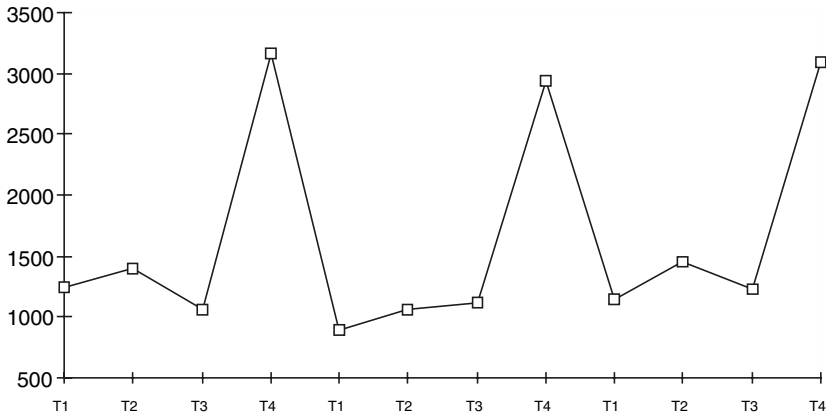


Figure 1.1 – Ventes trimestrielles en milliers d'unités

Le tableau de Buys-Ballot est un tableau complet à deux entrées dans lequel sont consignées les valeurs de x_t . Il est constitué (cf. Tableau 1.1) en ligne par les années et en colonne par le facteur à analyser (mois, trimestre...). Les moyennes et les écarts types¹ des années et des trimestres (ou des mois selon le cas) sont calculés ainsi que pour l'ensemble des observations de la chronique.

Tableau 1.1 – Exemple de constitution d'un tableau de Buys-Ballot pour des ventes trimestrielles (calculs arrondis)

Dates	T1	T2	T3	T4	Moyenne	Écart type
Année 1	1248	1392	1057	3159	1714	842,69
Année 2	891	1065	1118	2934	1502	831,02
Année 3	1138	1456	1224	3090	1727	795,48
Moyenne	1092	1304	1133	3061	Moyenne générale	Écart type général
Écart type	149	171	69	94	1647,7	829,74

Nous pouvons alors classer les trimestres pour chaque année par valeurs décroissantes (cf. Tableau 1.2).

Tableau 1.2 – Classement des trimestres en fonction de leur valeurs

Années				
Année 1	T4	T2	T1	T3
Année 2	T4	T3	T2	T1
Année 3	T4	T2	T3	T1

1. La formule de l'écart type utilisée est celle de la population.

La lecture du Tableau 1.2 indique la persistance du trimestre T4 à se classer en première position quelles que soient l'année et la position de « creux » occupée par le trimestre T1, ce qui nous conduit à retenir l'existence d'une saisonnalité rigide. Cette technique très simple permet la détection de la saisonnalité et aussi d'en préciser la nature.

B. Analyse de la variance et test de Fisher

L'examen visuel du graphique ou du tableau ne permet pas toujours de déterminer avec certitude l'existence d'une saisonnalité, de surcroît il interdit l'automatisme de traitement qui peut s'avérer nécessaire dans le cas d'un nombre important de séries à examiner. Le test de Fisher à partir de l'analyse de la variance permet de pallier ces deux inconvénients.

Ce test¹ suppose la chronique sans tendance ou encore sans extra-saisonnalité. Dans le cas contraire cette composante est éliminée par une régression sur le temps (extra-saisonnalité déterministe), ou par une procédure de filtrage (extra-saisonnalité aléatoire).

Soit :

N le nombre d'années,

p le nombre d'observations (la périodicité) dans l'année (trimestre $p = 4$, mois $p = 12$, etc.).

x_{ij} la valeur de la chronique pour la $i^{\text{ème}}$ année ($i = 1, \dots, N$) et la $j^{\text{ème}}$ période ($j = 1, \dots, p$) supposée telle que $x_{ij} = m_{ij} + e_{ij}$; les e_{ij} sont les résidus considérés comme aléatoires formés d'éléments indépendants : $e_{ij} \rightarrow N(0; \sigma^2)$.

Les m_{ij} sont les éléments d'une composante de la chronique qui s'écrivent : $m_{ij} = a_i + b_j$ avec b_j qui mesure l'effet période en colonne du tableau et a_i qui mesure l'effet année en ligne du tableau.

Deux effets absents sont testés contre deux effets significativement présents :

- si l'effet période (ici trimestre) est significatif, la série est saisonnière ;
- si l'effet année est significatif, ceci suggère deux interprétations.

- La chronique de départ n'a pas été transformée, elle possède alors des paliers horizontaux.
- La chronique a été transformée, des changements de tendance existent dans la chronique.

1. Laloire, 1972.

Le déroulement du test est le suivant :

1) Calcul de la variance totale du tableau

Soit S_T la somme totale des carrés :

$$S_T = \sum_{i=1}^N \sum_{j=1}^P (x_{ij} - x_{..})^2 \text{ avec}$$

$$x_{..} = \frac{1}{N \times P} \sum_{i=1}^N \sum_{j=1}^P x_{ij} : \text{moyenne générale de la chronique sur les } N \times p$$

observations.

$$x_{i.} = \frac{1}{P} \sum_{j=1}^P x_{ij} : \text{moyenne de l'année } i$$

$$x_{.j} = \frac{1}{N} \sum_{i=1}^N x_{ij} : \text{moyenne de la période } j$$

Comme $x_{ij} = m_{ij} + e_{ij}$ avec $e_{ij} \rightarrow N(0; \sigma^2)$ et $m_{ij} = \text{effet période} + \text{effet année}$, nous obtenons :

$$\begin{aligned} S_T &= \sum_i \sum_j (x_{ij} - x_{..})^2 \\ &= \sum_i \sum_j [(x_{i.} - x_{..}) + (x_{.j} - x_{..}) + (x_{ij} - x_{i.} - x_{.j} + x_{..})]^2 \\ &= \sum_i \sum_j (x_{i.} - x_{..})^2 + \sum_i \sum_j (x_{.j} - x_{..})^2 + \sum_i \sum_j (x_{ij} - x_{i.} - x_{.j} + x_{..})^2 \\ &\quad + (\text{terme rectangle nul}) \\ &= p \sum_i (x_{i.} - x_{..})^2 + N \sum_j (x_{.j} - x_{..})^2 + \sum_i \sum_j (x_{ij} - x_{i.} - x_{.j} + x_{..})^2 \end{aligned}$$

$$S_T = S_A \underset{\text{(année)}}{+} S_p \underset{\text{(période)}}{+} S_R \underset{\text{(résidus)}}{+}$$

Le tableau 1.3 présente les calculs intermédiaires avec les notations précédentes.

Tableau 1.3 – Calculs des moyennes par année et période

<i>Périodes</i>	<i>I</i>	...	<i>J</i>	...	<i>p</i>	<i>Moyennes années</i>
<i>Années</i>						
1	x_{11}		x_{1j}		x_{1p}	
...						
<i>i</i>	x_{i1}		x_{ij}		x_{ip}	$x_{i.} = \frac{1}{p} \sum_{j=1}^p x_{ij}$
...						
<i>N</i>	x_{N1}		x_{Nj}		x_{Np}	
<i>Moyennes périodes</i>			$x_{.j} = \frac{1}{N} \sum_{i=1}^N x_{ij}$			$x_{..} = \frac{1}{Np} \sum_{j=1}^p \sum_{i=1}^N x_{ij}$

Nous utilisons ces résultats pour effectuer l'analyse de la variance de la série (Tableau 1.4).

Tableau 1.4 – Analyse de la variance pour détecter une saisonnalité et/ou une tendance

Somme des carrés	Degré de liberté	Désignation	Variance
$S_p = N \sum_j (x_{.j} - x_{..})^2$	$p - 1$	Variance Période	$V_p = \frac{S_p}{p - 1}$
$S_A = P \sum_i (x_{i.} - x_{..})^2$	$N - 1$	Variance Année	$V_A = \frac{S_A}{N - 1}$
$S_R = \sum_i \sum_j (x_{ij} - x_{i.} - x_{.j} + x_{..})^2$	$(p - 1) \times (N - 1)$	Variance Résidu	$V_R = \frac{S_R}{(p - 1)(N - 1)}$
S_T	$N \times p - 1$	Variance Totale	$V_T = \frac{S_T}{N \times p - 1}$

À partir de ce tableau 1.4, nous pouvons construire les tests d'hypothèses.

2) Test de l'influence du facteur colonne, la période (mois ou trimestre : $H_0 = \text{pas d'influence}$)

Calcul du Fisher empirique $F_c = \frac{V_P}{V_R}$ que l'on compare au Fisher lu dans la table $F_{v_1;v_2}^\alpha$ à $v_1 = p - 1$ et $v_2 = (N - 1)(p - 1)$ degrés de liberté.

Si le Fisher empirique est supérieur au Fisher lu dans la table, on rejette l'hypothèse H_0 , la série est donc saisonnière.

3) Test de l'influence du facteur ligne, la tendance ($H_0 = \text{pas d'influence du facteur année}$)

Calcul du Fisher empirique $F_c = \frac{V_A}{V_R}$ que l'on compare au Fisher lu dans la table $F_{v_3;v_2}^\alpha$ à $v_3 = N - 1$ et $v_2 = (N - 1)(p - 1)$ degrés de liberté.

Si le Fisher empirique est supérieur au Fisher lu, on rejette l'hypothèse H_0 , la série est donc affectée d'une tendance.

Exercice n° 1.1

fichier C1EX1

Test de détection de la saisonnalité à partir des données du Tableau 1.1

On demande d'effectuer le test de détection de saisonnalité et de tendance à partir des données du Tableau 1.1.

Solution

Nous trouvons les résultats suivants ($p = 4, N = 3$) :

$$x_{..} = 1\,647,67 ;$$

$$x_{1.} = 1\,714 ; x_{2.} = 1\,502 ; x_{3.} = 1\,727$$

$$x_{.1} = 1\,092,33 ; x_{.2} = 1\,304,33 ; x_{.3} = 1\,133 ; x_{.4} = 3\,061$$

$$S_p = 8\,066\,015,31 ; S_A = 127\,650,68 ; S_R = 68\,026,67 ;$$

$$S_T = S_A + S_p + S_R = 8\,261\,692,66$$

a) Test de saisonnalité : $F_c = \frac{S_p/3}{S_R/6} = 237,14$ que l'on compare à $F_{3;6}^{0,05} = 4,76$. La série est donc saisonnière.

b) Test de tendance : $F_c = \frac{S_A/2}{S_R/6} = 5,63$ que l'on compare à $F_{2;6}^{0,05} = 5,14$, l'hypothèse H_0 est rejetée ; la chronique est affectée d'une tendance (à la limite de la significativité pour un seuil de 5 %). Dans ce cas il convient de refaire le test après élimination de la tendance, soit par régression sur le temps soit par un passage aux différences premières (cf. chapitre 5).

C. La fonction d'autocorrélation

On appelle coefficient d'autocorrélation d'ordre 1 le coefficient de corrélation linéaire simple calculé entre la série chronologique et cette même série décalée d'une période de temps. Les coefficients d'autocorrélation sont calculés pour des ordres allant de 0 à K , K étant le décalage maximum admissible. On choisit en général $\frac{n}{6} \leq K \leq \frac{n}{3}$ mais aussi K égal à 2 ou 3 fois la période de la saisonnalité, $K = \frac{n}{5}$ si $n \geq 150$ ou encore comme certains auteurs le suggèrent $K = \sqrt{n}$.

Tableau 1.5 – Exemple de calcul d'une corrélation

	Retards	Autocorrélation
1 2 3 ... t ... n	0	$r_0 = 1$
1 2 3 ... $t-1$... $n-1$	1	r_1
1 2 3 ... $t-2$... $n-2$	2	r_2
...
$t-k$... $n-k$	k	r_k
...
$t-K$... $n-K$	K	r_K

La représentation graphique de la fonction d'autocorrélation (notée FAC) est appelée corrélogramme.

Le coefficient d'autocorrélation d'ordre k est donné par :

$$r_k = \frac{\sum_{t=k+1}^n (x_t - \bar{x}_1)(x_{t-k} - \bar{x}_2)}{\sqrt{\sum_{t=k+1}^n (x_t - \bar{x}_1)^2 \sum_{t=k+1}^n (x_{t-k} - \bar{x}_2)^2}} =$$

$$\frac{\sum_{t=k+1}^n x_t x_{t-k} - (n-k)\bar{x}_1 \bar{x}_2}{\sqrt{\left(\sum_{t=k+1}^n x_t^2 - (n-k)\bar{x}_1^2\right)\left(\sum_{t=k+1}^n x_{t-k}^2 - (n-k)\bar{x}_2^2\right)}}$$

avec $\bar{x}_1 = \frac{1}{n-k} \sum_{t=k+1}^n x_t$ $\bar{x}_2 = \frac{1}{n-k} \sum_{t=k+1}^n x_{t-k}$

Le test de signification sur le coefficient r_k permet de sélectionner les coefficients d'autocorrélation significativement différents de 0 ; il peut, si n est suf-

fisamment grand, s'effectuer comme pour un coefficient de corrélation linéaire simple. Soit ρ_k la valeur vraie de r_k et l'hypothèse $H_0 : \rho_k = 0$ contre l'hypothèse alternative $H_1 : \rho_k \neq 0$.

Sous l'hypothèse H_0 la quantité $t_c = \frac{|r_k|}{\sqrt{1-r_k^2}} \sqrt{n-2}$ obéit à une loi de Student à $n-2$ degrés de liberté (ou à une loi normale centrée réduite si $n > 30$) dans lequel n est le nombre d'observations servant à calculer le coefficient d'autocorrélation. Si $t_c > t_{n-2}^{1-\alpha/2}$, l'hypothèse H_0 est rejetée ($t_{n-2}^{1-\alpha/2}$ valeur de la loi de Student au seuil α à $n-2$ degrés de liberté).

Pour n grand, si la série est sans tendance et de variance relativement constante au cours du temps, (cf. chapitre 6) nous pouvons poser $\bar{x}_1 = \bar{x}_2 = \bar{x}$ (\bar{x} = moyenne de la série chronologique), $\text{Var}(x_t) = \text{Var}(x_{t-k})$ et nous pouvons écrire la formule simplifiée suivante :

$$r_k = \frac{R_k}{R_0} \text{ avec } R_k = \frac{1}{n-k} \sum_{t=k+1}^n (x_t - \bar{x})(x_{t-k} - \bar{x}) \text{ et } R_0 = \text{Var}(x_t)$$

La représentation graphique de la fonction d'autocorrélation (notée FAC) est appelée corrélogramme comme l'illustre la figure 1.2.

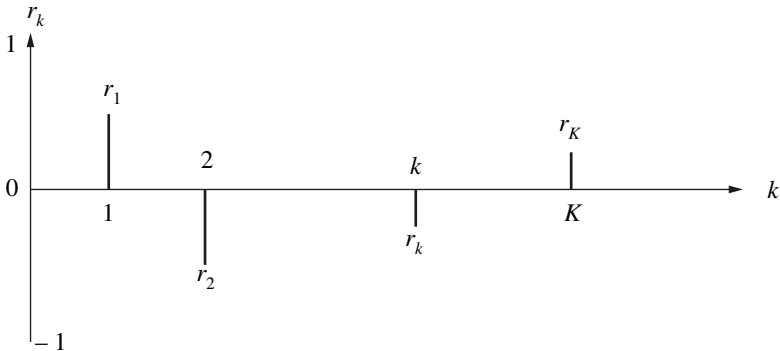


Figure 1.2 - Exemple de corrélogramme

Les graphiques suivants indiquent l'allure « attendue » de la fonction d'autocorrélation dans des cas simples. Ils permettent de vérifier que la FAC joue bien le rôle de décomposition temporelle de la chronique. Soit les quatre processus x_{1t} , x_{2t} , x_{3t} , et x_{4t} générés artificiellement et dont les représentations sont présentées figure 1.3.

$$x_{1t} = R_t \quad (R_t = \text{variable aléatoire gaussienne})$$

$$x_{2t} = a_1 t + a_0 + R_t \quad (t = \text{tendance } t = 1 \dots n)$$

$$x_{3t} = S_t + R_t \quad (S_t = \text{saisonnalité pure de quatre périodes})$$

$$x_{4t} = a_1 t + a_0 + S_t + R_t$$

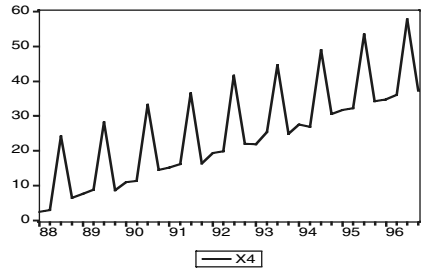
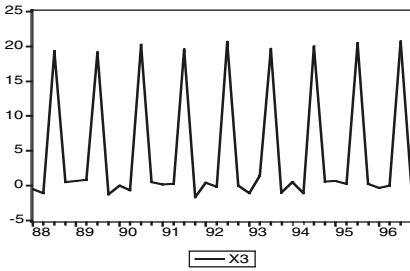
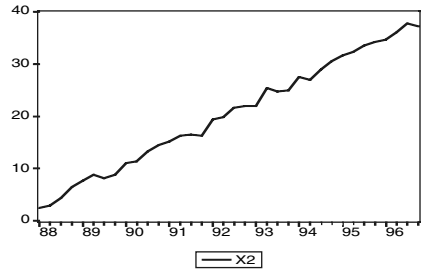
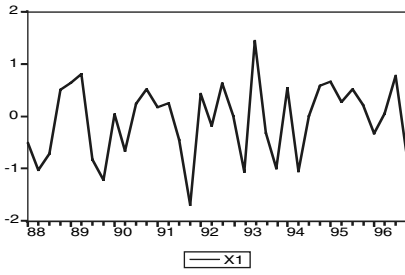


Figure 1.3 – Représentation des processus analysés

Les corrélogrammes respectifs de chacun des processus sont présentés sur la figure 1.4.

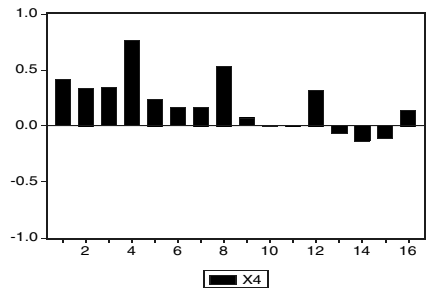
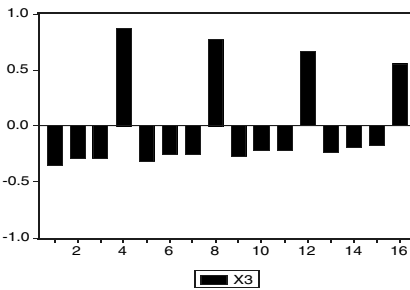
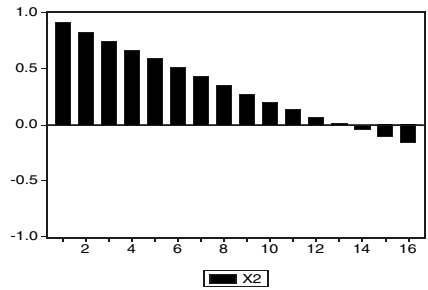
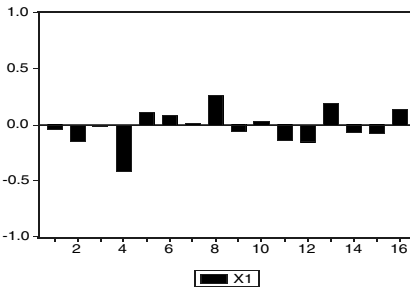


Figure 1.4 – Corrélogrammes des quatre processus

Les termes du corrélogramme de x_{1t} sont tous faibles à l'exception (fortuite) d'un seul, ceci est caractéristique d'une variable aléatoire de type « bruit blanc¹ ». Le corrélogramme de x_{2t} avec une décroissance lente de ses termes est spécifique d'une série affectée d'une tendance. Les termes très élevés (4, 8, 12 et 16) du corrélogramme de x_{3t} révèlent la saisonnalité de quatre périodes. Enfin, le corrélogramme de x_{4t} est une combinaison des trois précédents.

Exercice n° 1.2

 fichier C1EX1

Calcul d'une fonction d'autocorrélation

À partir des données relatives aux ventes du Tableau 1.1, on demande de calculer la FAC et de tester la significativité des coefficients d'autocorrélation par rapport à 0.

Solution

Nous allons détailler sur le tableau 1.6 les calculs pour un coefficient, par exemple r_2 , c'est-à-dire la corrélation entre x_t et x_{t-2} .

Tableau 1.6 – Exemple de calcul pour un coefficient d'autocorrélation d'ordre 2

t	x_t	x_{t-2}	$(x_t - \bar{x}_1)(x_{t-2} - \bar{x}_2)$	$(x_t - \bar{x}_1)^2$	$(x_{t-2} - \bar{x}_2)^2$
1	1 057	1 248	195 360,56	431 162,99	88 518,15
2	3 159	1 392	-222 259,17	209 0540,47	23 629,84
3	891	1057	402 349,24	676 391,15	239 336,23
4	1 065	3 159	-1 045 292,60	419 813,28	2 602 672,82
5	1 118	891	390 149,56	354 775,05	429 051,24
6	2 934	1 065	-586 700,47	1 490 767,59	230 899,47
7	1 138	1 118	246 239,35	330 659,57	183 372,34
8	1 456	2 934	-357 132,95	66 167,26	1 927 598,84
9	1 224	1 138	199 297,67	239 052,53	166 154,12
10	3 090	1 456	-123 679,42	1 896 046,27	8 067,63
Somme	17 132	1 5458	-901 668,23	7 995 376,17	5 899 300,68

1. Nous examinons, en détail, au chapitre 3 les caractéristiques d'un bruit blanc, succession de variables aléatoires homoscedastiques et non corrélées.

$$\bar{x}_1 = 1\,713,2 \text{ et } \bar{x}_2 = 1\,545,8$$

$$r_2 = \frac{\sum_{t=k+1}^n (x_t - \bar{x}_1)(x_{t-2} - \bar{x}_2)}{\sqrt{\sum_{t=k+1}^n (x_t - \bar{x}_1)^2 \sum_{t=k+1}^n (x_{t-2} - \bar{x}_2)^2}} = \frac{-901\,668,23}{\sqrt{7\,995\,376,17} \sqrt{5\,899\,300,68}} = -0,13$$

Le t empirique est égal pour $n = 10$ et $r_k = -0,13$: $t_c = \frac{|r_k|}{\sqrt{1 - r_k^2}} \sqrt{n - 2} = 0,38$. Le coefficient n'est pas significativement différent de 0.

Le Tableau 1.7 indique l'ensemble des valeurs de la fonction d'autocorrélation que l'on compare à la valeur lue dans la table de Student pour un seuil de 5 % et à $n - 2$ degrés de liberté. Seul le coefficient d'autocorrélation d'ordre 4 est significativement différent de 0 ; la périodicité des données étant trimestrielle, ce « pic » est donc attribué à la saisonnalité des données.

Tableau 1.7 – Calcul d'une fonction d'autocorrélation

Retards	r_k	n	$ t_c $	$d d l$	t lu à 0,95
0	1	12	–	–	–
1	– 0,395	11	1,29	9	2,262
2	– 0,132	10	0,38	8	2,306
3	– 0,33	9	0,95	7	2,365
4	0,952	8	7,62	6	2,447

D. Le spectre

En première analyse le spectre d'une chronique peut s'interpréter comme étant la décomposition de sa variance dans le domaine des fréquences. Le spectre est un graphique présentant en abscisse les fréquences d'apparition des cyclicités, ces fréquences s'exprimant :

- en radians pour la pulsation ω ;
- en hertz pour la fréquence λ ;
- en unité de temps pour la période T ,

en ordonnée sont portées les valeurs du spectre qui peuvent être assimilées aux différentes variances des cyclicités de la fréquence ω , T ou λ .

$$\omega = \frac{2\pi}{T} = 2\pi\lambda$$

À partir de ce graphe, nous pouvons mettre en évidence, dans la structure de la chronique, une cyclicité ($A\cos(\omega_0 t + \varphi)$) de fréquence ω_0 , d'amplitude A et de phase φ , si la valeur spectrale en ω_0 est élevée.

Nous verrons ultérieurement (chapitre 4) une définition plus rigoureuse de cet instrument ainsi que des simulations donnant la forme « attendue » du spectre.

II. La sélection du schéma

La saisonnalité d'une chronique peut parfois être influencée par l'extra-saisonnier et/ou le résidu. C'est de l'existence de ces interactions que sont nés les schémas de décomposition des séries chronologiques : additif, multiplicatif et multiplicatif complet. Nous présentons deux techniques simples pour sélectionner le schéma.

A. La procédure de la bande

La procédure de la bande consiste à partir de l'examen visuel du graphique de l'évolution de la série brute à relier, par une ligne brisée, toutes les valeurs « hautes » et toutes les valeurs « basses » de la chronique. Si les deux lignes sont parallèles, la décomposition de la chronique peut se faire selon un schéma additif ; dans le cas contraire, le schéma multiplicatif semble plus adapté. L'examen visuel du graphique des données du Tableau 1.1 révèle un schéma de type additif. Les figures 1.5 et 1.6 illustrent une interprétation de cette procédure.

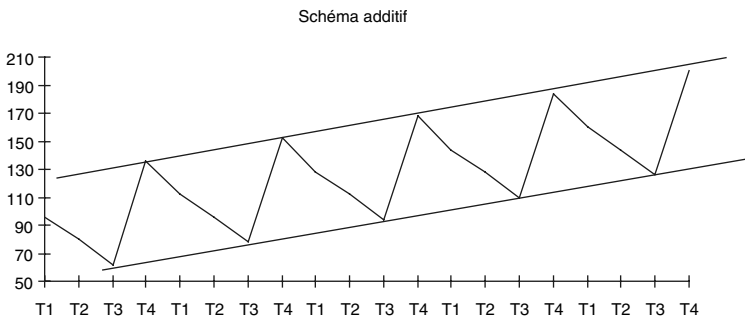


Figure 1.5 – Exemple de schéma additif

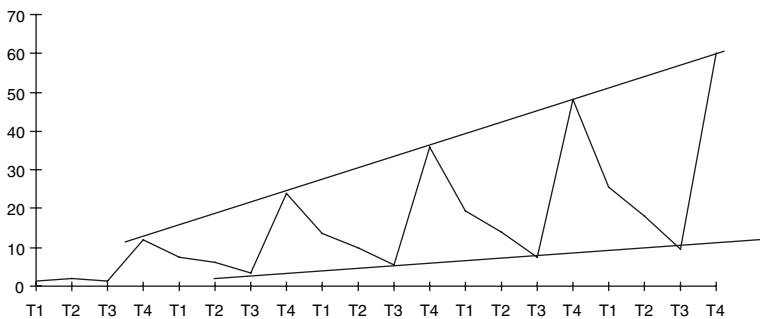


Figure 1.6 – Exemple de schéma multiplicatif

B. Le test de Buys-Ballot

Le test de Buys-Ballot est fondé sur les résultats du Tableau 1.1 (calcul des moyennes et des écarts types par année). Le schéma est, par définition, additif si l'écart type et la moyenne sont indépendants ; il est multiplicatif dans le cas contraire. Lorsque le nombre d'années est suffisant, nous pouvons estimer par la méthode des MCO les paramètres a_1 et a_0 de l'équation $\sigma_i = a_1 \bar{x}_i + a_0 + \varepsilon_i$ après avoir vérifié sur le graphique de régression (σ_i, \bar{x}_i) que le nuage peut être représenté par ce type de modèle. Dans le cas, où le coefficient a_1 n'est pas significativement différent de 0 (test de Student) alors on accepte l'hypothèse d'un schéma additif ; dans le cas contraire, nous retenons un schéma multiplicatif.

Bien qu'ici le nombre d'années (= 3) du Tableau 1.1 (Fichier C1EX1) soit très faible, nous procédons, à titre d'exemple, à l'estimation du modèle :

$$\sigma_i = -0,065\bar{x}_i + 929,57 + e_i$$

(-0,35)

$$R^2 = 0,12$$

$$n = 3$$

(.) = t de Student.

Le coefficient a_1 n'est pas significativement différent de 0, nous pouvons conclure à un schéma de type additif.

Ces deux procédures conduisent parfois à des résultats ambigus. Il est recommandé dans ce cas d'utiliser un schéma multiplicatif complet. De plus, lorsque le phénomène économique est observé sur un historique important, il arrive fréquemment que la structure de la chronique se modifie en passant d'un schéma additif à un schéma multiplicatif. Il convient donc de ne pas confondre la période historique avec le nombre d'observations : il est possible d'avoir beaucoup d'observations sur un laps de temps faible (des cours boursiers cotés en continu par exemple), où la structure de la chronique a peu de risque de se modifier.

III. Les méthodes de désaisonnalisation

Lorsque qu'une série chronologique est structurée par une saisonnalité, les comparaisons inter-temporelles du phénomène nécessitent une chronique Corrigée des Variations Saisonnières notée CVS ou encore désaisonnalisée.

Désaisonnaliser une chronique, c'est éliminer la saisonnalité sans modifier les autres composantes de la chronique. C'est une opération délicate ce qui explique le grand nombre de méthodes de désaisonnalisation.

Le choix de la technique la mieux appropriée dépend de la nature déterministe ou aléatoire (stochastique) de la saisonnalité de la chronique :

- lorsqu'elle est rigide (c'est-à-dire déterministe, bien marquée et répétitive), les méthodes de régression et l'emploi de coefficients saisonniers identiques sur la période historique sont adaptés ;
- lorsqu'elle est souple, c'est-à-dire aléatoire en amplitude et/ou en période, les techniques de filtrage par moyennes mobiles (CENSUS par exemple) doivent être utilisées.

A. Le principe de la conservation des aires

L'analyse de la saisonnalité a pour but une nouvelle répartition du profil « intra-annuel » de l'historique, sans modifier le niveau atteint en cumul annuel : les moyennes annuelles de la série brute et de la série CVS doivent être identiques. Ce principe de base – appelé le principe de la conservation des aires – permet de calculer des coefficients saisonniers définitifs.

- Dans le cas d'un schéma additif la somme des coefficients doit être nulle.
Soit S_j le $j^{\text{ème}}$ coefficient provisoire, on calcule la somme des coefficients :

$$S = \sum_{j=1}^p S_j \text{ avec } p \text{ la période de la saisonnalité.}$$

Si $S = 0$, les S_j sont les coefficients saisonniers définitifs.

Si $S \neq 0$, les coefficients saisonniers sont normés afin que leur somme soit nulle.

Les coefficients définitifs sont donnés par : $S_j^* = S_j - \bar{S}$.

- Dans le cas d'un schéma multiplicatif la moyenne des coefficients, pour une année donnée, doit être égale à 1.
Soit S_j le $j^{\text{ème}}$ coefficient provisoire, on calcule la somme des coefficients :

$$S = \sum_{j=1}^p S_j \text{ avec } p \text{ la période de la saisonnalité.}$$

Les coefficients définitifs sont donnés par : $S_j^* = S_j / \bar{S}$.

La méthode de désaisonnalisation doit être adaptée au type de saisonnalité : rigide ou souple.

B. Cas d'une saisonnalité rigide

1) Méthodes de désaisonnalisation par régression

a) Régression sur le temps

La figure 1.7 illustre le traitement d'une désaisonnalisation de ce type pour une série trimestrielle, schéma qu'il est facile d'adapter au cas d'une périodicité différente.

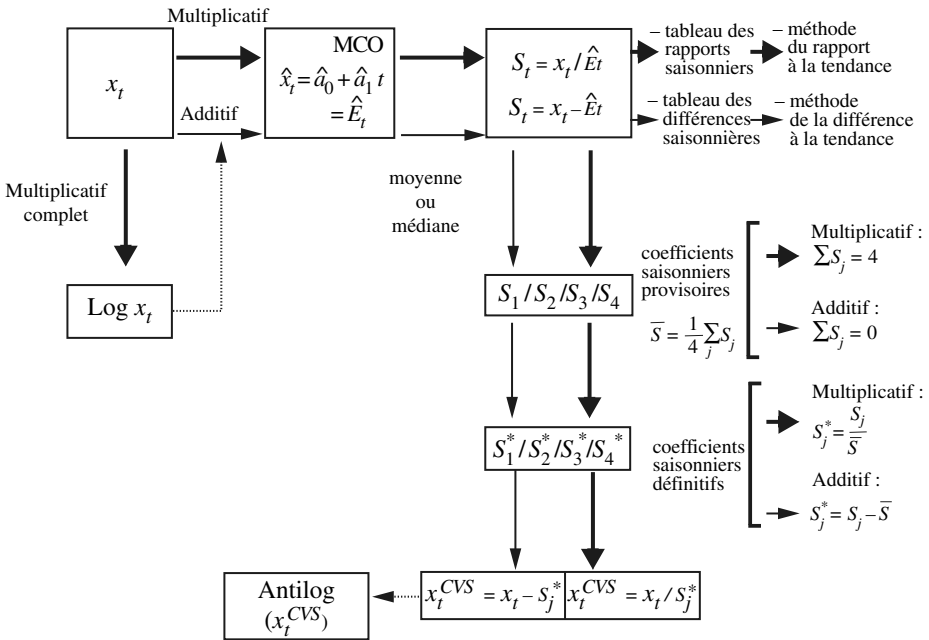


Figure 1.7 – Traitement pour une chronique trimestrielle

Exercice n° 1.3

fichier C1EX1

Désaisonnalisation selon un schéma additif

On demande de désaisonnaliser la série des ventes du Tableau 1.1 selon un schéma additif

Solution

Les résultats de cette méthode, à partir des données du Tableau 1.1, sont consignés sur le Tableau 1.8.

Tableau 1.8 – Calcul des coefficients saisonniers par le rapport à tendance – schéma additif

Dates	(2) Ventes	(3) \hat{E}_t	(4) Écart = $Ventes - \hat{E}_t$	(5) VENTES CVS
T1	1248,30	1308,79	- 60,49	1711,16
T2	1392,10	1370,42	21,68	1704,55
T3	1056,60	1432,04	- 375,44	1602,30
T4	3159,10	1493,66	1665,44	1838,09
T1	890,80	1555,29	- 664,49	1353,66
T2	1065,30	1616,91	- 551,61	1377,75
T3	1117,60	1678,54	- 560,94	1663,30
T4	2934,20	1740,16	1194,04	1613,19
T1	1138,20	1801,79	- 663,59	1601,06
T2	1456,00	1863,41	- 407,41	1768,45
T3	1224,30	1925,03	- 700,73	1770,00
T4	3090,20	1986,66	1103,54	1769,19
Moyenne	1647,73			1647,73
Écart type	866,72			151,60
CV ¹	0,53			0,09

S_1	$-462,85 = (-60,49 - 664,49 - 663,59)/3$
S_2	$-312,44 = (21,68 - 551,61 - 407,41)/3$
S_3	$-545,70 = (-375,44 - 560,94 - 700,73)/3$
S_4	$1321,00 = (1665,44 + 1194,04 + 1103,54)/3$

1. CV = coefficient de variation (rapport de l'écart type à la moyenne).

La tendance $\widehat{E}_t = \widehat{a}_0 + \widehat{a}_1 t$ est estimée à partir des données de la colonne (2). Les coefficients de régression sur le temps sont égaux à : $\widehat{a}_1 = 61,6$ et $\widehat{a}_0 = 1\,247,2$. Les écarts entre la série observée (colonne 2) et la tendance (colonne 3) donnent les coefficients saisonniers provisoires. Puis on rassemble les écarts relatifs aux mêmes trimestres afin de calculer leurs moyennes. Nous pouvons vérifier que la somme est égale à 0, le principe de la conservation des aires est vérifié, la normalisation des coefficients saisonniers ne s'impose pas.

Enfin, nous pouvons calculer la série CVS (colonne 5) par différence entre la série brute et le coefficient saisonnier du trimestre considéré. Le coefficient de variation – rapport de l'écart type à la moyenne – diminue considérablement entre la série brute et la série CVS : c'est l'effet réducteur de la variance de la désaisonnalisation.

L'utilisation de la méthode des MCO pour ajuster la tendance de la chronique nécessite un tableau de Buys-Ballot complet, ce qui limite (outre le choix de la forme rigide de la tendance et de la saisonnalité) la portée de cette méthode.

b) Régression sur fonction trigonométrique

La série chronologique est ici approchée au moyen de fonctions :

- l'extra-saisonnalité par un polynôme de degré k ,
- la saisonnalité par une somme de termes trigonométriques.

Dans cette dernière somme, le premier terme a pour période, la période de la saisonnalité ; les autres termes ont pour période, des périodes harmoniques. Ces fonctions trigonométriques sont des cycles du type $A \cos(\omega t + \varphi)$ avec :

A l'amplitude du cycle,

ω la pulsation, en radian, avec $\omega = \frac{2\pi}{T} = 2\pi\lambda$ où T , en unité de temps, est la

période de la saisonnalité ($T = 12$ pour une série mensuelle, $T = 4$ pour une série trimestrielle), λ la fréquence en hertz, φ la phase, en radian, du cycle à l'origine.

Cette expression contient deux inconnues A et φ . Il est difficile d'utiliser, dans ce cas, les MCO pour estimer les coefficients car l'expression est non linéaire ; c'est pourquoi on recourt à la forme développée :

$A \cos(\omega t + \varphi) = A \{ \cos \omega t \cos \varphi - \sin \omega t \sin \varphi \} = \beta_1 \cos(\omega t) + \gamma_1 \sin(\omega t)$
avec $\beta_1 = A \cos \varphi$ et $\gamma_1 = -A \sin \varphi$

À titre d'exemple, considérons une série chronologique mensuelle organisée selon un schéma additif : $x_t = E_t + S_t + R_t$

Avec cette méthode, la chronique est modélisée par un polynôme de degré quelconque k plus une somme de fonction trigonométrique.

$$x_t = \alpha_0 + \sum_{i=1}^k \alpha_i \frac{t^i}{i!} + \sum_{i=1}^m A_i \cos\left(\frac{2\pi}{T} i t + \varphi_i\right) + R_t$$

avec :

– périodicité mensuelle $T = 12$;

– k le degré du polynôme ;

– m le nombre d'harmoniques $\left(m = \frac{T}{2} - 1 = 5\right)$

Pour un extra-saisonnier linéaire, par exemple, le modèle s'écrit :

$$x_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^5 A_i \cos\left(\frac{2\pi}{12}it + \varphi_i\right) + R_t$$

$$x_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^5 \left[\beta_i \cos \frac{2\pi}{12}it + \gamma_i \sin \frac{2\pi}{12}it \right] + R_t$$

$$\begin{aligned} x_t = & \alpha_0 + \alpha_1 t + \beta_1 \cos \frac{2\pi}{12}t + \beta_2 \cos \frac{2\pi}{6}t + \beta_3 \cos \frac{2\pi}{4}t + \beta_4 \cos \frac{2\pi}{3}t \\ & + \beta_5 \cos \frac{2\pi}{2,4}t + \gamma_1 \sin \frac{2\pi}{12}t + \gamma_2 \sin \frac{2\pi}{6}t + \gamma_3 \sin \frac{2\pi}{4}t + \gamma_4 \sin \frac{2\pi}{3}t \\ & + \gamma_5 \sin \frac{2\pi}{2,4}t + R_t \end{aligned}$$

Nous sommes en présence d'un Modèle Linéaire Général Simple (MLGS) où x_t est la variable endogène et $t = 1, 2, \dots, n$.

L'application de la méthode des MCO à ce modèle donne les estimations des paramètres :

$$\hat{x}_t = \hat{\alpha}_0 + \hat{\alpha}_1 t + \sum_{i=1}^5 \hat{\beta}_i \cos \frac{2\pi}{12}it + \sum_{i=1}^5 \hat{\gamma}_i \sin \frac{2\pi}{12}it$$

La chronique désaisonnalisée est égale à

$$x_t^{CVS} = x_t - \left(\sum_{i=1}^5 \hat{\beta}_i \cos \frac{2\pi}{12}it + \sum_{i=1}^5 \hat{\gamma}_i \sin \frac{2\pi}{12}it \right)$$

Cette méthode porte le nom de régression sur fonction trigonométrique ou régression sur séries de Fourier.

Cette technique est aussi une modélisation complète de la structure de la série chronologique ; elle peut donc être utilisée comme méthode de prévision.

En tant que technique de désaisonnalisation, elle présente les intérêts de la simplicité de mise en œuvre (un logiciel de régression multiple suffit), de la rapidité des calculs, de la fiabilité par l'utilisation de la théorie des tests de la régression ; les inconvénients sont liés aux hypothèses de la méthode des MCO (en particulier l'hypothèse de non-colinéarité des variables explicatives) et à l'extra-saisonnalité qui est assimilée à une fonction de type polynôme.

Cette méthode convient bien pour des séries à la structure très rigide, mais elle n'est pas toujours bien adaptée pour des séries économiques où les composantes sont le plus souvent mouvantes.

Exercice n° 1.4

fichier C1EX4

Désaisonnalisation par fonction trigonométrique

On demande de désaisonnaliser par régression sur fonction trigonométrique les données de production d'électricité en Australie connue mensuellement sur 31 ans.

Solution

L'application de cette méthode sur les données de production d'électricité en Australie, entre janvier de l'année 1 et décembre de l'année 31, peut être programmée à partir du logiciel Eviews (fichier C1EX4.PRG) à l'aide des instructions suivantes (le schéma étant multiplicatif, nous prenons le logarithme de la série brute) :

```
GENR TREND = @trend
GENR LELEC = LOG(ELEC)
SCALAR PI = 3.14116
FOR !I = 1 TO 5
GENR CS!I = COS(2*PI/12*TREND*!I)
GENR SN!I = SIN(2*PI/12*TREND*!I)
NEXT
EQUATION EQ1.LS LELEC C TREND CS1 CS2 CS3 CS4 CS5 SN1 SN2 SN3 SN4 SN5
GENR LELECVS = LELEC - (C(3)*CS1+C(4)*CS2+C(5)*CS3+C(6)*CS4+C(7)*CS5+
C(8)*SN1+C(9)*SN2+C(10)*SN3+C(11)*SN4+C(12)*SN5)
GENR ELECVS = EXP(LELECVS)
```

Les résultats de la régression sont les suivants :

LS// Dependent Variable is LELEC
Included observations : 372

Variable	CoefficienStd.	Error	Statistic	Prob.
C	7.348962	0.008593	855.2390	0.0000
TREND	0.005601	4.01E-05	139.7001	0.0000
CS1	-0.053527	0.006087	-8.794380	0.0000
CS2	-0.004526	0.006086	-0.743596	0.4576
CS3	0.017199	0.006086	2.825898	0.0050
CS4	-0.000707	0.006086	-0.116124	0.9076
CS5	-0.019774	0.006084	-3.250100	0.0013
SN1	-0.094215	0.006087	-15.47682	0.0000
SN2	0.016409	0.006086	2.696039	0.0073
SN3	0.002221	0.006086	0.364936	0.7154
SN4	0.000238	0.006087	0.039103	0.9688
SN5	-0.012598	0.006088	-2.069302	0.0392
R-squared	0.982319	F-statistic	1818.237	

Les variables sont dans l'ensemble significatives à l'exception de CS4, SN3 et SN4 dont les probabilités critiques sont supérieures à 0,05.

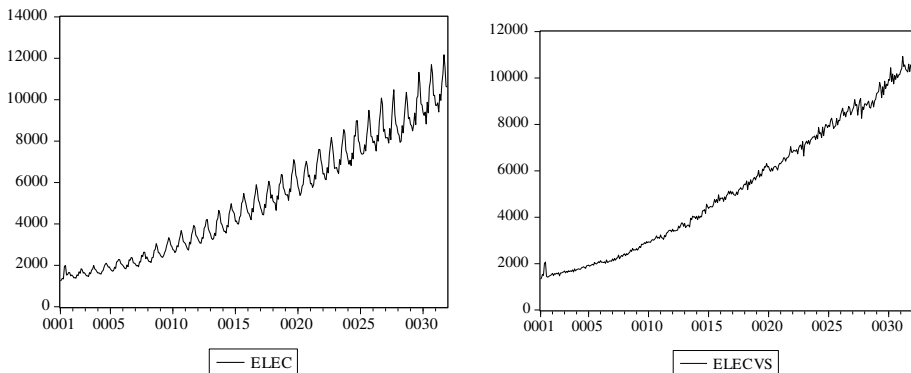


Figure 1.8 – Série de production d'électricité brute et CVS

L'examen de la figure 1.8 indique les limites de cette méthode : si la désaisonnalisation fonctionne bien au centre de l'historique, il demeure « un bruit » aux extrémités.

c) Régression sur variables indicatrices

Le principe est identique au précédent. L'extra-saisonnalité (E_t) est modélisée par une fonction polynôme, la composante saisonnière (S_t) par un ensemble de variables dichotomiques¹, notées $IND_{t,i}$.

$$x_t = \alpha_0 + \sum_{i=1}^k \alpha_i \frac{t^i}{i!} + \sum_{i=1}^T \delta_i IND_{t,i} + R_t \text{ avec } t = 1, 2, \dots, n \text{ et } T \text{ la période de}$$

la saisonnalité (la plus souvent $T = 12$).

$IND_{t,i} = 1$ si t correspond à la période saisonnière (ou t est congru à i modulo T),

= 0 sinon.

Soit une chronique trimestrielle ($T = 4$) obéissant à un schéma additif et possédant une extra-saisonnalité E_t de type linéaire ($k = 1$).

Le schéma structurant de la chronique s'écrit alors :

$$x_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^4 \delta_i IND_{t,i} + R_t$$

Sous forme matricielle ce modèle s'écrit :

$$\begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ \cdot & 2 & 0 & 1 & 0 & 0 \\ \cdot & 3 & 0 & 0 & 1 & 0 \\ \cdot & 4 & 0 & 0 & 0 & 1 \\ \cdot & \cdot & 1 & \cdot & \cdot & 0 \\ 1 & n & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix} + \begin{bmatrix} R_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ R_n \end{bmatrix}$$

Dans la pratique, la $T^{\text{ème}}$ variable indicatrice n'est pas utilisée car la somme des T variables indicatrices serait alors égale au vecteur unité et il y aurait donc colinéarité stricte avec le vecteur unité relatif au terme constant. Il existe plusieurs solutions à cette colinéarité². La plus utilisée consiste à ne calculer que

1. Les termes de variables indicatrices, variables muettes sont aussi employés, les Anglo-Saxons utilisent le terme de variable *dummy*. Pour des exemples d'application cf. Bourbonnais (2009).
2. Nous pouvons estimer aussi un modèle sans terme constant sur variables centrées.

$T - 1$ coefficients de saisonnalité et de vérifier pour le $T^{\text{ième}}$ le principe de conservation des aires : $\sum_{i=1}^4 \delta_i = 0 \Rightarrow \delta_4 = -(\delta_1 + \delta_2 + \delta_3)$

Le modèle sous forme matricielle s'écrit alors :

$$\begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 \\ \cdot & 3 & 0 & 0 & 1 \\ \cdot & 4 & -1 & -1 & -1 \\ \cdot & \cdot & 1 & 0 & 0 \\ 1 & n & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \delta_1 \\ \delta_2 \\ \delta_3 \end{pmatrix} + \begin{pmatrix} R_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ R_n \end{pmatrix}$$

Les deux méthodes (fonctions trigonométriques et variables indicatrices) conduisent à des résultats similaires. Elles présentent des avantages et des inconvénients identiques. Ces méthodes ne sont applicables que pour désaisonnaliser des séries chronologiques structurées selon un schéma additif ou multiplicatif complet.

Exercice n° 1.5

 **fichier C1EX1**

Désaisonnalisation à l'aide de variables dichotomiques

On demande de désaisonnaliser les ventes du Tableau 1.1 par la méthode des variables dichotomiques.

Solution

On génère donc trois variables indicatrices $IND1$, $IND2$ et $IND3$ puis on effectue la régression avec les ventes comme variable à expliquer et quatre variables explicatives ($IND1$, $IND2$, $IND3$ et une tendance).

Les résultats de la régression (cf. Tableau 1.9) indiquent les quatre coefficients saisonniers.

$$S_1 = -552,89$$

$$S_2 = -342,52$$

$$S_3 = -515,48$$

D'où $S_4 = 1410,89$ (puisque la somme des quatre coefficients saisonniers doit être nulle).

Tableau 1.9 – Résultat de la régression sur variables indicatrices (Eviews)

Dependent Variable : VENTE				
Included observations : 12				
Variable	Coefficient	Std. Error	t - Statistic	Prob.
C	1637.104	107.4081	15.24190	0.0000
TREND	1.625000	14.76525	0.110056	0.9155
IND1	- 552.8958	86.41142	- 6.398412	0.0004
IND2	- 342.5208	83.85051	- 4.084899	0.0047
IND3	- 515.4792	83.85051	- 6.147597	0.0005
R - squared		0.976356	F - statistic	72.26444

Nous remarquons que la tendance n'est pas significative.

d) La correction des variations saisonnières à effets saisonniers additif et multiplicatif simultanés¹

Dans la pratique les séries chronologiques ne sont pas toujours le résultat d'un schéma de décomposition additif pur ou multiplicatif pur. Le schéma de décomposition mixte peut être de la forme suivante : $x_t = (T_t \times S_t^M) + S_t^A$, avec $T_t =$ Tendance, $S_t^M =$ coefficient saisonnier multiplicatif, $S_t^A =$ coefficient saisonnier additif.

Estimateurs des moindres carrés pour le schéma mixte

Notations :

- n : nombre de cycles (d'années) d'indice courant i
- p : périodicité de la série, nombre de périodes (d'indice courant k) dans le cycle
- i : indice de cycle ; $i = 1, \dots, n$
- k : indice de période ; $k = 1, 2, \dots, p : k = (t - 1 \text{ modulo } p) + 1$
- t : indice de date d'observation, $t = (i - 1)p + k$ correspondant à la k^e période du i^e cycle
- $x_t = x_k^i$: série brute à la date $t = (i - 1)p + k$ correspondant à la k^e période du i^e cycle
- T_t : valeur de la vraie tendance à la date t
- $\widehat{T}_t = T_k^i$: tendance linéaire estimée à la date t
- S_k^M : coefficient saisonnier fixe multiplicatif associé à la période k

1. Cf. Bourbonnais R. et Vallin Ph., 2007.

- S_k^A : coefficient saisonnier fixe additif associé à la période k
- E_t : résidu, partie non expliquée de la série

Le principe consiste à déterminer simultanément par des estimateurs tenants compte de la mixité du modèle les deux jeux des coefficients saisonniers S_k^M et S_k^A .

En observant que, si la tendance est connue (ou estimée), le modèle est linéaire en fonction des coefficients saisonniers, nous pouvons estimer les coefficients saisonniers S_k^M et S_k^A ($k = 1, 2, \dots, p$) par les estimateurs des moindres carrés, solution du système linéaire classique à $2p$ inconnues et à $2p$ équations :

$$S = (A'A)^{-1} A'X$$

où :

- $X(np \times 1)$: le vecteur colonne des x_t pour n cycles de p périodes consé-

$$\text{cutifs, } X = \begin{bmatrix} x_1^1 \\ x_k^i \\ \dots \\ x_p^n \end{bmatrix}$$

- X est composé des n sous matrices associées aux cycles i : $X_i = \begin{bmatrix} x_1^i \\ \dots \\ x_p^i \end{bmatrix}$,

$$i = 1 \dots n$$

- $S(2p \times 1)$: le vecteur colonne de composantes $\begin{bmatrix} S_k^M \\ S_k^A \end{bmatrix}$; $k = 1 \dots p$

- $A(np \times 2p)$: la matrice $\begin{bmatrix} T^1 I \\ T^i I \\ T^n I \end{bmatrix}$; où T^i est la matrice diagonale, d'ordre

p , d'éléments $\{T_k^i\}$ pour le cycle i , k variant sur les p périodes

- I est la matrice unité d'ordre p
- A' transposée de A
- E : matrice colonne des résidus E_t

Les estimateurs de S_k^M et S_k^A sont :

$$\widehat{S}_k^M = \frac{\sum_{i=1}^n T_k^i x_k^i - n \bar{T}_k \bar{x}_k}{\sum_{i=1}^n (T_k^i)^2 - n \bar{T}_k^2}$$

$$\widehat{S}_k^A = \bar{x}_k - \bar{T}_k \widehat{S}_k^M$$

$$\text{Avec } \bar{T}_k = \frac{1}{n} \sum_{i=1}^n T_k^i \text{ et } \bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_k^i.$$

Nous retrouvons les estimateurs classiques du modèle linéaire à une variable explicative avec constante où, ici, la variable explicative est la tendance estimée à la date t : $\widehat{T}_t = T_k^i$. Ce résultat provient du fait que l'estimation des coefficients saisonniers s'effectue par la résolution de p blocs indépendants de n équations.

Le modèle de base est : $x_t = T_t \times S_k^M + S_k^A + E_t$ ou avec la notation à deux indices :

$$x_k^i = T_k^i \times S_k^M + S_k^A + E_k^i$$

Le système s'écrit matriciellement : $X = AS + E$ [1]

Chaque matrice carrée T^i , $i = 1, \dots, n$, de dimension $(p \times p)$ composant A est de la forme :

$$T^i = \begin{bmatrix} T_1^i & 0 & 0 \\ 0 & T_k^i & 0 \\ 0 & 0 & T_p^i \end{bmatrix}$$

On est donc ramené à p systèmes de n équations de régression indépendantes, chaque système k , $1 \leq k \leq p$, étant formé des n équations associées aux n observations du k^e cycle des n périodes permettant d'estimer les deux paramètres S_k^M et S_k^A par les résultats classiques de la régression linéaire à deux variables.

L'ensemble des équations de régression du système k s'écrit :

$$x_k^i = T_k^i \times S_k^M + S_k^A + E_k^i, \quad \text{pour } i = 1, 2, \dots, n$$

Remarque : si $A'A$ est inversible (c'est le cas en particulier si $\widehat{T}_t = a_1 t + a_0$ avec $a_1 \neq 0$), les estimateurs des coefficients saisonniers qui minimisent la somme des carrés des résidus sont solutions du système : $S = (A'A)^{-1} A'X$ [2]

Les sous-matrices diagonales permettent un calcul matriciel relativement aisé et la relation [2] fournit les estimateurs présentés.

Cette méthode d'estimation des coefficients saisonniers permet de tester la significativité des coefficients des deux types et, par suite, d'opter pour l'un des trois modèles (additif, multiplicatif, mixte).

Exercice n° 1.6

 fichier C1EX5 (énoncé et solution)

Un exemple d'application de saisonnalité mixte (additive et multiplicative) : le cas de la téléphonie mobile

Le marché de la téléphonie mobile est un marché récent, trois opérateurs interviennent sur le marché français : Orange, SFR, et Bouygues Telecom. Ce caractère oligopolistique engendre une activité marketing très forte et des offres nouvelles très fréquentes.

Nous nous situons fin 1999 (marché non encore mature), période pour laquelle la prévision du nombre d'abonnés (les achats des cartes Sim¹) s'avérait cruciale pour chacun des opérateurs. La saisonnalité (cf. figure 1.9) indique un pic très important au mois de décembre (cadeaux de fin d'année pour les particuliers), puisque le nombre d'abonnés est deux fois plus élevé que la moyenne des ventes de l'année.

Deux segments de marché bien distincts coexistent en termes de comportement et de saisonnalité : le professionnel et le particulier.

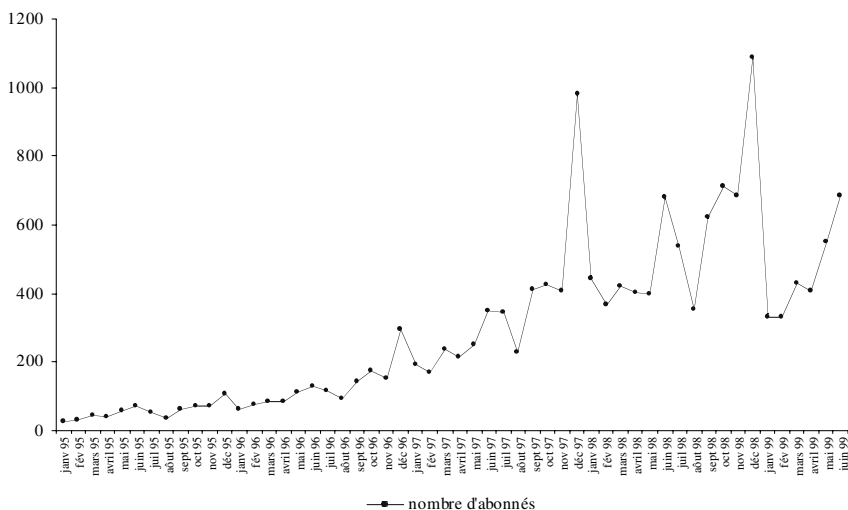


Figure 1.9 – Nombre d'abonnés (cartes Sim vendues) à la téléphonie mobile en France en milliers

Sur cet historique de ventes (janvier 1995 à juin 1999, 54 observations) nous allons appliquer les trois méthodes de désaisonnalisation (cf. Tableau 1.10), bien que le schéma additif s'avère, au vu du graphique, à proscrire.

1. La carte Sim (*Subscriber Identity Module*) est la carte à puce que l'on glisse dans l'appareil téléphonique et qui permet l'identification de l'abonné.

Tableau 1.10 – Coefficients saisonniers estimés selon trois méthodes

Mois	Additif Pur	Multiplicatif Pur	Schéma Mixte	
			Additif	Multiplicatif
Janvier	- 29,29	0,88	- 4,3	0,93
Février	- 67,27	0,81	22,4	0,70
Mars	- 35,87	0,93	33,8	0,78
Avril	- 67,24	0,83	24,6	0,72
Mai	- 64,94	0,89	61,0	0,63
Juin	45,67	1,14	- 23,1	1,20
Juillet	- 22,54	0,94	4,0	0,93
Août	- 113,40	0,63	10,3	0,60
Septembre	10,96	1,03	- 11,7	1,09
Octobre	40,16	1,13	- 21,2	1,21
Novembre	13,65	1,02	- 34,6	1,16
Décembre	290,12	1,78	- 61,3	2,05

À la lecture des coefficients saisonniers estimés, nous constatons de fortes différences. Il est intéressant de constater que dans le cadre du schéma mixte, les coefficients additifs et multiplicatifs ne connaissent pas le même profil d'évolution. Cela nous suggère que pour le marché professionnel, justifiable plutôt du schéma additif car déjà stabilisé dans les années 1998, le profil saisonnier n'est pas le même que pour les particuliers, relevant du schéma multiplicatif avec un marché en pleine expansion soumis à trois pointes : juin (avant les départs en vacances), octobre (rentrées scolaire et universitaire) et décembre (fêtes de fin d'année).

Afin d'évaluer la pertinence de ces trois méthodes, nous comparons la somme des carrés des écarts (*cf.* Tableau 1.11) entre valeurs observées et valeurs ajustées à l'aide d'un modèle à tendance linéaire et saisonnalité.

Tableau 1.11 – Somme des carrés des écarts entre les valeurs observées et les valeurs ajustées

	Schéma additif	Schéma multiplicatif	Schéma mixte
Somme des carrés des écarts	312 501	90 163	40 801
Degrés de liberté	54 - 11 = 43	43	32
Carré moyen	7 267,46	2 096,81	1 316,16

En faisant l'hypothèse de la normalité des écarts, nous pouvons procéder à un test de Fisher tel que :

$$F^* = \frac{(90\,163 - 40\,801)/11}{40\,801/32} = 3,51 > F_{11;32}^{0,05} = 2,10.$$

Ce test laisse présager qu'il existe une différence significative de la somme des carrés des écarts entre le schéma multiplicatif et le schéma mixte.

Comme nous pouvions l'anticiper, le schéma additif s'avère très mauvais et le schéma mixte préférable au schéma multiplicatif pur (carré moyen deux fois plus faible).

Nous comparons (cf. figure 1.10) les résidus d'estimation entre le schéma mixte et le schéma multiplicatif.

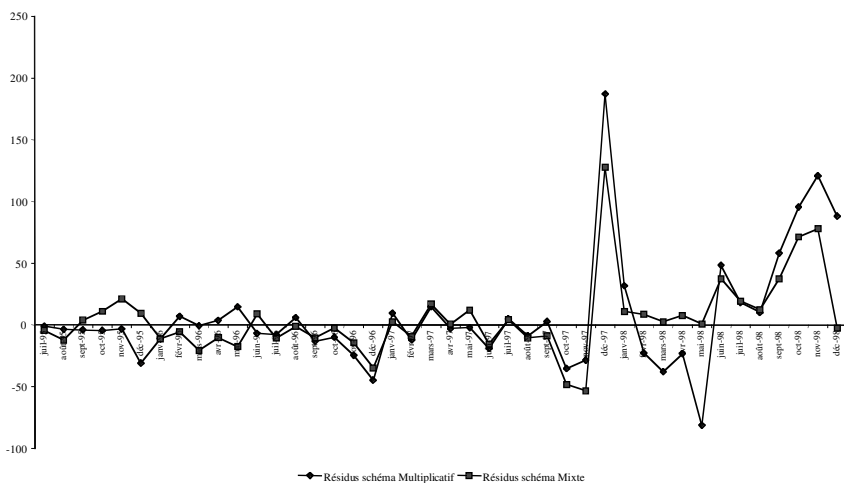


Figure 1.10 – Résidus d'estimation schéma multiplicatif et schéma mixte

Sur la période 1995 à 1998, le schéma multiplicatif procure des résidus d'estimation d'égale amplitude aux résidus du schéma mixte. En revanche, après 1998 – période d'émergence du marché pour les particuliers – le schéma mixte fournit un ajustement de meilleure qualité, comme l'illustre le Tableau 1.12.

Enfin, nous pouvons noter que dans le cas d'utilisation de cette méthode pour élaborer des prévisions, des variables d'intervention pour décembre 1997 et décembre 1998 devraient être intégrées au modèle.

Tableau 1.12 – Somme des carrés des écarts entre les valeurs observées et les valeurs ajustées sous-période janvier 1998 à juin 1999

	Schéma multiplicatif	Schéma mixte
Somme des carrés des écarts	47 709	14 713

Le schéma mixte, sur cet exemple et sur la totalité de la période, semble préférable à un schéma multiplicatif. De plus, il permet d'identifier pour ce marché les coefficients saisonniers multiplicatifs et additifs correspondants à deux marchés distincts (professionnels et particuliers) qui émergent à partir de 1998.

2) Les méthodes de filtrage

Un filtre est une transformation mathématique de la chronique. La chronique entrante porte le nom d'*input* du filtre et le résultat de la transformation (la chronique sortante) porte le nom d'*output*.



Le filtre le plus utilisé pour désaisonnaliser une chronique est celui de la moyenne mobile.

a) Les moyennes mobiles simples

La moyenne mobile simple est un filtre symétrique à horizon fini¹. Il s'agit d'une succession de moyennes arithmétiques de longueur choisie égale à L (appelée ordre de la moyenne mobile). Les formules générales de filtrage par moyenne mobile symétrique sont les suivantes :

- Si l'ordre de la moyenne mobile correspond à un nombre impair

$$(L = 2m + 1, \text{ avec } m \in N^*), y_t = \frac{1}{2m + 1} \left[\sum_{i=-m}^{i=m} x_{t+i} \right]$$

- Si l'ordre correspond à un nombre pair ($L = 2m$, avec $m \in N^*$), il faut recourir à un artifice de calcul afin de faire correspondre le terme central x_t à la valeur de la moyenne mobile y_t .

$$y_t = \frac{1}{2m} \left[\frac{1}{2}x_{t-m} + \sum_{i=-m+1}^{i=m-1} x_{t+i} + \frac{1}{2}x_{t+m} \right]$$

Les moyennes mobiles simples ont pour propriétés d'éliminer une saisonnalité de période égale à l'ordre de la moyenne mobile, de laisser « passer » l'extra-saisonnalité et de lisser le résidu. Ces propriétés fondent les méthodes de désaisonnalisation utilisant ces filtres. Par construction, les moyennes mobiles sont plus courtes – en nombre d'observations – du fait de la perte de m valeurs aux extrémités de la chronique.

b) Désaisonnalisation par les moyennes mobiles simples

L'organigramme de désaisonnalisation de la chronique par l'usage de la moyenne mobile simple est identique à la figure 1.7, seule change l'étape de calcul de l'extra-saisonnalité (la tendance) qui est appréhendée, dans ce cas, par une moyenne mobile d'ordre égal à la période de la saisonnalité.

1. Ce filtre est décrit plus précisément au chapitre 4.

Exercice n° 1.7

fichier C1EX1

Désaisonnalisation par moyenne mobile selon un schéma additif

On demande de désaisonnaliser la série des ventes du Tableau 1.1 par moyenne mobile (schéma additif)

Solution

Les résultats de cette méthode sont consignés sur le Tableau 1.13.

Exemple de calcul :

$$\hat{E}_3 = (0,5 \times 1248 + 1392 + 1057 + 3159 + 0,5 \times 891)/4 = 1\,669,34$$

Les écarts relatifs aux mêmes trimestres sont rassemblés et leurs moyennes sont calculées (coefficients saisonniers provisoires). Enfin, les coefficients saisonniers sont normalisés afin de respecter le principe de la conservation des aires.

Tableau 1.13 – Calcul des coefficients saisonniers par écart par rapport à la moyenne mobile

DATES	VENTES	\hat{E}_t	(4) Écart = Ventes – \hat{E}_t	VENTES CVS
T1	1 248,30			1 841
T2	1 392,10			1 745
T3	1 056,60	1 669,34	– 612,74	1 565
T4	3 159,10	1 583,80	1 575,30	1 705
T1	890,80	1 550,58	– 659,78	1 484
T2	1 065,30	1 530,09	– 464,79	1 418
T3	1 117,60	1 532,90	– 415,30	1 626
T4	2 934,20	1 612,66	1 321,54	1 480
T1	1 138,20	1 674,84	– 536,64	1 731
T2	1 456,00	1 707,67	– 251,67	1 809
T3	1 224,30			1 732
T4	3 090,20			1 636
Moyenne	1 647,73			1 647,73
Écart type	866,72			136,63
CV	0,53			0,08

Tableau 1.13 – Calcul des coefficients saisonniers par écart par rapport à la moyenne mobile (suite)

	Coefficients saisonniers provisoires	Coefficients saisonniers définitifs
S_1	$-598,21 = (-659,78 - 536,64)/2$	$-592,70$
S_2	$-358,23 = (-464,79 - 251,67)/2$	$-352,72$
S_3	$-514,02 = (-612,74 - 415,30)/2$	$-508,51$
S_4	$1\ 448,42 = (1\ 575,30 + 1\ 321,54)/2$	$1\ 453,93$
<i>Somme</i>	$-22,04$	$0,00$

Nous pouvons calculer la série CVS par différence entre la série brute et les coefficients saisonniers définitifs du trimestre. Le coefficient de variation, là encore, diminue considérablement entre la série brute et la série CVS.

L'utilisation de coefficients saisonniers identiques pour la période historique de la chronique montre que cette technique est surtout utilisable dans le cas d'une saisonnalité rigide.

C. Cas d'une saisonnalité souple

1) Les méthodes CENSUS¹

La désaisonnalisation par la moyenne mobile simple d'une chronique est souvent insuffisante pour diverses raisons (saisonnalité fluctuante, extra-saisonnalité complexe...).

En 1954, J. Shiskin, spécialiste des cycles économiques, propose une méthode de désaisonnalisation utilisant, de façon itérative, plusieurs moyennes mobiles. Shiskin, chercheur au *Bureau of Census* des États-Unis a donné le nom de Census-1 à cette méthode. Depuis 1954, elle a connu de nombreuses améliorations à partir de Census-2 et ses onze versions expérimentales. En 1967, Shiskin associé à A. H. Young et J. C. Musgrave propose la version Census \times 11 qui est restée longtemps la méthode la plus utilisée pour désaisonnaliser les chroniques (notamment par l'INSEE). En 1979, E. B. Dagum fait évoluer la version Census \times 11 en Census \times 11 ARIMA.

Enfin actuellement, le *bureau of Census* utilise la version X12-ARIMA (appelée aussi Reg-ARIMA) développée par Findley.

1. Pour approfondir sur la désaisonnalisation cf. Darne, 2002.

Comme les méthodes Censu sont bâties à partir d'itérations successives de moyennes mobiles d'ordre différent pour mieux appréhender l'extra-saisonnalité (E_t) de la chronique ainsi que les fluctuations de la saisonnalité, elles font perdre de l'information à l'extrémité terminale de la chronique. Cette perte d'information est comblée par une prévision de type Box et Jenkins (*cf.* chapitre 7), avant la désaisonnalisation de la chronique, d'un nombre de points égal à la perte d'information inhérente à l'utilisation des moyennes mobiles.

De façon concurrente, d'autres méthodes ont vu le jour, par exemple celle CPB X11 élaborée par le *Central Bureau of Statistics* des Pays Bas. Il s'agit d'une version corrigée de la technique CPB1 – créée en 1980 par le *Central Planning Bureau*. Elle a été construite pour des chroniques dont la saisonnalité évolue rapidement (instabilité du schéma de décomposition).

Nous présentons, à titre illustratif, l'algorithme¹ simplifié, en cinq étapes, de la version trimestrielle de Censu $\times 11$ selon un schéma additif.

a) Première estimation de la composante saisonnière

Application d'une moyenne mobile (M^0) centrée d'ordre 5 avec pour coefficients : 0,125 ; 0,250 ; 0,250 ; 0,250 ; 0,125 à la série x_t . Soit z_t la série filtrée (première estimation de la tendance) ; on calcule $S1_t = x_t - z_t$ puis, de nouveau, on applique à la série $S1_t$ une moyenne mobile (M^1) centrée de 5 ans avec pour coefficients : 0,111 (en $t - 8$ et $t + 8$) ; 0,222 (en $t - 4$ et $t + 4$) ; 0,333 (en t) ; (la somme des cinq valeurs étant divisée par 0,999). La série résultante est lissée par un filtre S de coefficients $-0,125$ (en $t - 2$ et $t + 2$) ; $-0,250$ (en $t - 1$ et $t + 1$) ; et $0,750$ en t . La série filtrée est une première estimation des coefficients saisonniers $CS1_t$.

b) Première correction pour valeurs aberrantes

Pour chaque trimestre, l'écart type σ_1 de $CS1_t$ est calculé et on corrige les valeurs de $S1_t$ à l'extérieur de l'intervalle $CS1 \pm \sigma_1$. La nouvelle série ainsi obtenue est lissée par le filtre M^1 puis par le filtre S . La série résultante est une deuxième estimation des coefficients saisonniers $CS2_t$, soit $x_t^{CVS1} = x_t - CS2$.

c) Estimation de la tendance

On applique à x_t^{CVS1} une moyenne mobile de Henderson d'ordre 5 ayant pour coefficients : -21 (en $t - 2$ et $t + 2$) ; 84 (en $t - 1$ et $t + 1$) ; et 160 en t , le total étant divisé par 286. La série résultante zz_t est une nouvelle estimation de la tendance.

d) Évaluation finale de la composante saisonnière

L'estimation de la saisonnalité est donnée par : $S2_t = x_t - zz_t$, on lui applique une moyenne mobile de 7 ans de coefficients 1 (en $t - 12$ et $t + 12$) ; 2 (en $t - 8$ et $t + 8$) ; 3 (en $t - 4$ et $t + 4$) ; et 3 en t ; afin d'obtenir une somme annuelle des coefficients nulle, on applique au résultat le filtre S et on appelle

1. Laroque 1977

CS_t , le résultat obtenu. Enfin, une dernière correction des valeurs aberrantes est effectuée sur la série CVS : $x_t^{CVS2} = x_t - CS_t$.

Toutes ces opérations sont itérées deux fois. Cette méthode utilise 21 moyennes mobiles dont il est démontré qu'elles se regroupent en une seule portant sur 28 trimestres de part et d'autre de l'instant considéré. Cette méthode d'analyse n'est donc applicable qu'à des séries longues de minimum huit années d'historiques. Le lecteur peut facilement par lui-même utiliser cette méthode fondée sur l'utilisation du filtre moyenne mobile.

Exercice n° 1.8

fichier C1EX4

Désaisonnalisation par la méthode de Census X11

On demande de désaisonnaliser par la méthode de Census X11, la série de production d'électricité en Australie, mensuellement sur 31 ans.

Solution

L'application de cette méthode (sur les vingt dernières années), selon un schéma multiplicatif, peut être utilisée directement (Fonction PROCS) à partir du logiciel Eviews. Nous ne pouvons pas présenter, dans le cadre restreint de ce manuel, l'ensemble des calculs. Seuls les graphiques de la série brute et CVS sont présentés (cf. figure 1.11).

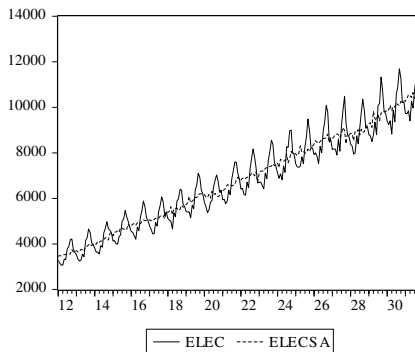


Figure 1.11 – Série de production d'électricité brute et CVS (méthode Census X11)

2) La démodulation complexe

C'est l'analyse des signaux physiques en radiodiffusion et en télécommunication. La démodulation est l'opération qui consiste à faire varier les caractéristiques de la fonction sinusoïdale, fonction qui représente la décomposition de Fourier d'un signal.

Nous pouvons résumer les principales étapes de la méthode¹ de la façon suivante.

Soit une chronique x_t , $t = 1, \dots, n$ qui présente une saisonnalité de période T_0 et x_t^1 la chronique obtenue de x_t en la multipliant par $e^{i\omega_0 t}$ où, $\omega_0 = \frac{2\pi}{T_0}$ est la fréquence de la saisonnalité :

$$x_t^1 = x_t e^{-i\omega_0 t} = x_t \cos(\omega_0 t) - i x_t \sin(\omega_0 t) = U_t - i V_t$$

Soit F un filtre passe-bande, c'est-à-dire un filtre qui conserve les fluctuations de basses fréquences (proche de zéro, comme par exemple la fréquence d'un extra-saisonnier) et qui élimine celles des moyennes et hautes fréquences.

On appelle démodulé complexe, noté $\tilde{x}_t^{(\omega_0)}$, l'application de ce filtre F à x_t^1 :

$$\begin{aligned} \tilde{x}_t^{(\omega_0)} &= F(x_t^1) = F(U_t) - i F(V_t) \\ &= z_t'' - i z_t' \end{aligned}$$

On caractérise le démodulé complexe d'un processus réel par son module (R_t) et par son argument (θ_t). Pour $\omega_0 \in [0; \pi]$ on a :

$$R_t = 2\sqrt{z_t'^2 + z_t''^2}$$

$$\theta_t = \text{Arctg} \left(-\frac{z_t'}{z_t''} \right)$$

R_t représente l'amplitude du démodulé complexe et θ_t sa phase.

On appelle remodulé complexe $x_t^{(\omega_0)}$ la série chronologique qui décrit l'évolution temporelle de la fluctuation de fréquence ω_0 , ici la saisonnalité. Elle est telle que :

$$x_t^{(\omega_0)} = \tilde{x}_t^{(\omega_0)} e^{i\omega_0 t} = (z_t'' - i z_t') (\cos \omega_0 t + i \sin \omega_0 t)$$

Pour un processus réel et pour $\omega_0 \in [0; \pi]$, nous avons :

$$x_t^{(\omega_0)} = 2z_t'' \cos \omega_0 t + 2z_t' \sin \omega_0 t$$

Ainsi, de manière pratique pour une série mensuelle, on calcule le remodulé pour la période $\frac{2\pi}{12}$ puis pour les périodes harmoniques $\frac{2\pi}{6}$, $\frac{2\pi}{4}$, $\frac{2\pi}{3}$, $\frac{2\pi}{2,4}$. La somme de tous ces remodulés constitue l'évolution de la saisonnalité. La chronique désaisonnalisée est alors : $x_t^{CVS} = x_t - \sum$ des remodulés complexes.

1. Mélard, 1975.

Cette méthode nécessite une forme additive ou une forme multiplicative complète de la structure de la chronique. Elle présente un intérêt certain puisqu'il est possible de calculer l'évolution de l'amplitude et de la phase du remodulé. Cette technique est cependant très peu employée pour désaisonnaliser les chroniques en raison de la nature du filtre F . En effet, il est indispensable de choisir un filtre « passe-bande » de type moyenne mobile. Or, la bande passante de ce filtre est d'autant plus étroite que la fréquence ω_0 est proche de 0 et donc que l'ordre de la moyenne mobile est élevé. Ceci entraîne une perte d'information importante au début et à la fin de la série.

Comme par ailleurs, la moyenne mobile est un filtre imparfait pour les hautes fréquences, on recourt à un compromis combinant deux moyennes mobiles (d'où le nom de double fenêtre attribué à cette méthode). Nous pouvons vérifier graphiquement, que pour une série mensuelle, la combinaison optimale est réalisée avec des moyennes mobiles¹ d'ordre 12 et 18 ; mais il y a une contradiction entre la finalité du résultat recherché et les longueurs des moyennes mobiles qui font « perdre » 15 observations aux extrémités de la chronique.

Exercice n° 1.9

fichier C1EX4

Désaisonnalisation par démodulation complexe

On demande de désaisonnaliser par la méthode de démodulation complexe, la série de production d'électricité en Australie, mensuelle sur 31 ans.

Solution

L'application de cette méthode, selon un schéma multiplicatif, peut être programmée à partir du logiciel EvIEWS à l'aide des instructions suivantes (C1EX9.PRG) :

```

GENR TREND = @trend
SCALAR PI = 3.14116
GENR LELEC = LOG(ELEC)'Schéma multiplicatif
GENR SY = 0
FOR !I = 1 TO 5
IF !I = 1 THEN SCALAR w0 = 12 ENDIF
IF !I = 2 THEN SCALAR w0 = 6 ENDIF
IF !I = 3 THEN SCALAR w0 = 4 ENDIF
IF !I = 4 THEN SCALAR w0 = 3 ENDIF
IF !I = 5 THEN SCALAR w0 = 2.4 ENDIF
GENR U = LELEC*COS(2*PI/w0*TREND)
GENR V = LELEC*SIN(2*PI/w0*TREND)
" CALCUL DES MOYENNES MOBILES D'ORDRE 12 ET 18
GENR F1 = @MOVAV(U,12)
GENR FU = @MOVAV(F1,18)

```

1. Terraza, 1981.

```

GENR F1 = @MOVAV(V,12)
GENR FV = @MOVAV(F1,18)
IF !I = 1 THEN GENR R = 2*sqr(FU^2+FV^2) ENDIF
GENR Y!I = 2*FU*COS(2*PI/w0*TREND)+2*FV*SIN(2*PI/w0*TREND)
GENR SY = SY+Y!I
DELETE Y!I
NEXT
" CALCUL DE LA SERIE CVS
GENR LELCVS = LELEC - SY
GENR ELECVS = EXP(LELCVS)

```

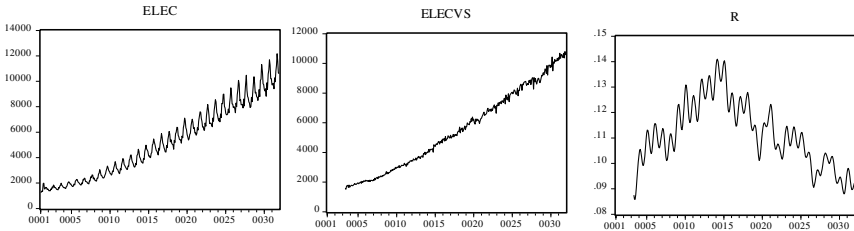


Figure 1.12 – Production d'électricité brute, CVS et module R_t pour $\omega_0 = \frac{2\pi}{12}$

3) La méthode TRAMO-SEATS

Cette technique fait partie des méthodes paramétriques fondées sur l'extraction du signal d'une chronique qui correspond à une de ses composantes. Elle repose essentiellement sur l'analyse spectrale (cf. chapitre 3) et s'apparente de ce fait à la démodulation complexe. La méthode TRAMO (« *Times series Regression with ARIMA noise, Missing observations and Outliers* ») est un logiciel qui modélise à l'aide d'un processus ARIMA (cf. chapitre 6) la série brute. Il est complété par le programme SEATS (« *Signal Extraction ARIMA Series* ») qui décompose la série et qui en extrait le signal choisi en utilisant l'analyse spectrale. Cette méthode¹ est à rapprocher de celle de Koopman² qui utilise le logiciel STAMP.

1. Cf. Gomez V., Maravall A., 1997.
2. Koopman S. J. et alii, 1995.

2. Prédiction d'une série chronologique

Dans ce chapitre nous examinons les méthodes traditionnelles de prédiction des séries chronologiques. Nous nous intéressons tout d'abord au cas d'une chronique dépourvue de saisonnalité et aux méthodes de prédiction susceptibles d'être utilisées : extrapolation de tendance, lissage exponentiel simple et double. Puis nous traitons le cas des séries chronologiques affectées d'un mouvement saisonnier, nous envisageons alors les méthodes de prédiction permettant l'extrapolation des différentes composantes et le modèle de Holt-Winters. Ces techniques, extrêmement simples à mettre en œuvre à partir d'un tableur, sont maintenant utilisées couramment en entreprise pour la prédiction des ventes, par exemple.

I. Prédiction d'une chronique non saisonnière

Si une chronique ne possède pas de saisonnalité ou a été désaisonnalisée, il est indispensable de tester la présence ou non d'une tendance. Les tests relatifs à cette détection utilisent la plupart des instruments déjà présentés pour l'existence de la saisonnalité, à savoir :

- la représentation graphique de la série brute ;
- l'analyse de la variance ;
- la fonction d'autocorrélation ;
- le spectre.

Le graphe du résidu de la chronique – hors tendance et hors saisonnalité – avec les hypothèses stochastiques classiques (espérance nulle ou constante, homoscedasticité, indépendance) se présente comme une série de nombres au hasard limitée en amplitude par deux bandes parallèles. On en déduit que toute autre forme du graphique est révélatrice de l'existence d'une extra-saisonnalité dans la chronique. L'utilisation de la décomposition de la variance (tableau de Buys-Ballot) conduit en présence d'une tendance à un effet ligne significatif qui s'élimine lorsque la série est transformée par les différences premières ou par une régression sur le temps (cf. chapitre 5).

Lorsqu'une tendance existe et présente une forme analytique assez prononcée, comme une tendance linéaire par exemple, la fonction d'autocorrélation de la chronique possède des valeurs élevées qui décroissent lentement¹. De même la fonction spectrale (cf. chapitre 4) se caractérise par des valeurs importantes aux très basses fréquences.

Ces instruments peuvent être nécessaires, mais ils sont insuffisants. En effet, la forme d'une fonction d'autocorrélation ou d'un spectre n'est pas toujours révélatrice d'une tendance qui néanmoins peut exister.

Cette recherche de tendance peut être complétée à partir de tests de détection d'une tendance.

A. Tests de détection d'une tendance

1) Le test du *Turning Point*

Pour une chronique x_t , un *Turning Point* est une observation telle que $(x_{t-1} < x_t$ et $x_t > x_{t+1})$ ou $(x_{t-1} > x_t$ et $x_t < x_{t+1})$. On appelle T le nombre de ces points de retournement. Si x_t est un bruit blanc alors la probabilité d'avoir un retournement au temps t est $2/3$, soit sur n observations $\frac{2(n-3)}{3}$ retournements. On peut donc démontrer que :

$$E[T] = \frac{2(n-3)}{3} \approx \mu_t$$

$$V[T] = \frac{16n-29}{90} \approx \sigma_T^2 \text{ et } T \rightarrow \mathcal{N}(E[T]; V[T])$$

Une valeur importante de $T - E[T]$ indique que la chronique fluctue beaucoup plus rapidement qu'un bruit blanc.

1. Les Anglo-Saxons parlent de l'absence de *cut-off*, c'est-à-dire de rupture brusque de la fonction d'autocorrélation.

Une valeur de $T - E[T]$ très négative indique l'existence d'une corrélation positive entre valeurs voisines.

On rejette l'hypothèse de bruit blanc au seuil de 5 % si : $\frac{|T - \mu_T|}{\sigma_T} > 1,96$.

2) Le test des signes

Ce test est fondé sur le nombre de fois où l'écart entre deux valeurs successives est positif, soit s cette fréquence. On démontre que :

$$E[s] = \frac{(n-1)}{2} \approx \mu_s$$

$$V[s] = \frac{n+1}{12} \approx \sigma_s^2 \text{ et } s \rightarrow \mathcal{N}(\mu_s; \sigma_s^2)$$

On peut vérifier qu'une valeur importante de $(s - \mu_s)$ indique la présence d'une tendance dans la chronique. On rejette alors l'hypothèse de bruit blanc au seuil de 5 % si $\frac{|s - \mu_s|}{\sigma_s} > 1,96$.

3) Le test de rang (Rank Test)

Ce test est souvent utilisé pour détecter des tendances linéaires dans les séries brutes. On appelle p le nombre de paires (i, j) tel que $x_j > x_i$ avec $j > i$ où

$$\begin{cases} i = 1, \dots, n-1 \\ j = 1, \dots, n-1 \end{cases}$$

On démontre que :

$$E[p] = \frac{n(n-1)}{4} \approx \mu_p$$

$$V[p] = \frac{n(n-1)(2n+5)}{8} \approx \sigma_p^2 \text{ et } p \rightarrow \mathcal{N}(\mu_p; \sigma_p^2)$$

Une valeur élevée de la différence $(p - \mu_p)$ indique la présence dans la chronique d'une tendance linéaire. L'hypothèse de bruit blanc est alors rejetée si

$\frac{|p - \mu_p|}{\sigma_p} > 1,96$ au seuil de 5 %.

Ces trois tests sont essentiellement utilisés pour détecter l'existence d'une tendance linéaire dans une chronique ; ils sont moins performants pour rechercher la présence de mouvements de long terme comme, en particulier, des cycles. Ils peuvent utilement compléter les méthodes traditionnelles de recherche d'une tendance dans une chronique non saisonnière (cf. partie I).

Si les tests décrits dans ce chapitre indiquent l'absence d'une tendance, cela signifie pour les méthodes classiques, que la chronique n'est structurée que par le résidu R_t . Ce résidu – par hypothèse – est une succession de nombres au hasard indépendants entre eux, donc sans structure et de moyenne nulle. La prévision de la chronique est donc la moyenne de ce résidu c'est-à-dire égale à 0 ou à une constante si la série n'est pas centrée.

En réalité, nous verrons ultérieurement que ce résidu peut avoir une structure sans qu'il possède une tendance. C'est par exemple une structure autorégressive « à mémoire », il est donc prévisible mais par les méthodes présentées lors de la partie 2.

B. Analyse par régression

Si une tendance existe suite aux tests effectués sur la série brute, l'utilisation correcte des techniques suppose que l'on puisse distinguer la tendance déterministe de la tendance aléatoire. La tendance déterministe est celle qui se présente sous une forme analytique pour l'ensemble des observations de la chronique. Elle peut être linéaire ou non.

Si la tendance est linéaire le modèle s'écrit :

$$x_t = at + b + \varepsilon_t$$

La technique des moindres carrés ordinaires (MCO) fournit les estimateurs de a et b respectivement notés : (\hat{a}, \hat{b}) ainsi que les nombreux tests associés (tests des paramètres, test des résidus, test d'adéquation...).

Si la tendance est non linéaire, nous pouvons distinguer deux cas :

- la tendance non linéaire est transformable en tendance linéaire : (fonction puissance, fonction logistique...) ; la méthode employée est celle des moindres carrés ordinaires sur les données transformées ;
- la tendance non linéaire est non transformable par anamorphose ; on recourt à des algorithmes d'optimisation non linéaire¹ afin d'estimer ses paramètres.

Ces méthodes postulent des paramètres constants dans le temps et la prévision porte plutôt le nom d'extrapolation de la chronique. Elle reste utile pour estimer les tendances « lourdes » des phénomènes économiques, « toutes choses étant égales par ailleurs »².

1. Ces algorithmes de recherche de racines dans les polynômes sont du type Newton-Raphson, Powell, Marquardt ou Nelder et Mead (cf. Himmelblau, 1972).
2. Un exemple d'extrapolation par les MCO est donné à l'exercice 2.3 de ce chapitre.

Exercice n° 2.1

Fichier C2EX1

Estimation d'un modèle de diffusion de type Logistique

Le Tableau 2.1 présente les ventes cumulées (en milliers d'unités) d'un CD audio d'un artiste à succès sur 19 semaines.

Tableau 2.1 – Ventes cumulées sur 19 semaines (milliers d'unités)

Semaines	Ventes
1	44,7
2	61,0
3	81,3
4	105,8
5	134,0
6	165,6
7	200,1
8	236,9
9	275,4
10	315,0
11	355,2
12	395,3
13	434,8
14	473,3
15	510,1
16	544,9
17	577,3
18	607,0
19	633,9

On demande d'estimer les ventes par un modèle de diffusion de type Logistique selon deux méthodes :

- en faisant l'hypothèse d'un seuil de saturation à 800 000 unités,
- en estimant l'ensemble des paramètres du modèle à l'aide du « solveur » du tableur Excel.

Solution

Le modèle logistique s'exprime de la manière suivante :

$$y_t = \frac{y_{\max}}{1 + br^t}$$

avec :

y_{\max} le seuil de saturation,

b et r deux paramètres caractéristiques du modèle ($b > 0$ et $-1 < r < 0$).

Les propriétés sont les suivantes :

Si $t \rightarrow -\infty$ alors $y_t \rightarrow 0$,

si $t \rightarrow \infty$ alors $y_t \rightarrow y_{\max}$.

Le point d'inflexion de la courbe est fixe, il est atteint lorsque les ventes cumulées représentent 50 % du seuil de saturation y_{\max} .

La représentation graphique est une courbe en « S ». La figure 2.1 présente l'évolution des ventes cumulées en fonction des semaines. Ce graphique peut suggérer une évolution selon un modèle de diffusion, nous serions alors à proximité du point d'inflexion.

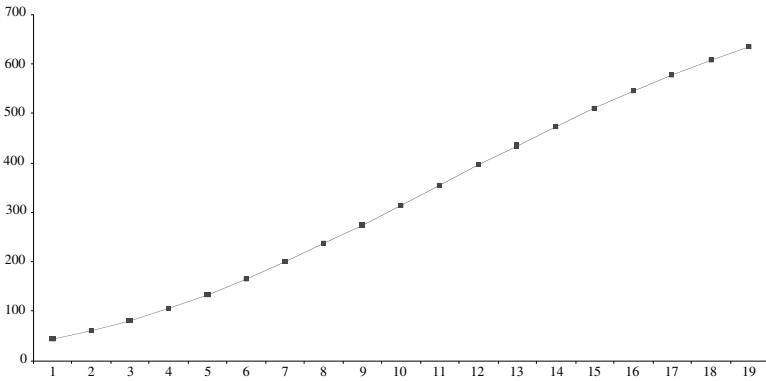


Figure 2.1 – Ventes cumulées

Les données correspondent visuellement à un modèle de type logistique dont nous allons estimer les paramètres. Par analogie avec un précédent CD du même artiste, on pense que les ventes maximums seront de 800 000 unités.

Pour procéder à l'estimation des paramètres avec $y_{\max} = 800$, nous allons linéariser la fonction : $y_t = \frac{y_{\max}}{1 + br^t}$.

$$\frac{y_{\max}}{y_t} = 1 + br^t \rightarrow \frac{y_{\max}}{y_t} - 1 = br^t \rightarrow \text{Ln} \left(\frac{y_{\max}}{y_t} - 1 \right) = \text{Ln}(b) + t \text{Ln}(r)$$

(Ln = logarithme népérien).

Soit à calculer la nouvelle variable à expliquer : $y_t = \text{Ln} \left(\frac{y_{\max}}{y_t} - 1 \right) = \text{Ln} \left(\frac{800}{y_t} - 1 \right)$

et le modèle linéaire s'écrit donc : $Y_t = a_0 + a_1 t + \varepsilon_t$ avec $b = e^{a_0}$ et $r = e^{a_1}$.

Les résultats de l'estimation de la régression sur le temps ($t = 1, 2, \dots, 19$) sont les suivants : $Y_t = 2,788 - 0,224t$

Le modèle estimé s'écrit : $y_t = \frac{y_{\max}}{1 + br^t} = \frac{800}{1 + 16,25 \times 0,7988^t}$

En effet : $\hat{b} = e^{\hat{a}_0} = e^{2,788} = 16,25$ et $\hat{r} = e^{\hat{a}_1} = e^{-0,224} = 0,7988$

La résolution analytique (cf. onglet solveur) donne les résultats suivants :

$$y_t = \frac{y_{\max}}{1 + br^t} = \frac{717}{1 + 14,72 \times 0,785^t}, \text{ le seuil de saturation est donc de 717 (avec somme des carrés des écarts = 699,34).}$$

C. Le lissage exponentiel

La technique du lissage exponentiel est utilisée dans le cas d'une chronique affectée d'une tendance aléatoire¹. Nous supposons, dans ce chapitre, que ce type de composante se présente lorsque son évolution n'est pas modélisable par une forme fonctionnelle.

Les techniques de lissage exponentiel ont été introduites par Holt en 1957 mais surtout par Brown en 1962.

Le lissage regroupe l'ensemble des techniques empiriques qui ont pour caractéristiques communes d'accorder un poids plus important aux valeurs récentes de la chronique. Ces méthodes portent aussi le nom de filtrage, car il s'agit d'une opération mathématique transformant un entrant x_t en une nouvelle chronique sortante y_t .

Les méthodes de lissage exponentiel non saisonnier supposent que la chronique de départ est structurée de la façon suivante :

$$x_t = f_t^{(k)} + \varepsilon_t$$

où : $E[\varepsilon_t] = 0$

$$E[\varepsilon_t^2] = \sigma_\varepsilon^2 \quad \forall t$$

$$E[\varepsilon_t \varepsilon_{t'}] = 0 \quad \forall t, \forall t', \forall t \neq t'$$

$f_t^{(k)}$ fonction polynomiale (de degré k) dont les paramètres dépendent du temps.

Les méthodes de lissage se différencient entre elles selon le degré de $f_t^{(k)}$; les degrés les plus utilisés sont $k = 0$ et $k = 1$.

- Ainsi, si $k = 0$, la série s'écrit :

$$x_t = a_t + \varepsilon_t \quad \forall t$$

avec $a_t = a$ constante si $t \in \{n_1; n_2\} \subset \{1; n\}$. L'ensemble $\{n_1; n_2\}$ porte le nom d'ensemble (ou intervalle) local. Cet intervalle peut correspondre au pas Δt de la chronique ou encore à la période de prévision. La méthode de lissage exponentiel adaptée à ce type de modèle est le lissage exponentiel simple (LES).

- Si $k = 1$ alors $x_t = a_t + b_t t + \varepsilon_t \quad \forall t$

avec :

$$a_t = a \text{ et } b_t = b \text{ si } t \in \{n_1; n_2\} \subset \{1; n\}$$

La méthode employée dans ce cas est le lissage exponentiel double (LED).

Lorsque k est quelconque, la méthode de lissage porte le nom de lissage exponentiel général, noté LEG.

1. Le chapitre 4 présente les tests statistiques permettant de distinguer les tendances déterministes des tendances aléatoires (ou stochastiques).

1) Le lissage exponentiel simple (LES)

a) Formulation générale

Supposons que x_t représente les ventes d'un produit quelconque en t . Ces ventes x_t peuvent être considérées comme le résultat d'une combinaison linéaire infinie de ses valeurs passées, le poids (ou l'influence) du passé sur le présent étant décroissant avec son ancienneté. Nous pouvons écrire la combinaison linéaire selon :

$$z_{t-1} = b_1 x_{t-1} + b_2 x_{t-2} + \dots + b_k x_{t-k} + \dots \quad \text{avec } b_1 > b_2 > \dots > b_k > \dots$$

Elle représente la valeur lissée de la chronique x_t calculée en $t - 1$. Par hypothèse et lorsque les paramètres sont estimés, cette valeur lissée peut être considérée comme la valeur prévue de x_t calculée en $t - 1$ pour t . Soit : $z_{t-1} = x_{t-1,t}^p = \widehat{x}_t$

Le problème est donc de déterminer les coefficients b_k conformément à l'hypothèse du lissage. Pour cela on suppose que la combinaison linéaire obéit à un modèle géométrique normalisé (modèle de Koyck) :

- il est géométrique car les poids de la combinaison linéaire décroissent exponentiellement ;
- il est normalisé car la somme des poids est égale à 1.

Pour que les poids (b_k) décroissent exponentiellement, on écrit :

$$b_k = b(1 - \lambda)^k \quad 0 \leq \lambda \leq 1 \quad k = 1, 2, \dots \quad \text{et } b_k \in \mathbb{R}$$

$$\text{D'où : } \sum_{k=1}^{\infty} b_k = b \sum_{k=1}^{\infty} (1 - \lambda)^k$$

Or $\sum_{k=1}^{\infty} (1 - \lambda)^k$ est la somme d'une progression géométrique de raison $(1 - \lambda)$

inférieure à 1, d'où $\sum_{k=1}^{\infty} b_k = b \frac{1 - \lambda}{\lambda}$, on constate donc que les poids ne sont pas normalisés.

Pour les normaliser on les divise par leur somme, tel que :

$$\omega_k = \frac{b_k}{\sum_{k=1}^{\infty} b_k} = \lambda(1 - \lambda)^{k-1} \Rightarrow \sum_{k=1}^{\infty} \omega_k = 1.$$

En remplaçant les poids ω_k dans le modèle géométrique infini :

$$\begin{aligned} z_{t-1} &= \widehat{x}_t = \lambda x_{t-1} + \lambda(1 - \lambda)x_{t-2} + \lambda(1 - \lambda)^2 x_{t-3} + \dots \\ &= \lambda \sum_{i=0}^{\infty} (1 - \lambda)^i x_{t-1-i} \end{aligned}$$

Cette formule peut aussi s'écrire¹ :

$$\hat{x}_t = \lambda x_{t-1} + (1 - \lambda)[\lambda x_{t-2} + \lambda(1 - \lambda)x_{t-3} + \dots] \text{ Soit :}$$

$$\hat{x}_t = \lambda x_{t-1} + (1 - \lambda)\hat{x}_{t-1}$$

Sous cette forme le lissage est une moyenne pondérée de la dernière réalisation et de la dernière valeur lissée.

$$\text{Ou encore : } \hat{x}_t = \hat{x}_{t-1} + \lambda(x_{t-1} - \hat{x}_{t-1})$$

Le lissage apparaît comme le résultat de la dernière valeur lissée corrigé par une pondération de l'écart entre la réalisation et la prévision.

Le paramètre λ , appelé la constante de lissage, joue un rôle important :

– lorsque λ est proche de 0, la pondération s'étale sur un grand nombre de termes du passé, la mémoire du phénomène étudié est forte et la prévision est peu réactive aux dernières observations ;

– lorsque λ est proche de 1, les observations les plus récentes ont un poids prépondérant sur les termes anciens, la mémoire du phénomène est faible et le lissage est très réactif aux dernières observations.

La figure 2.2 illustre le poids décroissant accordé au passé en fonction de différentes valeurs de λ .

À un instant donné de l'historique, le calcul de $\hat{x}_t = x_{t-1,t}^p$ est réalisable puisque tous les éléments de la formule sont connus. Pour la prévision de la chronique ($t = n + 1 \dots, n + h$, avec $n + h$ l'horizon de prévision du modèle) ce calcul ne peut être effectué que pour le premier pas du temps.

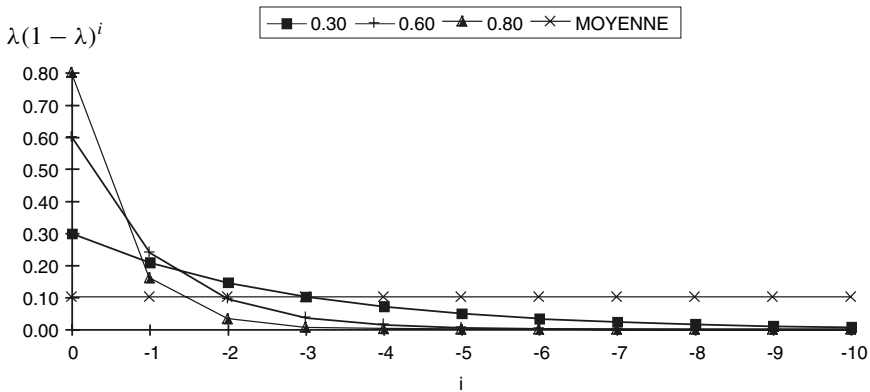


Figure 2.2 – Pondération décroissante de l'information avec son ancienneté : comparaison entre une moyenne classique (poids identique = 0,10) et trois valeurs de λ (poids géométriquement décroissant)

1. Dans certains ouvrages le lissé est z_t et la formule du LES s'écrit $\hat{x}_{t+1} = \lambda z_t + (1 - \lambda)\hat{x}_t$ ce qui harmonise l'ensemble des formules du lissage mais pose un problème d'écriture pour l'erreur de la prévision

En effet :

Pour $t = n + 1 \implies \hat{x}_{n+1} = \lambda x_n + (1 - \lambda)\hat{x}_n = x_{n,n+1}^p$ (λ , x_n , et \hat{x}_n sont connus).

En revanche, pour $t = n + 2 \implies \hat{x}_{n+2} = \lambda x_{n+1} + (1 - \lambda)\hat{x}_{n+1} = x_{n+1,n+2}^p$ la valeur de x_{n+1} est inconnue. La seule information dont nous disposons est sa prévision \hat{x}_{n+1} . En substituant cette valeur dans l'équation, on obtient $\hat{x}_{n+2} = \hat{x}_{n+1}$ et plus généralement $\hat{x}_{n+h} = \hat{x}_{n+1}$.

Les valeurs prévues à partir d'un LES sont donc identiques entre elles et égales à celle prévue en $n + 1$. Ce résultat est conforme à l'hypothèse de base du modèle $\hat{x}_t = a_t = a$ sur un intervalle local (celui de la prévision).

Il est possible de calculer le délai moyen de réaction – certains auteurs parlent d'âge moyen de l'information – qui est la moyenne pondérée des coefficients de lissage.

$$\bar{\lambda} = \frac{\sum_{i=0}^{\infty} i\lambda(1-\lambda)^i}{\sum_{i=0}^{\infty} \lambda(1-\lambda)^i} \quad \text{or} \quad \sum_{i=0}^{\infty} \lambda(1-\lambda)^i = 1 \quad (\text{par construction}).$$

$$\text{D'où :} \quad \bar{\lambda} = \sum_{i=0}^{\infty} i\lambda(1-\lambda)^i$$

$$\begin{aligned} \bar{\lambda} &= \lambda(1-\lambda) + 2\lambda(1-\lambda)^2 + 3\lambda(1-\lambda)^3 + \dots \\ &= \lambda[(1-\lambda) + 2(1-\lambda)^2 + 3(1-\lambda)^3 + \dots] \\ &= \lambda S \end{aligned}$$

Calculons l'expression : $S - (1-\lambda)S = \lambda S$

$$\begin{aligned} S &= (1-\lambda) + 2(1-\lambda)^2 + 3(1-\lambda)^3 + \dots \\ -(1-\lambda)S &= -(1-\lambda)^2 - 2(1-\lambda)^3 - 3(1-\lambda)^4 - \dots \end{aligned}$$

$$\begin{aligned} \lambda S &= (1-\lambda) + (1-\lambda)^2 + (1-\lambda)^3 + \dots \\ &= (1-\lambda) \times \frac{1}{1-(1-\lambda)} = \frac{1-\lambda}{\lambda} \end{aligned}$$

D'où :

$$\bar{\lambda} = \frac{1-\lambda}{\lambda}$$

Si $\lambda = 1$, l'âge moyen est nul puisque seule la dernière valeur est prise en compte ; si $\lambda = 0$ l'âge moyen est infini puisque seule la valeur initiale est prise en compte.

b) Utilisation du LES

• Choix de la constante λ

La valeur optimale du coefficient de lissage, à un instant donné, résulte d'un compromis entre l'inertie liée à l'intégration des données lointaines et la sensibilité aux valeurs récentes.

En cas d'erreur de prévision constatée, deux interprétations sont possibles :

- il s'agit d'un accident, le coefficient doit alors être diminué afin de gommer l'effet de cette valeur anormale pour le calcul des prochaines prévisions ;
- il s'agit d'une rupture de tendance durable, le coefficient doit être augmenté afin d'intégrer le plus rapidement cette nouvelle tendance pour le calcul de la prévision.

L'arbitrage est très délicat entre ces deux hypothèses¹.

Si nous étions tentés de choisir $\lambda = 1$ pour prendre en compte le passé immédiat et augmenter la réactivité, le rôle de filtrage des aléas que doit assurer le lissage serait inopérant.

$$\text{En effet, } \hat{x}_t = \lambda \sum_{i=0}^{\infty} (1 - \lambda)^i x_{t-1-i}$$

$$\text{Var}(\hat{x}_t) = \lambda^2 \text{Var} \left(\sum_{i=0}^{\infty} (1 - \lambda)^i x_{t-1-i} \right)$$

Or $x_t = a_t + \varepsilon_t$, d'où sur un intervalle local $\text{Var}(x_t) = \sigma_\varepsilon^2$ et $\forall t$ et $t' (t \neq t')$ les variables x_t et $x_{t'}$ sont indépendantes. D'où :

$$\begin{aligned} \text{Var}(\hat{x}_t) &= \lambda^2 \text{Var}(x_t) \sum_{i=0}^{\infty} (1 - \lambda)^{2i} \\ &= \lambda^2 \text{Var}(x_t) [1 + (1 - \lambda)^2 + (1 - \lambda)^4 + \dots] \\ &= \lambda^2 \text{Var}(x_t) \frac{1}{1 - (1 - \lambda)^2} \\ &= \lambda^2 \text{Var}(x_t) \frac{1}{1 - 1 + 2\lambda - \lambda^2} \\ &= \text{Var}(x_t) \frac{\lambda}{2 - \lambda} \end{aligned}$$

Nous avons donc :

$$\frac{\text{Var}(\hat{x}_t)}{\text{Var}(x_t)} = \frac{\lambda}{2 - \lambda}$$

1. Cf. Bourbonnais et Usunier, 2007, p. 70 à 78.

dans le cas où $\lambda = 1$, \hat{x}_t a la même variance¹ que x_t .

Pour assurer un rôle de filtrage efficace, il convient de choisir un λ faible, mais alors on perd en réactivité.

La figure 2.3 présente, en ordonnée les valeurs du rapport de l'écart type de la série lissée à l'écart type de la série brute $\left(\sqrt{\frac{\lambda}{2-\lambda}}\right)$ et, en abscisse, l'âge du lissage pour différentes valeurs de λ .

À la lecture de cette courbe, on comprend pourquoi la valeur de $\lambda = 0,3$ est très couramment choisie : lorsque le coefficient λ est inférieur à 0,20, l'âge du lissage s'accroît très vite pour un gain faible de filtrage ; au-delà de 0,4, la qualité du filtrage décroît très vite pour une faible réduction de l'âge moyen.

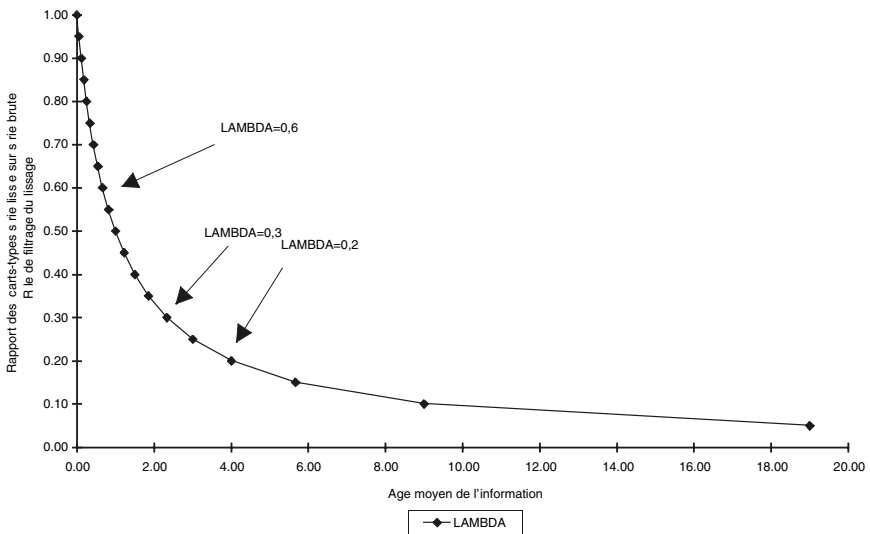


Figure 2.3 – Valeur du coefficient λ , âge moyen de l'information et filtrage

• Détermination de la constante de lissage λ

Nous ne citerons ici que deux approches de calculs permettant en partie de répondre au dilemme précédent.

– la valeur du coefficient de lissage est celle qui minimise la somme des car-

rés des erreurs de prévision passée : $\text{Min} \sum_{t=1}^n e_t^2 = \sum_{t=1}^n (x_t - \hat{x}_t)^2$: c'est la

1. Nous pouvons vérifier ce résultat à partir de la formule de lissage. Lorsque $\lambda = 1 \implies \hat{x}_t = x_{t-1,t}^p = x_{t-1}$ et $\text{Var}(x_{t-1}) = \text{Var}(x_t) = \text{Var}(\hat{x}_t) = \sigma_\varepsilon^2$. On parle dans ce cas de prévision naïve.

technique de calcul la plus couramment employée ; son principe est simple : pour l'intervalle donné (0; 1) des valeurs de λ , les prévisions sont simulées avec un « pas » assez fin (0,01 par exemple). Nous retenons la valeur de λ qui rend minimum la somme des carrés des écarts de prévision. L'utilisation d'un « solveur » sur les tableurs permet de résoudre ce problème. On peut aussi utiliser un algorithme d'optimisation de type dichotomique comme le font la plupart des logiciels ;

- procédure de régulation d'un coefficient de lissage : à partir d'un certain nombre d'indicateurs (amplitude de l'erreur de prévision, prévision biaisée...), une procédure de régulation du coefficient de lissage détermine sa valeur idéale à un instant donné. Contrairement à la méthode précédente, dans laquelle un coefficient optimal sur la totalité de la période historique est déterminé, cette approche permet d'estimer la valeur jugée la plus efficace à l'instant où la prévision est calculée.

• Initialisation des calculs

Dans la formule du LES $\hat{x}_t = \lambda x_{t-1} + (1 - \lambda)\hat{x}_{t-1}$

Pour $t = 1$: $\hat{x}_1 = \lambda x_0 + (1 - \lambda)\hat{x}_0$ on a x_0 et $\hat{x}_0 = x_{-1,0}^p$ inconnus.

Pour $t = 2$: $\hat{x}_2 = \lambda x_1 + (1 - \lambda)\hat{x}_1$ seul \hat{x}_1 est inconnu.

Plusieurs solutions existent pour démarrer les calculs à la période $t = 2$:

- prendre la moyenne arithmétique¹ des x_t ;
- poser $\hat{x}_1 = x_1$ Cette méthode est la plus simple et la plus utilisée² ;
- réaliser une prévision arrière de la chronique³, c'est-à-dire parcourir la chronique par le LES de la date n à la date 1 en utilisant comme démarrage des calculs la méthode précédente et faire la prévision pour la valeur 1.

• Choix du nombre d'observations en fonction de λ

Il est possible de déterminer le nombre d'observations en fonction de la valeur de λ . En effet, la pondération des observations décroît de manière géométrique jusqu'à devenir négligeable au delà d'un temps $t = K$. Soit A cet impact proche de 0. Par exemple, nous avons vu précédemment que la formule

du LES s'écrit : $\hat{x}_t = \lambda \sum_{i=0}^{\infty} (1 - \lambda)^i x_{t-1-i}$.

La somme des poids de ce lissage à partir de l'instant $t = K$ est égale à :

$$\sum_{i=K}^{\infty} \lambda (1 - \lambda)^i = \lambda \sum_{i=K}^{\infty} (1 - \lambda)^i = (1 - \lambda)^K$$

1. Brown, 1962 ; Montgomery et Johnson, 1976.
 2. Makridakis et Wheelwright, 1983.
 3. Abraham et Ledolter, 1983.

De ce fait nous avons : $(1 - \lambda)^K = A \rightarrow K \ln(1 - \lambda) = \ln A$ et $K = \frac{\ln A}{\ln(1 - \lambda)}$.

Par exemple, en finance, le système « *RISKMETRICS* » de J.P. Morgan utilise dans ses calculs la valeur $(1 - \lambda) = 0,94$ et $A = 1 \%$. Le nombre d'observations pour estimer les paramètres de leur modèle est alors de $K = \frac{\ln 0,01}{\ln 0,94} = 74$ jours.

• Intervalle de confiance de la prévision

Nous avons montré précédemment que : $\frac{\text{Var}(\widehat{x}_t)}{\text{Var}(x_t)} = \frac{\lambda}{2 - \lambda}$

Le résidu (erreur) de la prévision effectuée en $t - 1$ pour t est : $e_t = x_t - \widehat{x}_t$. D'où :

$\text{Var}(e_t) = \text{Var}(x_t) + \text{Var}(\widehat{x}_t)$ puisque x_t et \widehat{x}_t sont indépendants.

$$\text{D'où } \text{Var}(e_t) = \text{Var}(x_t) \left(1 + \frac{\lambda}{2 - \lambda} \right) = \text{Var}(x_t) \left(\frac{2}{2 - \lambda} \right) = \frac{2\sigma_x^2}{2 - \lambda}$$

L'intervalle de confiance à $100(1 - \alpha) \%$ de la valeur prévue de x_t en $t - 1$ (ou de $t = n + 1$ à $n + h$) s'écrit donc :

$$\text{Prob} \left\{ -t^{\alpha/2} \leq \frac{x_t - \widehat{x}_t}{\sigma_x \sqrt{\frac{2}{2 - \lambda}}} \leq t^{\alpha/2} \right\} = 100(1 - \alpha) \%$$

$$\implies \text{Prob} \left\{ x_t \in \left[\widehat{x}_t \pm t^{\alpha/2} \sigma_x \sqrt{\frac{2}{2 - \lambda}} \right] \right\} = 100(1 - \alpha) \%$$

avec $t^{\alpha/2}$ la valeur de loi normale centrée réduite. Pour un intervalle de confiance à 95 %, nous avons :

$$\text{IC} = \widehat{x}_t \pm 1,96\sigma_x \sqrt{\frac{2}{2 - \lambda}}$$

Exercice n° 2.2

Calcul d'une prévision par un LES

Soit les données observées de la série x_t consignées sur le Tableau 2.2. On demande de calculer une prévision par un LES – assortie de son intervalle de confiance à 95 % – à un horizon de trois périodes. La constante de lissage λ est égale 0,3.

Solution

Tableau 2.2 – Prévision à partir du modèle de lissage exponentiel simple

t	x_t	\hat{x}_t	$e_t = x_t - \hat{x}_t$
1	30	30	0
2	40	30,00	10,00
3	40	33,00	7,00
4	30	35,10	- 5,10
5	20	33,57	- 13,57
6	20	29,50	- 9,50
7	30	26,65	3,35
8	30	27,65	2,35
9		28,36	
10		28,36	
11		28,36	

On initialise $\hat{x}_1 = x_1 = 30$

Pour $t = 2$ (prévision calculée en $t = 1$ pour $t = 2$) nous avons :

$$\hat{x}_2 = x_{1,2}^p = 0,3 \times x_1 + 0,7 \times \hat{x}_1 = 30 = x_1$$

Pour $t = 3$: $\hat{x}_3 = 0,3 \times 40 + 0,7 \times 30 = 33$

...

Pour $t = 8$, $\hat{x}_8 = 0,3 \times 30 + 0,7 \times 26,65 = 27,65$

Pour $t = 9$, $\hat{x}_9 = 0,3 \times 30 + 0,7 \times 27,65 = 28,36$

Pour $t = 10, 11$ $\hat{x}_{10} = \hat{x}_{11} = \hat{x}_9 = 28,36$

Nous avons $\sigma_x = \sqrt{\frac{\sum(x_t - \bar{x})^2}{n - 1}} = 7,56$ (formule des petits échantillons) et

$$\sqrt{\frac{2}{2 - \lambda}} = 1,085$$

D'où : IC = $\hat{x}_t \pm 1,96\sigma_x \sqrt{\frac{2}{2 - \lambda}} = 28,36 \pm 1,96 \times 7,56 \times 1,085$

$$\text{IC} = 28,36 \pm 16,08 = [12,28; 44,44]$$

3) Le lissage exponentiel double (LED)

a) Formulation générale

Le modèle du lissage exponentiel double s'applique à une chronique du type :

$$x_t = a_t + b_t t + \varepsilon_t \text{ avec } a_t = a, b_t = b \forall t, t \in [n_1; n_2] \subset [1; n]$$

Comme son nom l'indique la technique du LED consiste à effectuer un lissage de la série déjà lissée. Supposons que l'on se situe sur un intervalle local et calculons la valeur lissée du modèle déterministe : $x_t = a + bt$

Nous avons :

$$\begin{aligned} z_{t-1} &= \lambda x_{t-1} + \lambda(1-\lambda)x_{t-2} + \lambda(1-\lambda)^2 x_{t-3} + \dots \\ &= \lambda \sum_{i=0}^{\infty} (1-\lambda)^i (a + b(t-1-i)) \\ &= \lambda a \sum_{i=0}^{\infty} (1-\lambda)^i + \lambda b t \sum_{i=0}^{\infty} (1-\lambda)^i - \lambda b \sum_{i=0}^{\infty} (1-\lambda)^i (1+i) \\ &= \lambda a \frac{1}{\lambda} + \lambda b t \frac{1}{\lambda} - \lambda b [1 + 2(1-\lambda) + 3(1-\lambda)^2 + \dots] \end{aligned}$$

$$\begin{aligned} \text{Or } 1 + 2(1-\lambda) + 3(1-\lambda)^2 + \dots &= \frac{1}{\lambda} + \frac{(1-\lambda)}{\lambda} + \frac{(1-\lambda)^2}{\lambda} + \dots \\ &= \frac{1}{\lambda} [1 + (1-\lambda) + (1-\lambda)^2 + \dots] \\ &= \frac{1}{\lambda^2} \end{aligned}$$

$$\text{D'où } z_{t-1} = a + bt - b \frac{1}{\lambda} = a + b(t-1+1) - b \frac{1}{\lambda}$$

$$z_{t-1} = a + b(t-1) - b \frac{1-\lambda}{\lambda}$$

$$z_{t-1} = a + b(t-1) - b\bar{\lambda}$$

$$\iff z_{t-1} = x_{t-1} - b\bar{\lambda} \iff (x_{t-1} - z_{t-1}) = b\bar{\lambda}$$

Ce résultat montre qu'il y a un biais égal à $b\bar{\lambda}$ entre la valeur observée et la valeur lissée sur un intervalle local quand le modèle est linéaire. Le LES est donc une méthode inadaptée pour prévoir les séries brutes localement linéaires ; ce résultat va nous permettre d'établir le LED.

On appelle zz_t le lissé de z_t et on effectue la même démonstration que précédemment puisque z_{t-1} est encore une forme linéaire : $z_{t-1} = (a - b\bar{\lambda}) + b(t-1)$.

On obtient :

$$\begin{aligned} zz_{t-1} &= (a - b\bar{\lambda}) + b(t-1) - b\bar{\lambda} \\ &= (a - 2b\bar{\lambda}) + b(t-1) \end{aligned}$$

La figure 2.4 illustre ce résultat.

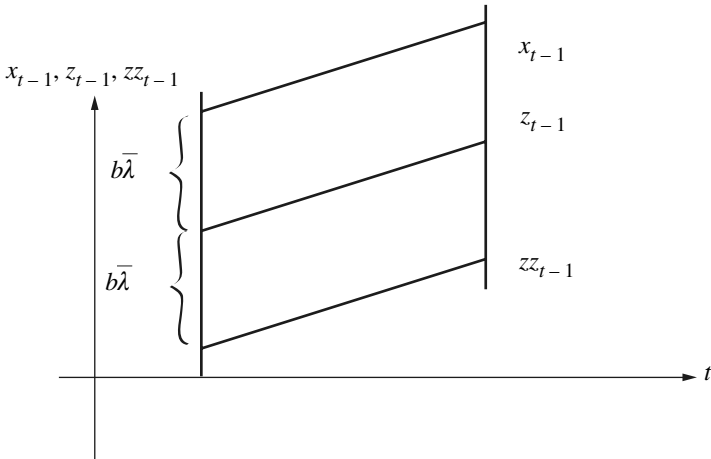


Figure 2.4 – Écart entre série brute, le LES et le LED en $t - 1$

Les séries lissées z_{t-1} et zz_{t-1} présentent donc un écart $b\bar{\lambda}$ sur l'intervalle local. Nous pouvons donc écrire : $x_{t-1} - z_{t-1} = z_{t-1} - zz_{t-1} = b\bar{\lambda}$

Cet intervalle local peut s'interpréter comme un pas du temps. Chacun des points de la chronique x_t est le départ d'une droite $a + bt$ et peut être considéré comme son ordonnée à l'origine. Les formules précédentes peuvent donc s'écrire quel que soit t :

$$x_{t-1} = 2z_{t-1} - zz_{t-1}.$$

Soit : $a_{t-1} = 2z_{t-1} - zz_{t-1}$

De même : $b_{t-1} = \frac{1}{\lambda}(z_{t-1} - zz_{t-1})$.

Dans ces expressions z_{t-1} et zz_{t-1} sont des valeurs lissées ; les résultats des calculs sont donc affectés à l'instant $t - 1$. Dans ces conditions la valeur trouvée par le LED n'est pas une valeur prévisionnelle. On choisit donc, comme pour le LES :

$$z_{t-1} = \widehat{x}_t = \lambda x_{t-1} + (1 - \lambda)\widehat{x}_{t-1}$$

et

$$\begin{aligned} zz_{t-1} &= \widehat{\widehat{x}}_t = \lambda z_{t-1} + \lambda(1 - \lambda)z_{t-2} + \lambda(1 - \lambda)^2 z_{t-3} + \dots \\ &= \lambda z_{t-1} + (1 - \lambda)zz_{t-2} = \lambda \widehat{x}_t + (1 - \lambda)\widehat{\widehat{x}}_{t-1} \end{aligned}$$

Les formules générales du LED sont donc les suivantes :

$$\begin{aligned}
\widehat{x}_t &= \lambda x_{t-1} + (1 - \lambda) \widehat{x}_{t-1} \\
\widehat{\widehat{x}}_t &= \lambda \widehat{x}_t + (1 - \lambda) \widehat{\widehat{x}}_{t-1} \\
a_t &= 2\widehat{x}_t - \widehat{\widehat{x}}_t \\
b_t &= \frac{1}{\lambda} (\widehat{\widehat{x}}_t - \widehat{\widehat{x}}_t); \quad \bar{\lambda} = \frac{1 - \lambda}{\lambda} \\
x_{t-1,t}^p &= a_t + b_t \quad (t = 2, \dots, n + 1) \\
x_{n,n+h}^p &= a_{n+1} + b_{n+1} \times h \quad (h = 2, 3, \dots)
\end{aligned}$$

La prévision par un LED est donc une droite qui a pour ordonnée à l'origine et pour pente les dernières valeurs trouvées de a_t et b_t . On peut, par conséquent, extrapoler cette droite à l'horizon h ; mais il convient de limiter cet horizon de prévision car la qualité de la prévision se dégrade rapidement avec l'accroissement de h .

b) Utilisation du LED

• Initialisation

En $t = 1$ nous avons :

$$\begin{cases} \widehat{x}_1 = \lambda x_0 + (1 - \lambda) \widehat{x}_0 \\ \widehat{\widehat{x}}_1 = \lambda \widehat{x}_1 + (1 - \lambda) \widehat{\widehat{x}}_0 \end{cases}$$

Comme x_0 , \widehat{x}_0 , \widehat{x}_1 et $\widehat{\widehat{x}}_0$ sont inconnus, les calculs débutent en $t = 2$ où seuls \widehat{x}_1 et $\widehat{\widehat{x}}_1$ ne sont pas déterminés.

Il existe différentes procédures d'initialisation pour le LED :

- la plus simple consiste à choisir : $\widehat{x}_1 = x_1$ et $\widehat{\widehat{x}}_1 = \widehat{x}_1$ soit $\widehat{x}_1 = \widehat{\widehat{x}}_1 = x_1$;
- une autre approche cherche à déduire \widehat{x}_1 et $\widehat{\widehat{x}}_1$ en utilisant les formules a_t et b_t pour $t = 1$:

$$t = 1 \quad \begin{cases} a_1 = 2x_1 - \widehat{x}_1 \\ b_1 = \frac{1}{\lambda} (\widehat{\widehat{x}}_1 - \widehat{x}_1) \end{cases} \implies \begin{cases} \widehat{x}_1 = a_1 - \bar{\lambda} b_1 \\ \widehat{\widehat{x}}_1 = a_1 - 2\bar{\lambda} b_1 \end{cases}$$

Comme a_1 et b_1 sont l'ordonnée à l'origine et la pente de $x_t = a_1 + b_1 t + \varepsilon_t$, nous pouvons les estimer par les MCO à partir de l'historique de la chronique x_t pour $t = 1 \dots n$. Néanmoins il convient de noter que cette procédure est en contradiction avec le LED qui s'applique sur une chronique localement linéaire.

- Choix de la constante λ : les méthodes présentées pour le LES sont employées de manière identique pour le LED.
- Construction d'un intervalle de confiance de la prévision

L'intervalle de confiance¹ de la prévision x_{n+h} de x_t notée $x_{n,n+h}^P$ (ou sur un intervalle local) est calculée à partir de la variance de l'erreur de prévision donnée par :

$$\text{Var}(e_{n,n+h}) = \sigma_x^2 C_h^2$$

avec $\sigma = \text{Var}(x_t)$, pour $t = 1 \dots, n$ et

$$C_h^2 = 1 + \frac{1 - \lambda'}{(1 + \lambda')^3} \left[(1 + 4\lambda' + 5\lambda'^2) + 2h(1 - \lambda')(1 + 3\lambda') + 2h^2(1 - \lambda')^2 \right]$$

avec $\lambda' = 1 - \lambda$

L'intervalle de confiance à $100(1 - \alpha) \%$ de la valeur prévue de x_t en $t - 1$ (ou de $t = n + 1$ à $n + h$) s'écrit donc :

$$\text{Prob} \left\{ -t^{\alpha/2} \leq \frac{x_{n+h} - x_{n,n+h}^P}{\sigma_x C_h} \leq t^{\alpha/2} \right\} = 100(1 - \alpha) \%$$

$$\implies \text{Prob} \{ x_{t+h} \in [x_{n,n+h}^P \pm t^{\alpha/2} \sigma_x C_h] \} = 100(1 - \alpha) \%$$

avec $t^{\alpha/2}$ la valeur de loi normale centrée réduite. Pour un intervalle de confiance à 95 %, nous avons :

$$\text{IC} = x_{n,n+h}^P \pm 1,96 \sigma_x C_h$$

Cet intervalle de confiance dépend de l'horizon de prévision h contrairement au LES.

Exercice n° 2.3

Prévision d'une chronique à partir d'un LED

Soit les données observées de la série x_t consignées sur le Tableau 2.3. On demande de calculer une prévision par une méthode de lissage double (LED) à un horizon de trois périodes. La constante de lissage λ est égale 0,5.

Solution

Exemples de calculs :

$$\text{a) } \hat{x}_1 = x_1 = 10$$

$$\hat{x}_2 = 0,5x_1 + 0,5\hat{x}_1 = 0,5 \times 10 + 0,5 \times 10 = 10$$

$$\hat{x}_3 = 0,5x_2 + 0,5\hat{x}_2 = 0,5 \times 20 + 0,5 \times 10 = 15$$

1. Abraham et Ledolter, 1983, p. 128-129.

Tableau 2.3 – Prédiction à partir du modèle de lissage exponentiel double

t	x_t	\hat{x}_t	$\hat{\hat{x}}_t$	a_t	b_t	$x_{t-1,t}^P$	e_t
1	10	10	10				
2	20	10,00	10,00	10,00	0,00	10,00	10,00
3	20	15,00	12,50	17,50	2,50	20,00	0,00
4	30	17,50	15,00	20,00	2,50	22,50	7,50
5	40	23,75	19,38	28,12	4,37	32,49	7,51
6	40	31,88	25,63	38,13	6,25	44,38	- 4,38
7	50	35,94	30,79	41,09	5,15	46,24	3,76
8	50	42,97	36,88	49,06	6,09	55,15	- 5,15
9		46,49	41,69	51,29	4,80	56,09	
10						60,89	
11						65,69	
12						70,49	

...

$$\hat{x}_9 = 0,5x_8 + 0,5\hat{x}_8 = 0,5 \times 50 + 0,5 \times 42,97 = 46,49$$

b) $\hat{\hat{x}}_1 = \hat{x}_1 = x_1 = 10$

$$\hat{\hat{x}}_2 = 0,5\hat{x}_2 + 0,5\hat{\hat{x}}_1 = 0,5 \times 10 + 0,5 \times 10 = 10$$

$$\hat{\hat{x}}_3 = 0,5\hat{x}_3 + 0,5\hat{\hat{x}}_2 = 0,5 \times 15 + 0,5 \times 10 = 12,5$$

...

$$\hat{\hat{x}}_9 = 0,5\hat{x}_9 + 0,5\hat{\hat{x}}_8 = 0,5 \times 46,49 + 0,5 \times 36,88 = 41,69$$

c) $\bar{\lambda} = \frac{1 - 0,5}{0,5} = 1$

d) $x_{8,9}^P = a_9 + b_9 = 51,29 + 4,80 = 56,09$

e) $x_{8,10}^P = a_9 + b_9 \times 2 = 51,29 + 4,80 \times 2 = x_{8,9}^P + b_9 = 56,09 + 4,80 = 60,89$

$$x_{8,11}^P = a_9 + b_9 \times 3 = 51,29 + 4,80 \times 3 = 60,89 + 4,80 = 65,69$$

$$x_{8,12}^P = a_9 + b_9 \times 4 = 51,29 + 4,80 \times 4 = 65,69 + 4,80 = 70,49$$

f) $\text{Var}(x_t) = \sigma_{\bar{x}}^2 = 193,75 \Rightarrow \sigma_x = 13,92$

$$C_1^2 = 2,07 ; C_2^2 = 2,67 ; C_3^2 = 3,41 ; C_4^2 = 4,30$$

$$C_1 = 1,44 ; C_2 = 1,63 ; C_3 = 1,85 ; C_4 = 2,07$$

$$\text{IC}_1 = 56,09 \pm 1,96 \times 1,44 \times 13,92 = [16,81; 95,38]$$

$$\text{IC}_2 = 60,89 \pm 1,96 \times 1,63 \times 13,92 = [16,43; 105,37]$$

$$\text{IC}_3 = 65,69 \pm 1,96 \times 1,85 \times 13,92 = [15,23; 116,18]$$

$$\text{IC}_4 = 70,49 \pm 1,96 \times 2,07 \times 13,92 = [14,03; 126,98]$$

c) Le modèle avec tendance de Holt

La méthode que nous venons de présenter est celle de Brown. Nous pouvons aussi utiliser le lissage de Holt qui comprend deux paramètres : un pour l'ordonnée à l'origine a_t et l'autre pour la pente b_t .

Deux lissages distincts sont effectués :

- le lissage de la moyenne a avec un coefficient de lissage α , $\alpha \in [0; 1]$;
- le lissage de la tendance b avec un coefficient de lissage β , $\beta \in [0; 1]$.

On peut démontrer qu'il existe une correspondance entre le coefficient de lissage de Brown et les deux coefficients de Holt¹.

• Formulation

Lissage de la moyenne : $a_t = \alpha x_t + (1 - \alpha)(a_{t-1} + b_{t-1})$

Lissage de la tendance : $b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1}$

Prévision calculée en t à un horizon de h périodes : $\hat{x}_{t+h} = a_t + hb_t$ avec :

x_t = valeur observée de la série en t

a_t = moyenne lissée de la série en t

b_t = tendance estimée en t .

• Initialisation (pour $t = 1$)

– initialisation de la moyenne lissée : $a_1 = x_1$;

– initialisation de la tendance : $b_1 = 0$.

Les formules générales peuvent ensuite être utilisées. Un exemple complet de calcul est présenté, par la suite, à partir du modèle de Holt-Winters.

4) Le lissage exponentiel généralisé (LEG)

Le degré k du polynôme $f_1^{(k)}$ peut être augmenté ce qui permet de définir des lissages exponentiels de degré supérieur à 1. Il en est ainsi du LET (lissage exponentiel triple) pour lequel $k = 2$. D'une manière plus générale, on peut définir un lissage exponentiel général non saisonnier noté LEG qui s'applique à une chronique du type :

$$x_t = f(t, \beta) + \varepsilon_t$$

où t = temps

β = vecteur $(k + 1, 1)$ des paramètres inconnus

f = une fonction dépendant des paramètres et du temps

1. Si $\alpha = \alpha^B(2 - \alpha^B)$ et si $\beta = \alpha^B/(2 - \alpha^B)$ avec α^B = coefficient de lissage double de Brown, le modèle de Holt est équivalent au lissage exponentiel double de Brown.

Ce modèle s'écrit sous la forme matricielle suivante :

$$x_t = f'(t)\beta + \varepsilon_t$$

$f'(t) = [f_0(t), \dots, f_i(t), \dots, f_k(t)]$ transposé du vecteur $f(t)$ de dimension $(k + 1, 1)$.

Vecteur de dimension $(k + 1, 1)$

Dans cette écriture, le temps est une variable dépendante, $f'(t)$ est le vecteur des fonctions d'ajustement et il vérifie l'écriture suivante : $f'(t + 1) = L \cdot f'(t)$

$L_{(k+1, k+1)}$ est appelée matrice de passage.

k est le degré du polynôme qui doit s'ajuster aux données :

$$f^{(k)}(t) = \frac{t^k}{k!} \quad \forall t, \forall k$$

L est une matrice triangulaire inférieure qui a pour élément générateur l_{ij} . Il est tel que :

$$l_{ij} = 0 \quad \text{si } i < j$$

$$l_{ij} = \frac{1}{(i - j)!} \quad \text{si } i \geq j$$

Exemples d'écriture :

- si $k = 0$ $x_t = \beta_0 + \varepsilon_t$
 $f_0(t) = 1$ $f'(t) = 1$ $L = 1$.

- si $k = 1$ $x_t = \beta_0 + \beta_1 t + \varepsilon_t$
 $f_0(t) = 1$ $f_1(t) = t$ $f'(t) = [1; t]$ $L_{(2,2)} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

Avec cette écriture très générale, les valeurs du LEG de Brown sont obtenues à partir de :

$$\text{Min} \sum_{t=0}^{n-1} (1 - \alpha)^t [x_{n-t} - f'(t)\beta]^2$$

$f'(t)\beta$: prévision de x_{n-1}

Lorsque les paramètres sont connus, nous pouvons résoudre cette équation et obtenir les estimations suivantes pour $t = n$:

$$\hat{\beta}_n = F^{-1}k_n$$

$$F_n = \sum_{t=0}^{n-1} (1 - \alpha)^t f(-t)f'(-t)$$

$$k_n = \sum_{t=0}^{n-1} (1 - \alpha)^t f(-t)x_{n-t}$$

La prévision calculée en $t = n$ pour la date $t = n + h$ est alors :

$$x_{n,n+h}^P = f'(h)\hat{\beta}_n$$

Il est possible – comme pour les autres méthodes – de construire des intervalles de confiance prévisionnels. On démontre que :

$$\text{Var}(e_{n,n+h}) = \sigma_x^2[1 + f'(h)F_n^{-1} \times F_n^{-1}f(h)] = \sigma_x^2 C_h^2$$

avec $\sigma_x^2 = \text{Var}(x_t)$, pour $t = 1 \dots, n$ et $e_{n,n+h}$ les erreurs de prévisions.

L'intervalle de confiance à $100(1 - \alpha) \%$ de la valeur prévue de x_t en $t - 1$ (ou de $t = n + 1$ à $n + h$) s'écrit donc :

$$\text{Prob} \left\{ -t^{\alpha/2} \leq \frac{x_{n+h} - x_{n,n+h}^P}{\sigma_x C_h} \leq t^{\alpha/2} \right\} = 100(1 - \alpha) \%$$

$$\implies \text{Prob} \{ x_{t+h} \in [x_{n,n+h}^P \pm t^{\alpha/2} \sigma_x C_h] \} = 100(1 - \alpha) \%$$

avec $t^{\alpha/2}$ la valeur de loi normale centrée réduite. Pour un intervalle de confiance à 95 %, nous avons :

$$\text{IC} = \hat{x}_t \pm 1,96\sigma_x C_h$$

Cet intervalle de confiance dépend de l'horizon de prévision h .

L'intérêt du lissage exponentiel réside dans sa très grande facilité de mise en œuvre et dans la restitution des résultats aisément compréhensibles et donc maîtrisables.

Cependant, il faut noter que cette technique ne peut être employée que sur des historiques sans saisonnalité ou préalablement désaisonnalisés.

Nous allons aborder maintenant des méthodes de prévision traitant simultanément les composantes extra-saisonnaire et saisonnière d'une série chronologique.

II. Prévision d'une chronique saisonnière

En toute rigueur, une chronique saisonnière doit être traitée selon le schéma précédent à savoir distinguer le cas d'une saisonnalité déterministe du cas d'une saisonnalité aléatoire.

Les tests qui permettent cette distinction sont récents. Ils sont présentés dans la seconde partie de l'ouvrage. La saisonnalité déterministe est donc, ici, celle

qui se présente sous la forme d'une cyclicité très marquée régulière en période et en amplitude (selon le schéma) ; tous les autres cas seront considérés comme des saisonnalités aléatoires.

A. Analyse par régression

Lorsque la saisonnalité est considérée comme déterministe, la méthode à utiliser pour la modéliser est la méthode des moindres carrés ordinaires (MCO). Si on suppose que la chronique est structurée selon une composante extra-saisonnière et une saisonnalité de type déterministe (simultanément) alors on peut employer les techniques de la régression sur fonction trigonométrique et la régression sur variables dichotomiques pour modéliser cette structure. L'analyse du MLGS fournit tous les instruments permettant de prévoir de façon efficace ce type de série brute : tests de paramètres, tests du R^2 , tests de colinéarité, tests des résidus, etc.

Il peut arriver que la tendance de la chronique soit de type aléatoire alors que la saisonnalité reste déterministe ; c'est sur ce constat que Harrison a construit son modèle de prévision qui combine la technique du LED de Holt (lissage à deux paramètres) avec la régression sur fonction trigonométrique pour la partie saisonnière.

B. Utilisation des coefficients saisonniers

Les coefficients saisonniers calculés par les méthodes présentées au chapitre 1 représentent le profil saisonnier moyen de la chronique. En faisant l'hypothèse que la saisonnalité se reproduit à l'identique, ils peuvent être utilisés pour prévoir les valeurs de cette composante. La chronique CVS est alors modélisée par une fonction cinématique¹ ou par une technique de lissage exponentiel. La prévision de la chronique est ensuite obtenue par agrégation des différentes composantes en fonction du schéma de décomposition retenu.

1. Sauf dans le cas de la désaisonnalisation par la différence ou le rapport à la tendance qui suppose le choix, *a priori*, de cette fonction.

Exercice n° 2.4

Fichier C1EX1

Prévision par agrégation des composantes extra-saisonnalité et saisonnalité

En reprenant les données du tableau 1.1 du chapitre 1 et la série désaisonnalisée selon la technique des moyennes mobiles (cf. exercice 1.7 du chapitre 1), on demande de calculer une prévision pour la quatrième année re-saisonnalisée selon trois méthodes :

- une extrapolation d'une droite de tendance,
- une prévision par un LES ($\lambda = 0,3$),
- une prévision par un LED ($\lambda = 0,3$).

Solutions

a) Calcul de la tendance (T_t)

À partir des données CVS calculées au chapitre 1 par la méthode des moyennes mobiles, nous pouvons utiliser la fonction « DROITREG » du tableur Excel qui permet de calculer l'ensemble des paramètres relatif à une tendance. Les résultats figurent sur le Tableau 2.4.

Tableau 2.4 – Résultats de la régression sur le temps et correspondance avec les statistiques usuelles

- 1,18	1 655,34	\hat{a}_1	\hat{a}_0
11,97	88,14	$\hat{\sigma}_{\hat{a}_1}$	$\hat{\sigma}_{\hat{a}_0}$
0,000972	143,219	R^2	$\hat{\sigma}_\varepsilon$
0,00973	10	F^*	ddl
199,727	205 108,93	SCE	SCR

Le modèle d'extrapolation s'écrit donc :

$$x_t^{CVS} = 1655,34 - 1,18t + e_t$$

$R^2 = 0,000972$ (La droite de régression n'explique que 0,09 % de la variance de x_t^{CVS}).

Validation de l'estimation de la régression :

- Tests de significativité du coefficient de régression et du coefficient de corrélation

$t_{\hat{a}_1}^* = |-1,18|/11,97 = 0,099 < t_{10}^{0,05} = 2,23$. On accepte l'hypothèse H_0 , le coefficient de la pente de la droite n'est pas significativement différent de 0.

Dans le cas d'un modèle de régression simple, nous avons l'équivalence de ce test de

Student avec celui du Fisher entre $F^* = \frac{R^2/1}{(1-R^2)/(n-2)} = \frac{SCE/1}{SCR/(n-2)} =$

$$0,00973 = (t_{\hat{a}_1}^*)^2$$

$$F^* = 0,00973 < F_{1;10}^{0,05} = 4,96,$$

on accepte l'hypothèse H_0 , le coefficient de détermination R^2 n'est pas significativement différent de 0.

Le tableau 2.5 présente l'ensemble des résultats avec :

- colonne (4), $T_t = 1655,34 - 1,18t$ pour $t = 1 \dots, 16$.
- colonne (5), e_t les résidus d'estimation ($x_t^{CVS} - T_t$)
- colonne (6), S_t les coefficients saisonniers calculés à l'exercice 5 du chapitre 1.
- colonne (7), x_t^P la prévision re-saisonnalisée soit $T_t + S_t$

Le test de Durbin et Watson d'autocorrélation des erreurs ne peut être mené du fait du nombre trop faible d'observations, cependant la lecture des résidus e_t - colonne (5) - semble indiquer une autocorrélation visible par les séquences de résidus positifs et négatifs. En définitive, la tendance linéaire estimée sur la chronique désaisonnalisée n'est pas significative, son utilisation génère même une autocorrélation des résidus. La prévision de la chronique se résume donc à la valeur moyenne des ventes CVS sur les trois ans - moyenne de la colonne (3) $\bar{x}^{CVS} = 1\ 647,67$ - à laquelle on ajoute la saisonnalité ; soit $x_t^{PMoy} = \bar{x}^{CVS} + S_t$ calculé à la colonne (9).

Les résidus entre la série brute et les prévisions par extrapolation de tendance et saisonnalité sont calculés à la colonne (8) et ceux entre la série brute et les prévisions

Tableau 2.5 - Prévision par extrapolation de la tendance et ajout de la saisonnalité

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
t	x_t	x_t^{CVS}	T_t	e_t	S_t	x_t^{PTen}	e_t^{Ten}	x_t^{PMoy}	e_t^{Moy}
1	1 248	1 841	1 654,17	1 86,83	- 592,70	1 061	186,83	1 055	193,33
2	1 392	1 745	1 652,98	92,02	- 352,72	1 300	91,84	1 295	97,16
3	1 057	1 565	1 651,80	- 86,80	- 508,51	1 143	- 86,69	1 139	- 82,56
4	3 159	1 705	1 650,62	54,38	1 453,93	3 105	54,55	3 102	57,51
5	891	1 484	1 649,44	- 165,44	- 592,70	1 057	- 165,94	1 055	- 164,17
6	1 065	1 418	1 648,26	- 230,26	- 352,72	1 296	- 230,24	1 295	- 229,64
7	1 118	1 626	1 647,08	- 21,08	- 508,51	1 139	- 20,97	1 139	- 21,56
8	2 934	1 480	1 645,89	- 165,89	1 453,93	3 100	- 165,62	3 102	- 167,39
9	1 138	1 731	1 644,71	86,29	- 592,70	1 052	86,18	1 055	83,23
10	1 456	1 809	1 643,53	165,47	- 352,72	1 291	165,19	1 295	161,06
11	1 224	1 732	1 642,35	89,65	- 508,51	1 134	90,46	1 139	85,14
12	3 090	1 636	1 641,17	- 5,17	1 453,93	3 095	- 4,89	3 102	- 11,39
13			1 639,98		- 592,70	1 047		1 055	
14			1 638,80		- 352,72	1 286		1 295	
15			1 637,62		- 508,51	1 129		1 139	
16			1 636,44		1 453,93	3 090		3 102	

par moyenne et saisonnalité sont calculés à la colonne (10). Nous pouvons remarquer que la somme des carrés des résidus du modèle « naïf » moyenne et saisonnalité

$$SCR^{Moy} = \sum_{t=1}^{12} (e^{Moy})_t^2 = 205\,357,68 \text{ est à peine supérieure à celle du modèle par extra-}$$

$$\text{polation de tendance } SCR^{LES} = \sum_{t=1}^{11} e_t^2 = 205\,170,05.$$

L'indicateur $MSE = \frac{SCR}{n}$ (« Mean Squared Error ») est égal à :

$$MSE^{Moy} = \frac{205\,357,68}{12} = 17\,113,14 \quad MSE^{Ten} = \frac{205\,170,05}{11} = 17\,097,50$$

b) Prédiction par un LES ($\lambda = 0,3$)

L'application du lissage simple à la série lissée \hat{x}_t^{CVS} (tableau 2.6) ne présente aucune difficulté, puis la saisonnalité est ajoutée. La somme des carrés des résidus (écarts entre série brute et série prévue) est égale à :

$$SCR^{LES} = \sum_{t=2}^{12} e_t^2 = 276\,478,20.$$

$$MSE^{LES} = \frac{276\,478,20}{11} = 25\,134,38$$

Tableau 2.6 – Prédiction par un LES et ajout de la saisonnalité

t	x_t	x_t^{CVS}	\hat{x}_t^{CVS}	S_t	x_t^P	e_t
1	1 248	1 841				
2	1 392	1 745	1 841,00	- 352,72	1 488	- 96
3	1 057	1 565	1 812,20	- 508,51	1 304	- 247
4	3 159	1 705	1 738,04	1 453,93	3 192	- 33
5	891	1 484	1 728,13	- 592,70	1 135	- 245
6	1 065	1 418	1 654,89	- 352,72	1 302	- 237
7	1 118	1 626	1 583,82	- 508,51	1 075	42
8	2 934	1 480	1 596,48	1 453,93	3 050	- 116
9	1 138	1 731	1 561,53	- 592,70	969	169
10	1 456	1 809	1 612,37	- 352,72	1 260	196
11	1 224	1 732	1 671,36	- 508,51	1 163	61
12	3 090	1 636	1 689,55	1 453,93	3 143	- 53
13			1 673,49	- 592,70	1 081	
14			1 673,49	- 352,72	1 321	
15			1 673,49	- 508,51	1 165	
16			1 673,49	1 453,93	3 127	

c) Prédiction par un LED ($\lambda = 0,3$)

L'application du lissage double (tableau 2.7) ne présente, là encore, aucune difficulté, à la série prévue par le LED \hat{x}_t^P , on ajoute la saisonnalité. La somme des carrés des résidus (écarts entre série brute et série prévue) est égale à :

$$SCR^{LED} = \sum_{t=2}^{12} e_t^2 = 275\,688,33$$

$$MSE^{LED} = \frac{275\,688,33}{11} = 25\,062,58$$

Tableau 2.7 – Prédiction par un LED et ajout de la saisonnalité

t	x_t	x_t^{CVS}	\hat{x}_t^{CVS}	$\hat{\hat{x}}_t^{CVS}$	a_t	b_t	\hat{x}_t^P	S_t	x_t^P	e_t
1	1 248	1 841								
2	1 392	1 745	1 841,00	1 841,00	1 841,00	0,00	1 841,00	- 352,72	1 488	- 96
3	1 057	1 565	1 812,20	1 832,36	1 792,04	- 8,64	1 783,40	- 508,51	1 275	- 218
4	3 159	1 705	1 738,04	1 804,06	1 672,02	- 28,30	1 643,72	1 453,93	3 098	61
5	891	1 484	1 728,13	1 781,28	1 674,97	- 22,78	1 652,19	- 592,70	1 059	- 169
6	1 065	1 418	1 654,89	1 743,37	1 566,41	- 37,92	1 528,50	- 352,72	1 176	- 110
7	1 118	1 626	1 583,82	1 695,50	1 472,14	- 47,86	1 424,28	- 508,51	916	202
8	2 934	1 480	1 596,48	1 665,79	1 527,16	- 29,71	1 497,45	1 453,93	2 951	- 17
9	1 138	1 731	1 561,53	1 634,52	1 488,55	- 31,28	1 457,27	- 592,70	8 65	274
10	1 456	1 809	1 612,37	1 627,87	1 596,87	- 6,64	1 590,23	- 352,72	1 238	218
11	1 224	1 732	1 671,36	1 640,92	1 701,80	13,05	1 714,85	- 508,51	1 206	18
12	3 090	1 636	1 689,55	1 655,51	1 723,60	14,59	1 738,19	1 453,93	3 192	- 102
13			1 673,49	1 660,90	1 686,07	5,39	1 691,46	- 592,70	1 099	
14							1 696,86	- 352,72	1 344	
15							1 702,25	- 508,51	1 194	
16							1 707,64	1 453,93	3 162	

Les résultats obtenus par le LES ou le LED sont moins performants sur la période historique selon le critère de recherche du minimum du MSE. Par ailleurs les résidus prévisionnels sont affectés dans les deux cas d'une tendance.

La « bonne » méthode de prévision est donc la plus simple ; celle qui consiste à ajouter les coefficients saisonniers à la valeur moyenne des ventes CVS. La figure 2.5 illustre le graphique des ventes brutes et prévues à l'aide de cette méthode.

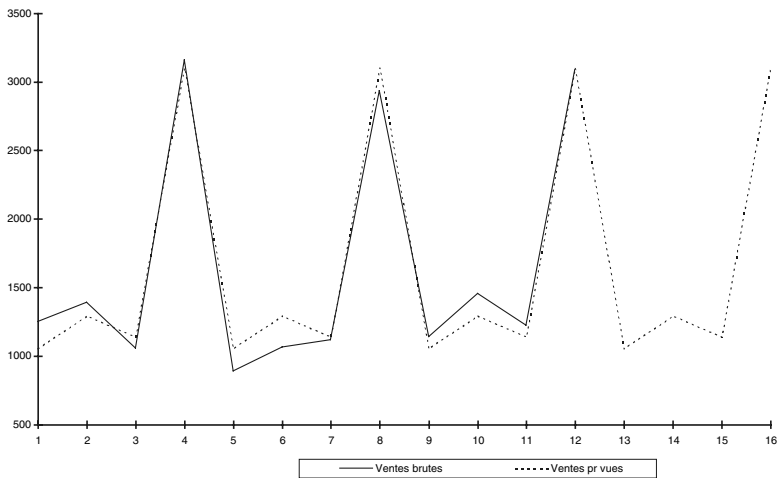


Figure 2.5 – Graphique des ventes brutes et ventes prévues

C. Prévision par lissage exponentiel de Holt-Winters

Si la saisonnalité est de type aléatoire (ce qui est fréquent en économétrie), nous pouvons recourir aux techniques de désaisonnalisation par filtrage (méthode Census ou démodulation complexe) ou encore à des méthodes de lissage exponentiel.

Si nous appliquons une méthode de type Census, on obtient non pas des coefficients saisonniers mais une matrice de coefficients qui autorise la prévision de la saisonnalité. Cette méthode est donc à la fois une méthode de désaisonnalisation et une technique de prévision de la chronique. Elle est classée parmi les méthodes qui permettent de prendre en compte des saisonnalités aléatoires puisqu'on peut démontrer qu'elle se résume à un filtre linéaire de type Moyennes Mobiles¹.

Il en est de même des méthodes de démodulation complexe et TRAMO-SEAT qui permettent de séparer la partie saisonnière de l'extra-saisonnalité de la chronique. Ces deux composantes peuvent alors être modélisées séparément et combiner différentes approches de prévision pour chacune d'entre elles.

Dans le prolongement du paragraphe précédent nous pouvons considérer que la tendance et la saisonnalité sont aléatoires et utiliser le lissage exponentiel pour en effectuer une prévision combinée. Cette technique a été mise au point par Holt et Winters (1960). Il s'agit d'un lissage exponentiel double (LED) de Holt à deux paramètres pour la partie non saisonnière et d'un lissage exponen-

1. Cf. Laroque, 1977, ou Gourieroux et Monfort, 1990.

tiel saisonnier à un paramètre de Winters. Cette méthode de lissage exponentiel comporte, donc, trois paramètres à estimer et il en existe deux versions : une version multiplicative et une version additive.

1) Le schéma multiplicatif

La chronique s'écrit dans ce cas :

$$x_t = (a_t + b_t t) S_t + \varepsilon_t$$

Trois lissages distincts sont effectués :

- le lissage de la moyenne a avec un coefficient de lissage α , avec $\alpha \in [0; 1]$;
- le lissage de la tendance b avec un coefficient de lissage β , avec $\beta \in [0; 1]$;
- le lissage de la saisonnalité S avec un coefficient de lissage γ , avec $\gamma \in [0; 1]$.

• Formulation

Lissage de la moyenne¹ : $a_t = \alpha(x_t/S_{t-p}) + (1 - \alpha)(a_{t-1} + b_{t-1})$

Lissage de la tendance : $b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1}$

Lissage de la saisonnalité : $S_t = \gamma(x_t/a_t) + (1 - \gamma)S_{t-p}$

Prévision à un horizon de h périodes :

$$\hat{x}_{t+h} = (a_t + hb_t)S_{t-p+h} \quad \text{si } 1 \leq h \leq p$$

$$\hat{x}_{t+h} = (a_t + hb_t)S_{t-2p+h} \quad \text{si } p + 1 \leq h \leq 2p$$

avec² :

$$\hat{x}_{t+h} = x_{t,t+h}^P$$

a_t = moyenne lissée de la série en t

x_t = valeur observée de la série en t

S_t = coefficient saisonnier en t

p = périodicité des données ($p = 12$ en mensuel, $p = 4$ en trimestriel)

b_t = tendance estimée en t .

Dans certaines écritures du modèle les coefficients saisonniers vérifient la propriété : $\sum_{i=1}^p S_i = p$ selon le principe de la conservation des aires.

1. On utilise S_{t-p} car S_t n'est pas encore connue.

2. Il convient de noter la différence d'écriture retenue pour la prévision avec celle du LES et du LED de Brown.

2) Le schéma additif

La chronique s'écrit dans ce cas : $x_t = a_t + b_t t + S_t + \varepsilon_t$

- Formulation

Lissage de la moyenne : $a_t = \alpha(x_t - S_{t-p}) + (1 - \alpha)(a_{t-1} + b_{t-1})$

Lissage de la tendance : $b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1}$

Lissage de la saisonnalité : $S_t = \gamma(x_t - a_t) + (1 - \gamma)S_{t-p}$

Prévision à un horizon de h périodes :

$$\hat{x}_{t+h} = (a_t + hb_t) + S_{t-p+h} \quad \text{si } 1 \leq h \leq p$$

$$\hat{x}_{t+h} = (a_t + hb_t) + S_{t-2p+h} \quad \text{si } p + 1 \leq h \leq 2p$$

Dans ce cas le principe de la conservation des aires implique : $\sum_{i=1}^p S_i = 0$.

Les paramètres α , β et γ sont optimisés comme pour les méthodes non saisonnières en minimisant la somme des carrés des erreurs prévisionnelles entre la valeur observée de la chronique et les valeurs prévues.

3) Initialisation du modèle de Holt-Winters

Comme pour les autres méthodes de lissage, dans la pratique, il y a un problème de démarrage de la technique : les valeurs de départ peuvent être estimées par la méthode des moindres carrés ordinaires ou plus simplement initialisées pour la première année ($t = 1 \dots, p$) de la manière suivante :

- Initialisation de la saisonnalité

Les coefficients saisonniers pour la première année sont estimés par la valeur observée en $t(x_t)$ divisée par la moyenne \bar{x} des p premières observations de la première année.

$$S_t = x_t / \bar{x} \quad \text{pour } t = 1 \dots, p$$

- Initialisation de la moyenne lissée : $a_p = \bar{x}$
- Initialisation de la tendance : $b_p = 0$

La technique d'initialisation proposée par Montgomery et Johnson est préférable bien que plus complexe : Il s'agit de calculer les moyennes arithmétiques des k premières périodes $\bar{x}_1, \dots, \bar{x}_k$.

Puis :

$$b_0 = \frac{\bar{x}_k - \bar{x}_1}{(k-1)p} \quad a_0 = \bar{x}_1 - \frac{p}{2} b_0$$

$$s_{j-p} = \frac{pR_j}{\sum_{j=1}^p R_j}$$

$$R_j = \frac{1}{k} \sum_{t=1}^k \frac{x_{(t-1)p+j}}{\bar{x}_t - \left(\frac{p+1}{2} - j\right) b_0}$$

Exercice n° 2.5

Fichier C1EX1

Prévision par le modèle de Holt-Winters

Utiliser la version additive du modèle de Holt-Winters avec $\alpha = 0,4$; $\beta = 0,1$ et $\gamma = 0,3$ pour calculer une prévision à un horizon de cinq trimestres à partir des données du tableau 1.1 du chapitre 1.

Solution

Le tableau 2.8 illustre le calcul de la prévision à l'aide du modèle de Holt-Winters (l'initialisation est effectuée par la méthode la plus simple).

• Initialisation ($p = 4$)

$$a_p = a_4 = \text{moyenne de la première année} = 1\,714$$

$$b_p = b_4 = 0$$

$$S_t = x_t - 1\,714 \text{ pour } t = 1, 2, 3, 4.$$

$$\alpha = 0,4 ; \beta = 0,1 ; \gamma = 0,3$$

• Exemples de calcul¹ pour la ligne $t = 7$

$$a_t = \alpha(x_t - S_{t-p}) + (1 - \alpha)(a_{t-1} + b_{t-1})$$

$$a_7 = \alpha(x_7 - S_3) + (1 - \alpha)(a_6 + b_6)$$

$$a_7 = 0,4 \times (1\,118 - (-657)) + 0,6 \times (1\,488,95 + (-21,08)) = 1\,590,73$$

$$b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1}$$

$$b_7 = \beta(a_7 - a_6) + (1 - \beta)b_6$$

$$b_7 = 0,1 \times (1\,590,73 - 1\,488,95) + 0,9 \times (-21,08) = 1\,590,73$$

$$S_t = \gamma(x_t - a_t) + (1 - \gamma)S_{t-p}$$

$$S_7 = \gamma(x_7 - a_7) + (1 - \gamma)S_3$$

$$S_7 = 0,3 \times (1\,118 - 1\,590,73) + 0,7 \times (-657) = -601,72$$

$$S_t^* = S_t - \frac{\sum_{i=1}^p S_i}{P}$$

1. Des différences concernant les décimales peuvent apparaître du fait que le tableur utilise tous les chiffres après la virgule.

$$S_7^* = S_7 - \frac{\sum_{i=5}^{i=8} S_i}{4} = -601,72 - \frac{-56,3}{4} = -587,64$$

$$\widehat{x}_{t+h} = (a_t + hb_t) + S_{t-p+h}^*$$

$$\widehat{x}_{6+1} = (a_6 + 1 \times b_6) + S_{6-4+1}^*$$

$$\widehat{x}_7 = (1\,488,95 + (-21,08)) + (-657,00) = 810,88$$

• Prévision à partir de $t = 12$

$$\widehat{x}_{t+h} = (a_t + hb_t) + S_{t-p+h}^* \quad 1 \leq h \leq p$$

$$\widehat{x}_{13} = (1\,696,85 + 1 \times 3,88) + (-512,67) = 1\,188,058$$

$$\widehat{x}_{14} = (1\,696,85 + 2 \times 3,88) + (-317,15) = 1\,387,46$$

$$\widehat{x}_{15} = (1\,696,85 + 3 \times 3,88) + (-579,33) = 1\,129,15$$

$$\widehat{x}_{16} = (1\,696,85 + 4 \times 3,88) + 1\,409,15 = 3\,121,51$$

$$\widehat{x}_{t+h} = (a_t + hb_t) + S_{t-2p+h} \quad p+1 \leq h \leq 2p$$

$$\widehat{x}_{17} = (1\,696,85 + 5 \times 3,88) + (-512,67) = 1\,203,56$$

Tableau 2.8 – Exemple de prévision à partir du modèle de Holt-Winters (schéma additif) avec $\alpha = 0,4$; $\beta = 0,1$; $\gamma = 0,3$

t	Dates	x_t	a_t	b_t	S_t	S_t^*	\widehat{x}_t	e_t
1	Année 1 : T_1	1 248,30			-466,00	-466,00		
2	T_2	1 392,10			-322,00	-322,00		
3	T_3	1 056,60			-657,00	-657,00		
4	T_4	3 159,10	1 714,00	0,00	1 445,00	1 445,00		
5	Année 2 : T_1	890,80	1 571,20	-14,28	-530,26	-516,19		
6	T_2	1 065,30	1 488,95	-21,08	-352,59	-338,51	1 234,92	-169,92
7	T_3	1 117,60	1 590,73	-8,79	-601,72	-587,64	810,88	307,12
8	T_4	2 934,20	1 544,76	-12,51	1 428,27	1 442,34	3 026,93	-92,93
9	Année 3 : T_1	1 138,20	1 581,03	-7,63	-504,09	-512,67	1 016,06	121,93
10	T_2	1 456,00	1 661,84	1,21	-308,56	-317,15	1 234,88	221,11
11	T_3	1 224,30	1 722,49	7,16	-570,75	-579,33	1 075,41	148,59
12	T_4	3 090,20	1 696,85	3,88	1 417,74	1 409,15	3 171,99	-81,99
13	Année 4 : T_1						1 188,05	
14	T_2						1 387,46	
15	T_3						1 129,15	
16	T_4						3 121,51	
17	Année 5 : T_1						1 203,56	

En combinant les différentes possibilités structurelles de l'extra-saisonnalité et de la saisonnalité, nous pouvons décrire la plupart des méthodes de cette première partie.

- Si l'extra-saisonnier et la saisonnalité sont rigides, la méthode de la régression sur fonction trigonométrique ou sur variable dichotomique se justifie.
- Si l'extra-saisonnier et la saisonnalité sont souples, on recourt au lissage exponentiel de Holt-Winters.
- Si les composantes sont de natures différentes, la désaisonnalisation par une méthode de moyenne mobile ou de Census est une étape préalable avant de traiter la nature rigide ou souple de l'extra-saisonnier.

Cela résume le traitement général d'une chronique sachant qu'il n'est pas possible (en l'état) de traiter simultanément les deux composantes : saisonnière et extra-saisonnière. Il est envisageable de contourner cette difficulté en procédant à une désaisonnalisation préalable de la chronique par une méthode de moyenne mobile d'ordre égal à la périodicité des données ou par une méthode de Census ou TRAMO-SEAT.

Nous ne pouvons pas traiter dans le cadre de ce manuel l'exhaustivité des méthodes classiques de prévision à court terme en particulier pour le lissage exponentiel¹ ; il donne cependant de l'analyse d'une chronique les principaux aspects retrouvés dans les logiciels spécialisés dans le traitement des séries chronologiques et d'économétrie.

1. Mélard, 1990.

Partie II

Traitement des séries temporelles, réalisations de processus aléatoires

Les techniques traditionnelles de prévision des séries temporelles se révèlent dans la plupart des cas insuffisantes pour prévoir les phénomènes économiques. En dehors des méthodes de lissage, elles sont construites à partir de l'extrapolation de composantes déterministes qui mutilent la réalité économique. Les statisticiens ont pu observer que ces instruments forgeaient le plus souvent mieux qu'ils n'informaient sur les réalités intrinsèques des séries. C'est ce qui explique l'engouement des chercheurs à investir ce domaine dans les années 1970 suite aux travaux de Box et Jenkins qui proposent une nouvelle « philosophie » du traitement des séries temporelles. Ils considèrent en effet qu'à un instant t , la valeur x_t de la chronique est un état d'une variable aléatoire. Ce faisant sur une période d'observation, la chronique est constituée par l'ensemble d'une succession d'états de variables aléatoires. Cette famille de variables aléatoires indexées par le temps, porte le nom de processus aléatoire et la chronique, succession temporelle d'états, celle d'échantillon ou de trajectoire du processus aléatoire. En procédant de cette façon, il est alors possible de recourir à l'ensemble des tra-

vaux concernant les processus aléatoires pour en faire bénéficier les chroniques économiques. Mais en économie l'expérience n'est pas renouvelable. On dispose donc, en tout et pour tout, du seul échantillon temporel : la chronique réalisation du phénomène. Il faut, en conséquence, recourir à l'hypothèse forte d'ergodicité qui suppose que l'estimation des moments du processus pour chacun des instants t tend vers l'estimation des moments le long du processus et donc vers les moments temporels de l'échantillon.

La théorie des processus aléatoires a débuté en 1927 avec les travaux de Yule quand ce dernier a introduit la notion de choc aléatoire ou encore d'impulsion dans les séries temporelles. Cette notion, associée à celle de composante non périodique, a permis de caractériser le processus aléatoire dont les principales propriétés ont été décrites par Kolmogorov en 1933. En 1930, Wiener généralise l'analyse harmonique introduite en 1850 par le mathématicien français J. B. Fourier au cas stochastique. L'idée de cette généralisation est que, sous certaines conditions, le processus stochastique peut être décomposé en une somme infinie de sinusoides de fréquences déterministes différentes mais d'amplitude aléatoire (alors que pour l'analyse harmonique de Fourier les amplitudes sont aussi déterministes). C'est le caractère périodique de la décomposition harmonique qui a conduit Khintchine en 1934 à introduire la classe des processus aléatoires stationnaires. En particulier est définie la stationnarité d'ordre 2 qui impose au premier moment du processus (l'espérance mathématique) d'être indépendant du temps et à la covariance de n'être fonction que du retard entre deux instants. Cela se traduit, par exemple, au niveau de la chronique échantillon par une absence de tendance et de variance explosive ou amortie. En 1938, Wold décompose dans le domaine du temps les processus aléatoires selon deux processus orthogonaux dont l'un est dit déterminable et l'autre indéterminable. Le premier est appelé ainsi car il peut être exactement prévu. Le second est le résultat d'une combinaison linéaire infinie de chocs aléatoires : c'est le processus de Yule. Il s'agit d'un processus qui – à cause de ses caractéristiques – ne peut être exactement prévu. En 1954, Wold bâtit, à partir de la classe des processus indéterminables, les modèles linéaires ARMA stationnaires. La partie autorégressive de ces processus notée AR (*AutoRegressive Process*) est constituée par une combinaison linéaire finie des valeurs passées du processus. La partie moyenne mobile notée MA (*Moving Average Process*) est constituée d'une combinaison linéaire finie en t des valeurs passées d'un bruit blanc, c'est-à-dire d'un processus aléatoire formé d'une succession de variables aléatoires indépendantes et d'espérance mathématique nulle. Wold montre que la classe ARMA est suffisante pour représenter la plupart des cas concrets. En 1960, le théorème de Wold-Cramer résout le problème de la prédiction linéaire au sens des Moindres Carrés des processus ARMA (théorème de la décomposition canonique des processus ARMA). En 1970, dans la première édition de leur ouvrage puis en 1976 dans la deuxième édition, Box et Jenkins rassemblent tous ces travaux épars dans une méthodologie itérative ayant pour but la prévision d'une chronique. Dans les années 1980, plusieurs auteurs, en particulier Granger et Sims, ont étendu l'étu-

de des processus aléatoires univariés à celle des processus aléatoires multivariés, c'est-à-dire des processus représentés par plusieurs chroniques échantillons.

Dans le cas univarié, l'algorithme utilisé est celui de Box et Jenkins. La figure 1 le présente sous sa forme la plus traditionnelle.

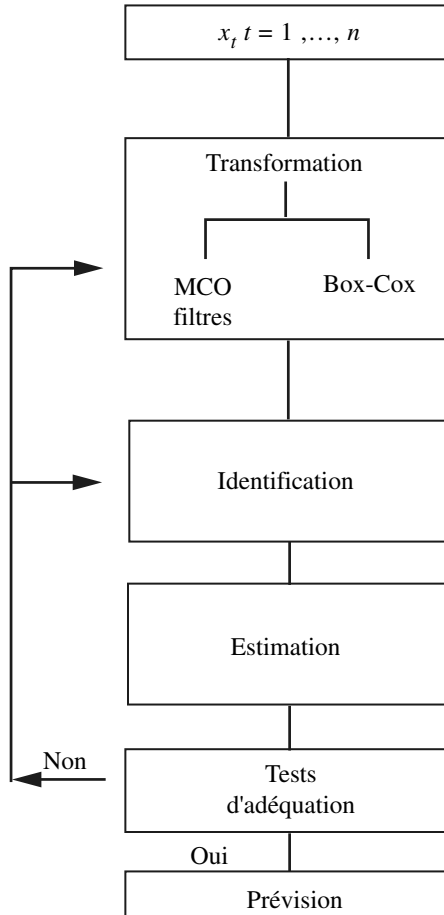


Figure 1 – Traitement dans le cas univarié

Il s'agit d'un algorithme itératif en cinq étapes. La première a pour objet de transformer la chronique échantillon afin d'éliminer d'éventuelles non stationnarité (tendance, saisonnalité...). Les caractéristiques de la chronique sont alors rapprochées de celles théoriques des processus ARMA afin d'identifier le processus générateur. L'estimation des paramètres – réalisée à partir du critère des moindres carrés – permet de valider ou d'invalider le modèle retenu.

Les itérations entre ces quatre étapes permettent de sélectionner un processus qui après prise en compte des différentes transformations de départ (recolation) peut être utilisé pour prévoir, dans une 5^e étape, la chronique.

Dans cet algorithme, les modèles sélectionnés (les modèles ARMA) sont des modèles linéaires. Plus récemment, certains auteurs (Tong, 1990, par exemple) ont associé à cette démarche des processus non linéaires. Cette non linéarité peut concerner le moment d'ordre 1 du processus (la moyenne pour l'échantillon), on les appelle modèles TAR (*Threshold – AR*), modèles Bilinear, modèles EXAR (*AR EXponentiels*), etc.

Cette non linéarité peut aussi concerner la variance de la chronique échantillon (moment d'ordre 2 du processus) ; ces modèles de type ARCH (*AutoRegressive Conditional Heteroscedasticity*) ont été introduits par Engle en 1982. Ils ont donné lieu à la construction de nombreux autres modèles (Guegan, 1994) combinant ce principe avec ceux des modèles ARMA (*CHARMA, GARCH...*).

Dans cette partie nous définissons les processus ARMA dans le domaine du temps (chapitre 3) puis dans celui des fréquences (chapitre 4). Nous présentons les tests de non stationnarité (chapitre 5) qui permettent de transformer les chroniques échantillons pour les identifier (chapitre 6) dans la classe des processus ARMA. L'estimation, les tests et la prévision font l'objet du chapitre 7 et enfin, le chapitre 8 concerne quelques extensions aux modèles à mémoire longue (ARFIMA), à structure chaotique et enfin, aux modèles à erreurs ARCH.

3. *P*rocessus aléatoires stationnaires et processus ARMA

I. Définition d'un processus stochastique

Un processus aléatoire est une application X qui associe au couple (ω, t) la quantité $x_t(\omega)$. Elle est telle que $\forall t \in T$ fixé, x_t est une variable aléatoire définie sur un espace probabilisé.

Un processus stochastique est donc une famille de variables aléatoires indexées par t noté $(x_t, t \in T)$ ou encore x_t .

Dans la suite de l'exposé, l'espace des indices T est le temps, t est alors l'instant d'observation de la variable aléatoire x sur l'individu ω , $\omega \in \Omega$ avec Ω l'espace fondamental permettant de définir la variable aléatoire. Si T est l'ensemble des réels, le processus est dit continu. Si $T = \mathbb{Z}$ ou \mathbb{N} ou \mathbb{N}^* ou \mathbb{N}^* borné, le processus est dit discret.

Pour ω fixé, $x_t(\omega)$ porte le nom de trajectoire (nous dirons réalisation) de x pour l'individu ω .

On suppose, par la suite, que la série temporelle notée x_t , $t \in \mathbb{N}^*$ bornée (soit une succession d'observations régulièrement espacées dans le temps d'une valeur économique) est une réalisation d'un processus stochastique discret, univarié, de même nom et notée aussi et sans ambiguïté x_t , $t \in \mathbb{N}^*$ borné.

La chronique est dite échantillon ou réalisation du processus aléatoire et ce dernier est appelé le processus générateur de la chronique.

II. Les processus stationnaires

A. Définition d'un processus stationnaire au sens strict : la stationnarité forte

Soit un processus aléatoire réel $x_t, t \in T$.

Le processus x_t est dit strictement ou fortement stationnaire si \forall le n -uple du temps $t_1 < t_2 < \dots < t_n$ tel que $t_i \in T$ et pour tout temps $h \in T$ avec $t_i + h \in T, \forall i, i = 1, \dots, n$, la suite $(x_{t_1+h}, \dots, x_{t_n+h})$ a la même loi de probabilité que la suite $(x_{t_1}, \dots, x_{t_n})$.

La loi de probabilité qui correspond à la suite $(x_{t_1}, \dots, x_{t_n})$ est caractérisée par sa fonction de répartition, d'où la définition équivalente de la stationnarité forte :

$$\forall (x_1, \dots, x_n), \forall (t_1, \dots, t_n), \text{ et } \forall h :$$

$$P[x_{t_1} < x_1, \dots, x_{t_n} < x_n] = P[x_{t_1+h} < x_1, \dots, x_{t_n+h} < x_n].$$

Ainsi un processus aléatoire est strictement stationnaire si toutes ses caractéristiques, c'est-à-dire *tous ses moments*, sont invariants pour tout changement de l'origine du temps.

Dans le cas où un processus $x_t, t \in T$ est tel que $T = \mathbb{R}, \mathbb{Z}$ ou \mathbb{N} alors on peut vérifier que si x_t est un processus strictement stationnaire :

$$E[x_t] = m \quad \forall t \in T$$

$$V[x_t] = \sigma^2 \quad \forall t \in T$$

$$\text{cov}[x_t, x_s] = \gamma[|t - s|] \quad \forall t \in T, \forall s \in T, t \neq s$$

$$\text{Ou encore : } \text{cov}(x_1, x_{1+k}) = \text{cov}(x_2, x_{2+k}) = \dots = \text{cov}(x_t, x_{t-k})$$

La covariance dépend de la différence de temps seule et non du temps.

$\text{cov}[x_t, x_s]$ existe si $E[x_t^2] < \infty$ et $E[x_s^2] < \infty$ (d'après l'inégalité de Schwarz).

Il s'agit donc d'une famille de variables aléatoires réelles homoscedastiques et corrélées.

B. La stationnarité d'ordre deux des processus : la stationnarité faible

Le processus $x_t, t \in T$ est dit faiblement stationnaire si seuls les moments d'ordre 1 et d'ordre 2 sont stationnaires. Par exemple, si $E[x_t^3]$ dépend du temps t alors le processus est faiblement stationnaire.

Les processus stationnaires d'ordre 2 sont des processus générateurs de chronique sans tendance en moyenne et sans tendance en variance mais cela ne signifie pas que les séries temporelles ont une représentation graphique stable.

C. Le processus Bruit Blanc (*White Noise*)

Soit le processus, $x_t, t \in T$. Si pour tout n -uplet du temps $t_1 < t_2 < \dots < t_n$, les variables aléatoires réelles suivantes $x_{t_2} - x_{t_1}, \dots, x_{t_n} - x_{t_{n-1}}$ (différences premières) sont indépendantes, il s'agit d'un processus à accroissements indépendants.

Le processus $x_t, t \in T$ est dit à accroissements indépendants stationnaires si de plus la loi de probabilité de $(x_{t+h} - x_t) \forall h \in T$ ne dépend pas de t .

Un Bruit Blanc¹ est un processus stochastique à accroissements non corrélés. Il est dit Bruit Blanc « fort » si les accroissements sont indépendants. Il s'agit donc d'une suite de variables aléatoires réelles homoscédastiques et indépendantes. On l'appelle aussi processus *i.i.d.* (processus discret formé de variables mutuellement indépendantes et identiquement distribuées). Si la loi de probabilité de x_t est normale alors le Bruit Blanc est nécessairement *i.i.d.* Il est parfois dit Bruit Blanc gaussien et noté alors *n.i.d.* (normalement et identiquement distribuée).

Un Bruit Blanc est donc tel que :

$$\begin{aligned} E[x_t] &= m & \forall t \in T \\ V[x_t] &= \sigma^2 & \forall t \in T \\ \text{cov}[x_t, x_{t+\theta}] &= \gamma_x(\theta) = 0 & \forall t \in T, \forall \theta \in T \end{aligned}$$

Si $E[x_t] = 0$, le Bruit Blanc est centré, ce que l'on suppose par la suite.

Un processus *i.i.d.* ou *n.i.d.* est nécessairement stationnaire mais tous les processus stationnaires ne sont pas *i.i.d.* ou *n.i.d.*, dans ce dernier cas le processus stationnaire est dit à mémoire, c'est-à-dire qu'il existe une loi de reproduction interne au processus qui est donc modélisable.

1. Appellation donnée par les ingénieurs par référence au spectre de la lumière blanche (cf. Chatfield, 1984).

Le tableau 3.1 illustre les différents types de processus.

Tableau 3.1 – Les différents types de processus aléatoires

<i>Processus non stationnaire</i>	<i>Processus stationnaire</i>
	Processus bruit blanc
	Processus bruit blanc gaussien

Il existe aussi des processus non stationnaires (*cf.* chapitre 5), mais également des processus qui ne sont pas des bruits blancs bien que l'espérance soit indépendante du temps et les covariances nulles mais dont la variance dépend du temps (hétéroscédasticité).

D. L'ergodicité

Les processus aléatoires recourent à l'hypothèse d'ergodicité pour les échantillons qui les constituent. Pour bien comprendre cette hypothèse que l'on utilise souvent de façon naturelle, il convient de définir ce que l'on entend par « moment dans le processus » et « moment le long du processus ».

Pour cela schématisons des processus aléatoires stationnaires et non stationnaires en supposant qu'à un instant t la loi de probabilité de la variable aléatoire est gaussienne.

Les figures 3.1 et 3.2 montrent que les moments du processus se calculent dans le processus pour l'instant donné et que la réalisation simultanée d'un nombre d'épreuves doit être suffisant pour que ce calcul ait lieu. Théoriquement, un processus peut être assimilé à une expérience, nous pouvons donc obtenir plusieurs échantillons résultant de ces expériences. La théorie de l'inférence statistique garantit la stabilité des estimations de ces moments. Mais

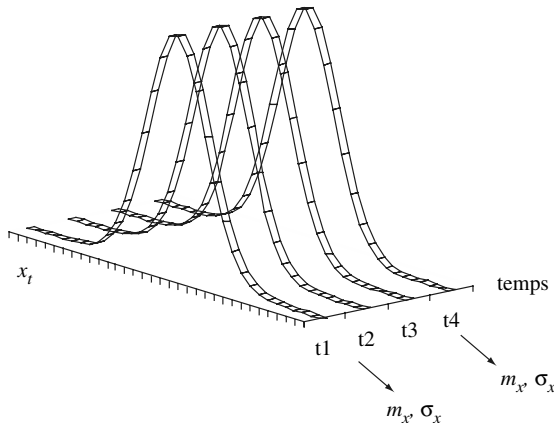


Figure 3.1 – Processus aléatoires stationnaires

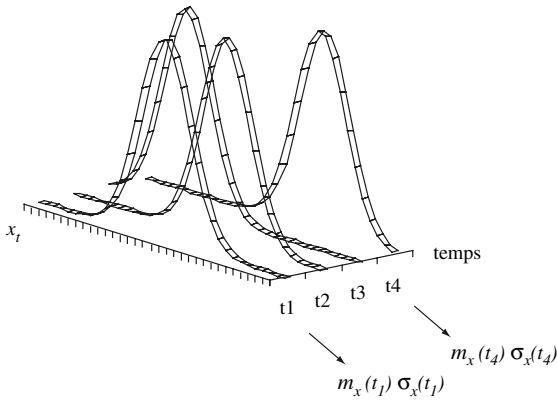


Figure 3.2 – Processus aléatoires non stationnaires

on peut aussi calculer les moments de l'échantillon, par exemple la moyenne temporelle \bar{x} , dans le cas discret : $\bar{x} = \lim_{n \rightarrow \infty} \frac{1}{2n+1} \sum_{-n}^{+n} x_t \quad t \in \mathbb{Z}$.

Lorsque ces moyennes existent, elles sont indépendantes du temps tout en pouvant être différentes d'un échantillon à l'autre.

On dit qu'un processus aléatoire est ergodique si on peut obtenir des estimateurs absolument corrects (*i.e.* sans biais et convergents) de ses caractéristiques (les moments) à partir d'un seul échantillon : la série temporelle.

III. La fonction d'autocorrélation et la fonction d'autocorrélation partielle

A. La fonction d'autocorrélation

La fonction d'autocovariance d'un processus x_t centré est une fonction de :

$$k \in \mathbb{Z} \rightarrow \gamma_k = E[x_t x_{t-k}] \in \mathbb{R}.$$

Elle est estimée par :

$$R_k = \frac{1}{n-k} \sum_{t=k+1}^n x_t x_{t-k} \quad k \in \mathbb{Z}^+$$

avec $R_0 = \frac{1}{n} \sum_{t=1}^n x_t^2 = V[x_t]$

La fonction d'autocorrélation (notée FAC) de x_t est une fonction de :

$$k \in Z \rightarrow \rho_k \in [-1; +1]$$

avec $\rho_k = \frac{\gamma_k}{\gamma_0}$. Elle est estimée par $r_k = \frac{R_k}{R_0}$

Le graphe de la fonction d'autocorrélation porte le nom de corrélogramme. La variance des autocorrélations est donnée par¹ :

$$V[\rho_k] = \frac{1}{n} \sum_{i=-K}^{+K} \rho_i^2 \quad K < k.$$

Remarque : On note ρ_i le coefficient d'autocorrélation théorique (inconnu) et r_i le coefficient d'autocorrélation estimé à partir de la chronique échantillon.

En utilisant la symétrie des ρ_i , l'estimation r_i de ρ_i et en supposant qu'au-delà de K les ρ_i sont nuls, on obtient :

$$V[r_k] = \frac{1}{n} \left(1 + 2 \sum_{i=1}^K r_i^2 \right)$$

Pour n grand, la distribution d'échantillon de l'estimateur r_k suit asymptotiquement une loi normale. La variable normale centrée réduite est :

$$t^* = \frac{|r_k - \rho_k|}{\sqrt{V[r_k]}}$$

Sous $H_0 : \rho_k = 0$, l'indicateur $t^* = \frac{r_k}{\sqrt{V[r_k]}}$ ($K < k$) est alors calculé et sa valeur absolue est comparée au t^α lu dans la table de la loi normale ($t^{\alpha/2} = -1,96$, valeur de la table de la loi normale pour un seuil de $\alpha = 5\%$). Si $|t^*| \geq 1,96$ on rejette l'hypothèse H_0 (pour les calculs on prend $K = k - 1$).

On peut aussi construire un intervalle de confiance pour ρ_k à 95 % :

IC = $\pm 1,96 \times \sqrt{V[r_k]}$ avec $K < k$. Comme la variance de r_k se modifie avec k , cet intervalle n'est pas constant.

Si on suppose que x_t est *i.i.d.*, alors, par définition, tous les coefficients d'autocorrélation sont nuls et donc $K = 0$ et $V[r_k] = \frac{1}{n} \forall k$. Sous cette condition

IC = $\pm 1,96 \times \sqrt{\frac{1}{n}}$, et c'est cet intervalle qui figure sur le corrélogramme (test de l'hypothèse $H_0 : \rho_k = 0$).

1. Bartlett, 1946.

B. La fonction d'autocorrélation partielle

La fonction d'autocorrélation partielle (notée FAP) d'un processus x_t centré est une fonction de $k \in \mathbb{Z} \rightarrow \rho_{kk} \in [-1; +1]$. Elle s'analyse de manière analogue au coefficient de corrélation partielle : il s'agit de la corrélation entre x_t et x_{t-k} , l'influence des variables x_{t-k-i} pour $(i < k)$ ayant été retirée. Les autocorrélations partielles empiriques, notées r_{kk} , sont obtenues des autocorrélations partielles théoriques ρ_{kk} en remplaçant les ρ_i par leurs estimations r_i .

Soit la matrice P_k symétrique formée des $(k - 1)$ premières autocorrélations de x_t :

$$P_k = \begin{bmatrix} 1 & \rho_1 & \dots & \dots & \rho_{k-1} \\ \vdots & 1 & & & \vdots \\ & & \ddots & & \\ & & & \ddots & \\ \rho_{k-1} & & & & 1 \end{bmatrix} \quad k \in \mathbb{N}$$

La FAP est la succession des $\rho_{kk} = \frac{|P_k^*|}{|P_k|}$ avec :

$|P_k^*|$ = déterminant de la matrice P_k dans laquelle on a remplacé la dernière colonne par le vecteur $[\rho_1 \dots \rho_k]$.

$$P_k^* = \begin{bmatrix} 1 & \rho_1 & \dots & \dots & \rho_1 \\ \vdots & & & & \vdots \\ & & & & \vdots \\ & & & & \vdots \\ \rho_{k-1} & & & & \rho_k \end{bmatrix}$$

Par exemple, calculons ρ_{11} , ρ_{22} et ρ_{33}

$$\rho_{11} = \frac{|P_1^*|}{|P_1|} \quad P_1 = [1] \quad P_1^* = [\rho_1]$$

$$\rho_{11} = \frac{\rho_1}{1} = \rho_1$$

On constate que la première valeur de l'autocorrélation partielle est égale à la première valeur de l'autocorrélation.

$$\rho_{22} = \frac{|P_2^*|}{|P_2|} \quad P_2 = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix} \quad P_2^* = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{bmatrix}$$

$$\rho_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

$$\rho_{33} = \frac{|P_3^*|}{|P_3|} \quad P_3 = \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix} \quad P_3^* = \begin{bmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{bmatrix}$$

$$\rho_{33} = \frac{\rho_1^3 - \rho_1 \rho_2 (2 - \rho_2) + \rho_3 (1 - \rho_1^2)}{1 - \rho_2^2 - 2\rho_1^2(1 - \rho_2)}$$

Rapidement, les calculs se compliquent, c'est pourquoi on utilise l'écriture récurrente de ρ_{ii} :

$$\rho_{ii} = \begin{cases} \rho_1 & \text{si } i = 1. \\ \frac{\rho_i - \sum_{j=1}^{i-1} \rho_{i-1,j} \rho_{i-j}}{1 - \sum_{j=1}^{i-1} \rho_{i-1,j} \rho_j} & i = 2, \dots, k. \end{cases}$$

où $\rho_{ij} = \rho_{i-1,j} - \rho_{ii} \rho_{i-1,i-j}$ avec $j = 1, 2, \dots, i-1$ et $\rho_{i,i} = \rho_{ii}$.

Dans la pratique le calcul des éléments de la FAP utilise les équations de **Yule-Walker**, définies au chapitre 6.

Pour des grands échantillons, les coefficients ρ_{kk} d'un processus AR(p) sont asymptotiquement distribués selon une loi normale d'espérance nulle et de

variance $V[\rho_{kk}] = \frac{1}{n}$ pour $k > p$ (Quenouille, 1949). Ce résultat est utilisé pour un ρ_{kk} quelconque. Il est alors possible de construire un test d'hypothèses pour ρ_{kk} .

Sous l'hypothèse $H_0 : \rho_{kk} = 0$.

$$\text{On calcule } t^* = \frac{|r_{kk}|}{\sqrt{\frac{1}{n}}}$$

Comme pour l'autocorrélation, le test de signification est effectué en utilisant la loi normale $\mathcal{N}(0, 1)$ par comparaison avec un t lu dans la table ($t^{\alpha/2} = -1,96$, valeur de la table de la loi normale pour un seuil de $= 5\%$). On déduit aussi un intervalle de confiance à 95% de ρ_{kk} : $\text{IC} = \left[\pm 1,96 \sqrt{\frac{1}{n}} \right]$. On constate que cet intervalle est constant et identique à celui de FAC lorsque x_t est un bruit blanc.

C. Analyse des fonctions d'autocorrélation

Lorsque nous étudions la fonction d'autocorrélation d'une série chronologique, la question qui se pose est de savoir quels sont les termes r_k qui sont significativement différents de 0.

En effet, par exemple si aucun terme n'est significativement différent de 0, on peut en conclure que le processus étudié est sans mémoire et donc qu'à ce titre il n'est affecté ni de tendance ni de saisonnalité. Ou encore si une série mensuelle présente une valeur élevée pour r_{12} (corrélacion entre x_t et x_{t-12}), la série étudiée est certainement affectée d'un mouvement saisonnier¹.

1) Test d'un coefficient d'autocorrélation

Le test d'hypothèses pour un terme ρ_k est le suivant :

$$H_0 : \rho_k = 0$$

$$H_1 : \rho_k \neq 0$$

Nous avons vu précédemment que si x_t est *i.i.d.* et pour un échantillon de taille importante ($n > 30$) le coefficient ρ_k tend de manière asymptotique vers une loi normale de moyenne 0 et d'écart type $1/\sqrt{n}$. Sous l'hypothèse H_0 l'intervalle de confiance du coefficient ρ_k est alors donné par :

$$\rho_k = 0 \pm t^{\alpha/2} \frac{1}{\sqrt{n}}$$

n = nombre d'observations.

Si le coefficient calculé r_k est à l'extérieur de cet intervalle de confiance, il est significativement différent de 0 au seuil α (en général $\alpha = 0,05$). La plupart des logiciels fournissent, avec le corrélogramme, l'intervalle de confiance ce qui autorise une interprétation instantanée.

Nous devons souligner une limite des tests à 5 %. En effet, lorsqu'une fonction d'autocorrélation est calculée pour un nombre important de retards, nous pouvons nous attendre à ce que quelques-uns soient, de manière fortuite, significativement différents de 0. Si h est le nombre de retards, le nombre possible de faux rejets est alors de $0,05 \times h$, pour un seuil de confiance de 5 %.

2) Test d'un ensemble de coefficients d'autocorrélation

Le test de Box-Pierce permet d'identifier les processus de marche au hasard.

1. Il s'agit même d'un test de détection de saisonnalité (cf. chapitre 1).

Si x_t , est une marche au hasard alors $\text{cov}(x_t, x_{t-k}) = 0$ ou encore $\rho_k = 0 \forall k$, d'où le test d'hypothèses :

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_K = 0$$

H_1 : il existe au moins un ρ_i significativement différent de 0.

Pour effectuer ce test, on recourt à la statistique Q (due à Box-Pierce¹) qui est donnée par :

$$Q = n \sum_{k=1}^h r_k^2$$

h = nombre de retards, r_k = autocorrélation empirique d'ordre k , n = nombre d'observations.

La statistique Q est distribuée de manière asymptotique comme un χ^2 (chi-deux) à h degrés de liberté. Nous rejetons, donc, l'hypothèse de bruit blanc, au seuil α , si la statistique Q est supérieure au χ^2 lu dans la table au seuil $(1 - \alpha)$ et h degrés de liberté.

Nous pouvons utiliser aussi une autre statistique, dont les propriétés asymptotiques sont meilleures, dérivée de la première qui est le Q' de Ljung et Box² :

$$Q' = n(n+2) \sum_{k=1}^h \frac{r_k^2}{n-k}$$

qui est aussi distribué selon un χ^2 à h degrés de liberté et dont les règles de décisions sont identiques au précédent. Ces tests sont appelés par les Anglo-Saxons : « portmanteau test ».

Exercice n° 3.1

Calcul des termes d'un corrélogramme et de leur intervalle de confiance à partir des éléments d'une chronique

Soit la chronique x_t , suivante :

t	1	2	3	4	5	6	7	8	9	10
x_t	10	3	-1	3	2	5	3	2	-1	3

1. Box et Pierce, 1970.
2. Ljung et Box, 1978.

- a) Calculer les trois premiers termes de la FAC (r_1, r_2 et r_3).
- b) Donner un intervalle de confiance à 95 % pour ρ_1, ρ_2 et ρ_3 et tester la significativité de r_1, r_2 et r_3 .
- c) Calculer les trois premiers termes de la FAP (r_{11}, r_{22} et r_{33}).
- d) Donner un intervalle de confiance à 95 % pour ρ_{11}, ρ_{22} et ρ_{33} et tester la significativité de r_{11}, r_{22} et r_{33} .
- e) Calculer la statistique de Ljung-Box pour $h = 3$, la série x_t est-elle un bruit blanc ?

Solution

- a) Le coefficient d'autocorrélation d'ordre k est donné par :

$$r_k = \frac{\sum_{t=k+1}^n (x_t - \bar{x}_1)(x_{t-k} - \bar{x}_2)}{\sqrt{\sum_{t=k+1}^n (x_t - \bar{x}_1)^2 \sum_{t=k+1}^n (x_{t-k} - \bar{x}_2)^2}}$$

$$\bar{x}_1 = \frac{1}{n-k} \sum_{t=k+1}^n x_t \quad \bar{x}_2 = \frac{1}{n-k} \sum_{t=k+1}^n x_{t-k}$$

Soit $r_1 = 0,021$; $r_2 = -0,502$; $r_3 = -0,348$

- b) En utilisant la formule de Bartlett, on obtient :

$$V[r_1] = \frac{1}{n} = \frac{1}{10} = 0,1 ;$$

$$V[r_2] = \frac{1}{n} (1 + 2r_1^2) = \frac{1}{10} (1 + 2 \times (0,021)^2) \approx 0,1$$

$$V[r_3] = \frac{1}{n} (1 + 2(r_1^2 + r_2^2)) = \frac{1}{10} (1 + 2(0,021^2 + 0,502^2)) \approx 0,15$$

Les intervalles de confiance à 95 % sous l'hypothèse $H_0 : \rho_k = 0$ pour $k = 1, 2, 3$ sont égaux à :

$$k = 1 : \left[-1,96\sqrt{\frac{1}{n}}; +1,96\sqrt{\frac{1}{n}} \right] = [-1,96\sqrt{0,10}; +1,96\sqrt{0,10}] = [-0,620; 0,620] \Rightarrow r_1$$

n'est pas significativement différent de 0.

$$k = 2 : [-1,96\sqrt{0,10}; +1,96\sqrt{0,10}] = [-0,620; 0,620] \Rightarrow r_2 \text{ n'est pas significativement différent de 0.}$$

$$k = 3 : [-1,96\sqrt{0,15}; +1,96\sqrt{0,15}] = [-0,759; 0,759] \Rightarrow r_3 \text{ n'est pas significativement différent de 0.}$$

- c) En utilisant les estimations r_1, r_2 et r_3 de ρ_1, ρ_2 et ρ_3 , nous obtenons : $r_{11} = r_1 = 0,021$

$$\begin{bmatrix} r_{21} \\ r_{22} \end{bmatrix} = \begin{bmatrix} 1 & r_1 \\ r_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 1 & 0,021 \\ 0,021 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0,021 \\ -0,502 \end{bmatrix} \rightarrow r_{22} = -0,502$$

$$\begin{bmatrix} r_{31} \\ r_{32} \\ r_{33} \end{bmatrix} = \begin{bmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0,021 & -0,502 \\ 0,021 & 1 & 0,021 \\ -0,502 & 0,021 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0,021 \\ -0,502 \\ -0,348 \end{bmatrix} = \begin{bmatrix} -0,185 \\ -0,489 \\ -0,431 \end{bmatrix} \Rightarrow r_{33} = -0,431$$

d) Sous l'hypothèse $H_0 : \rho_{kk} = 0$ pour $k = 1, 2, 3$, nous avons $\rho_{kk} \rightarrow \mathcal{N}\left(0, \sqrt{\frac{1}{n}}\right)$;

les intervalles de confiance à 95 % sont alors égaux à : $\left[-1,96\sqrt{\frac{1}{n}}; +1,96\sqrt{\frac{1}{n}}\right] = [-0,620; 0,620]$

Aucun coefficient n'est à l'extérieur de l'intervalle, ils sont donc non significativement différents de 0.

e) Soit le calcul suivant :

$$Q' = n(n+2) \sum_{k=1}^3 \frac{r_k^2}{n-k} = 10 \times 12 \left[\frac{(0,021)^2}{9} + \frac{(-0,502)^2}{8} + \frac{(-0,348)^2}{7} \right] = 5,86$$

Cette statistique est à comparer au chi-deux lu dans la table à trois degrés de liberté ($5,86 < \chi_3^{0,95} = 7,81$), on accepte l'hypothèse H_0 de nullité des coefficients d'auto-corrélation. La série x_t est donc un bruit blanc, ce qui est cohérent avec les résultats du point b.

Exercice n° 3.2

Fichier C5EX3

Exemple d'analyse des FAC et FAP au CAC40

On demande de calculer et d'analyser les corrélogrammes simple et partiel de l'indice CAC40 en niveau et en différences premières (cf. exercice 5.3 du chapitre 5).

Solution

La figure 3.3 présente les corrélogrammes simple (*Autocorrelation*) et partiel (*Partial Correlation*) du CAC40.

Nous constatons que la fonction d'autocorrélation simple (colonne AC) décroît très lentement, cela est typique d'une série non stationnaire. En revanche, la fonction d'autocorrélation partielle (colonne PAC) a seul son premier terme significativement différent de 0 (l'intervalle de confiance est stylisé par les pointillés).

La statistique de Ljung-Box est calculée (colonne Q -Stat) pour chaque retard h , assortie de sa probabilité critique.

Par exemple calculons la Q -Stat pour $h = 4$:

$$Q' = n(n+2) \sum_{k=1}^4 \frac{r_k^2}{n-k}$$

$$= 1160 \times 1162 \left[\frac{0,989^2}{1159} + \frac{0,976^2}{1158} + \frac{0,964^2}{1157} + \frac{0,953^2}{1156} \right] = 4387,86$$

Cette statistique est à comparer au chi-deux lu dans la table à quatre degrés de liberté ($4387,8 > \chi_4^{0,95} = 9,48$), on rejette l'hypothèse H_0 , il existe au moins un coefficient d'autocorrélation significativement différent de 0 (évidemment ici le test est inutile, la simple lecture du corrélogramme nous fait rejeter l'hypothèse H_0).

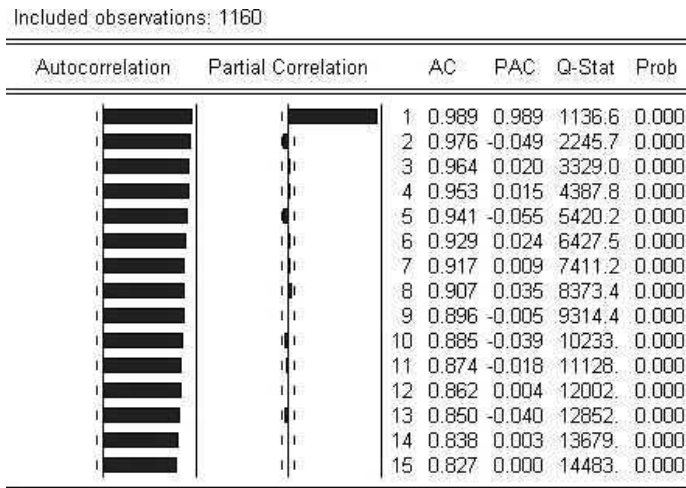


Figure 3.3 – Corrélogrammes simple et partiel du CAC40

Nous calculons ensuite, les fonctions d'autocorrélation simple et partielle sur le CAC40 en différences premières (figure 3.4). Nous constatons que – presque – tous les termes sont dans l'intervalle de confiance à l'exception des retards 1 et 9, et que la probabilité critique de la statistique de Ljung-Box est, pour presque tous les retards, supérieure à 0,05 (acceptation de H_0). Le CAC40 en différences premières est bien sûr stationnaire et peut être considéré ici comme un processus de marche au hasard.

Included observations: 1159

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.062	0.062	4.4625	0.035
		2 -0.010	-0.014	4.5772	0.101
		3 -0.039	-0.038	6.3597	0.095
		4 0.059	0.064	10.409	0.034
		5 -0.015	-0.024	10.665	0.058
		6 -0.020	-0.018	11.145	0.084
		7 -0.017	-0.010	11.493	0.119
		8 -0.008	-0.012	11.572	0.171
		9 0.044	0.046	13.813	0.129
		10 0.021	0.016	14.347	0.158
		11 0.018	0.016	14.712	0.196
		12 0.040	0.043	16.558	0.167
		13 -0.008	-0.018	16.634	0.217
		14 -0.051	-0.049	19.652	0.141
		15 -0.009	0.001	19.741	0.182

Figure 3.4 – Corrélogrammes simple et partiel de D (CAC40)

IV. La classe des processus aléatoires ARMA linéaires et stationnaires

Il est possible de définir la classe des processus ARMA à partir du théorème de décomposition des processus de Wold.

A. Le théorème de décomposition de Wold¹

Soit le processus centré réel ou complexe Z_t stationnaire et de variance finie. Il existe trois processus T_t, x_t, a_t qui vérifient les propriétés suivantes :

$$\bullet Z_t = T_t + x_t$$

où T_t et x_t sont deux processus orthogonaux (indépendants).

Le processus T_t est dit processus singulier (ou encore processus déterminable). Il s'agit d'une composante dont chaque valeur peut se calculer à partir d'une combinaison linéaire finie ou infinie de ses valeurs passées. C'est donc un processus dont nous pouvons déterminer exactement la prévision. Par opposi-

1. Wold, 1954.

tion au processus précédent, x_t porte le nom de processus indéterminable.

- a_t est un Bruit Blanc centré, $E[a_t] = 0$.

$$\bullet x_t = \sum_{j=\begin{cases} -\infty \\ 0 \end{cases}}^{+\infty} \psi_j a_{t-j} \quad \text{avec } \psi_0 = 1, \psi_j \in R \quad (1)$$

j varie de $-\infty$ à $+\infty$ dans le cas complexe, j varie de 0 à $+\infty$ dans le cas réel

et $\sum_{j=\begin{cases} -\infty \\ 0 \end{cases}}^{+\infty} \psi_j^2 < +\infty$.

Un processus x_t , centré, stationnaire d'ordre 2, peut s'écrire selon l'expression (1) précédente.

La condition sur les ψ_j assure l'existence des moments d'ordre 2 du processus, on dit que x_t est convergent en moyenne quadratique.

En effet, considérons la formule (1) sous sa forme développée dans le cas réel :

$$x_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots$$

et calculons, par exemple, la variance de x_t

$$V[x_t] = E[(x_t - E(x_t))^2]. \text{ Or } E(x_t) = \sum_j \psi_j E(a_{t-j}) = 0.$$

D'où

$$\begin{aligned} V[x_t] &= E[x_t^2] \\ &= E[a_t^2 + \psi_1^2 a_{t-1}^2 + \psi_2^2 a_{t-2}^2 + \dots + \psi_0 \psi_1 a_t a_{t-1} + \dots +] \end{aligned}$$

Supposons que la suite des ψ_j est telle que :

$$\psi_0 = 1 \quad \psi_1 = \sqrt{\frac{1}{2}} \quad \dots \quad \psi_j = \psi_1^j$$

La somme $(1 + \psi_1^2 + \psi_2^2 + \dots) = (1 + \psi_1^2 + \psi_1^4 + \dots)$ est, dans ce cas, la somme d'une progression géométrique de raison $\frac{1}{2}$, elle est convergente et la variance de x_t est finie.

Supposons que la suite des ψ_j est telle que :

$$\psi_0 = 1 \quad \psi_1 = \sqrt{2} \quad \dots \quad \psi_j = \psi_1^j$$

La somme $(1 + \psi_1^2 + \psi_2^2 + \dots)$ est, à présent, la somme des termes d'une progression géométrique de raison 2 ; elle est non convergente et la variance de x_t est infinie.

Cette condition de l'équation (1) sur les ψ_j est également importante dans la suite de l'exposé¹ car sous cette forme le processus x_t peut être considéré comme la sortie d'un filtre linéaire dont l'entrée est un Bruit Blanc a_t et de fonction de réponse impulsionnelle les valeurs successives des ψ_j . Or ce filtre est stable si la fonction de réponse est constituée d'une suite de valeurs dont la somme est convergente.

Enfin, le processus x_t peut être réécrit en utilisant l'opérateur de retard B tel que :

$$B^j y_t = y_{t-j} \quad \text{où } y_t \text{ est un processus quelconque.}$$

On aura dans le cas réel :

$$x_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots$$

$$x_t = a_t + \psi_1 B a_t + \psi_2 B^2 a_t + \dots$$

$$x_t = (1 + \psi_1 B + \psi_2 B^2 + \dots) a_t$$

$$x_t = \psi(B) a_t \quad (2)$$

$\psi(B)$ porte le nom de polynôme d'opérateur et on peut lui appliquer la plupart des théorèmes courants de mathématiques.

B. Propriétés de l'opérateur retard

Examinons les propriétés de l'opérateur B .

$Ba = a$, l'opérateur d'une constante a est une constante

$$B^0 x_t = x_t$$

$(1 - B)^2 x_t \neq (1 - B^2) x_t$ en effet :

$$(1 - B)^2 x_t = (1 - 2B + B^2) x_t = x_t - 2x_{t-1} + x_{t-2} \neq x_t - x_{t-2}$$

$$(B^i + B^j) x_t = B^i x_t + B^j x_t = x_{t-i} + x_{t-j}$$

$$B^i (B^j x_t) = B^i x_{t-j} = x_{t-i-j}, \text{ de même } B^i (B^j x_t) = B^{i+j} x_t = x_{t-i-j}$$

$$B^{-i} x_t = x_{t+i}$$

Si $|a| < 1$ la somme infinie $(1 + aB + a^2 B^2 + a^3 B^3 + \dots) x_t = \frac{x_t}{(1 - aB)}$.

Démonstration : multiplions de chaque côté la relation précédente par $(1 - aB)$.

$$(1 - aB)(1 + aB + a^2 B^2 + a^3 B^3 + \dots) x_t = (1 - aB) \frac{x_t}{(1 - aB)} = x_t$$

$$(1 - aB + aB - a^2 B^2 + a^2 B^2 - a^3 B^3 + \dots) x_t = x_t$$

1. Cf. chapitre 4

Comme $|a| < 1$, $a^n B^n x_t$ converge vers 0 lorsque n tend vers l'infini. Ainsi, les deux membres de la relation sont égaux.

C. Définition des processus ARMA

1) Les processus MA et AR

Le processus stationnaire défini par les formules (1) ou (2) est appelé processus Moyenne Mobile infini. Il est noté $MA(\infty)$.

Considérons le processus simple suivant pour lequel $\psi(B)$ est un polynôme de degré 1.

$$x_t = (1 + \psi_1 B)a_t \text{ et par commodité posons } \psi'_1 = -\psi_1.$$

Le polynôme $(1 - \psi'_1 B)^{-1}$ est inversible si et seulement si $|\psi'_1| < 1$. Dans le cas où $|\psi'_1| < 1$ on peut écrire¹ :

$$(1 - \psi'_1 B)^{-1} = \frac{1}{1 - \psi'_1 B} = \sum_{i=0}^{\infty} \psi_1^i B^i$$

Nous pouvons calculer :

$$a_t = (1 - \psi'_1 B)^{-1} x_t$$

$$a_t = (1 + \psi'_1 B + \psi_1^2 B^2 + \dots)x_t$$

$$a_t = \sum_{j=0}^{+\infty} \psi_1^j x_{t-j}. \quad (3)$$

Cette formule (3) est équivalente à la formule (1) pour l'écriture du processus a_t . Pour que a_t soit un processus stationnaire il faut donc que $\sum_{j=0}^{+\infty} \psi_1^j$ soit convergente.

Il s'agit d'une progression géométrique de raison ψ_1 et sa convergence est assurée lorsque $|\psi'_1| < 1$, ou encore lorsque $|\psi_1| < 1$.

Considérons la formule (3) précédente sous sa forme développée.

$$a_t = x_t + \psi'_1 x_{t-1} + \psi_1^2 x_{t-2} + \dots + \psi_1^j x_{t-j} + \dots$$

et posons $\pi_j = \psi_1^j$. Alors :

$$a_t = x_t + \pi_1 x_{t-1} + \pi_2 x_{t-2} + \dots + \pi_j x_{t-j} + \dots$$

1. Cf. Gouriéroux et Monfort, 1990, p. 172-176.

$$\begin{aligned}
 a_t &= x_t + \pi_1 B x_t + \pi_2 B^2 x_t + \dots + \pi_j B^j x_t + \dots \\
 a_t &= (1 + \pi_1 B + \pi_2 B^2 + \dots + \pi_j B^j + \dots) x_t \\
 \mathbf{a_t} &= \mathbf{\pi(B)x_t} \tag{4}
 \end{aligned}$$

On appelle processus autorégressif noté AR, l'écriture (4) du processus et lorsque le polynôme d'opérateur particulier $\pi(B)$ comporte une infinité de termes, il est dit processus AR(∞).

Enfin, observons que la formule (4) de l'AR(∞) découle de l'écriture $x_t = (1 + \psi_1 B)a_t$. Or cette écriture est celle d'une MA d'ordre 1 puisque le polynôme d'opérateur a pour degré 1. Nous pouvons également montrer (en utilisant l'écriture (4) avec un polynôme $\pi(B)$ de degré 1) que l'écriture d'un MA(∞) peut être issue d'un AR(1) d'où la correspondance :

$$AR(1) \implies MA(\infty)$$

$$MA(1) \implies AR(\infty)$$

Considérons la formule (4) précédente $a_t = \pi(B)x_t$ et multiplions les deux membres par $\psi(B)$:

$$\psi(B)a_t = \psi(B)\pi(B)x_t. \text{ Or, } \psi(B)a_t = x_t, \text{ d'après la relation (2), d'où :}$$

$$\psi(B)\pi(B) = 1$$

Nous pouvons résumer les résultats précédents selon la figure 3.5.

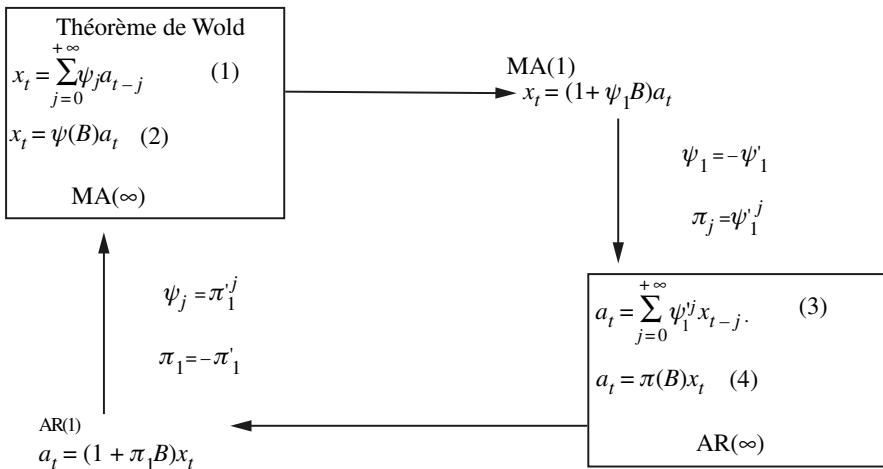


Figure 3.5 – Relation entre théorème de Wold, processus AR et processus MA

2) Les processus ARMA non saisonniers

Les processus ARMA – mélanges de processus AR et MA – sont nécessairement, en pratique, finis. Ils sont issus de deux processus :

- Le processus MA(q) défini à partir de la formule (1) :

$$\begin{aligned}x_t &= a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \\x_t &= (1 - \theta_1 B - \dots - \theta_q B^q) a_t \\x_t &= \theta_q(B) a_t\end{aligned}\tag{5}$$

θ_j , élément du polynôme d'opérateur fini $\theta_q(B)$, joue le rôle de $-\psi_j$ dans la formule (1).

- Le processus AR(p) défini à partir de la formule (4) :

$$\begin{aligned}x_t &= \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + a_t \\ \text{Il s'écrit aussi : } (1 - \phi_1 B - \dots - \phi_p B^p) x_t &= a_t \\ \phi_p(B) x_t &= a_t\end{aligned}\tag{6}$$

ϕ_j élément du polynôme d'opérateur fini $\phi_p(B)$ joue le rôle de $-\pi_j$ dans la formule (4).

- Le processus ARMA(p, q) s'écrit alors :

$$\phi_p(B) x_t = \theta_q(B) a_t\tag{7}$$

Dans cette écriture le processus est centré ($E[x_t] = 0$). Plus généralement on peut considérer des processus avec constante :

$$\phi_p(B) x_t = \mu + \theta_q(B) a_t\tag{7'}$$

On remarque que :

$$\begin{aligned}\text{ARMA}(0, q) &\equiv \text{MA}(q) \\ \text{ARMA}(p, 0) &\equiv \text{AR}(p)\end{aligned}$$

À partir de la formule (7), nous pouvons écrire à condition que $\phi_p^{-1}(B)$ existe :

$x_t = \frac{\theta_q(B)}{\phi_p(B)} a_t$ ce qui montre qu'il est possible avec des degrés peu élevés de p et q de définir un nombre assez important de modèles (c'est le principe dit de « parcimonie de Box et Jenkins »).

Exemples d'écritures :

$$- \text{AR}(1) : (1 - \phi_1 B)x_t = a_t \iff x_t = x_{t-1} + a_t$$

$$- \text{AR}(2) : (1 - \phi_1 B - \phi_2 B^2)x_t = a_t \iff x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + a_t$$

$$- \text{MA}(1) : x_t = (1 - \theta_1 B)a_t \iff x_t = a_t - \theta_1 a_{t-1}$$

$$- \text{MA}(2) : x_t = (1 - \theta_1 B - \theta_2 B^2)a_t \iff x_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

$$- \text{ARMA}(1, 1) : (1 - \phi_1 B)x_t = (1 - \theta_1 B)a_t \iff x_t = \phi_1 x_{t-1} + a_t - \theta_1 a_{t-1}$$

- ARMA(1, 2) avec constante :

$$(1 - \phi_1 B)x_t = \mu + (1 - \theta_1 B - \theta_2 B^2)a_t$$

$$\iff x_t = \mu + \phi_1 x_{t-1} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

D. La stationnarité et l'inversibilité des processus

1) Conditions de stationnarité et d'inversibilité

Considérons le MA(1) précédent de la formule (2) : $x_t = (1 + \psi_1 B)a_t$.

Nous avons vu qu'il peut s'écrire selon la forme (3) d'un AR (∞) sous la condition que $|\psi_1| < 1$, condition qui permet d'inverser le polynôme $(1 + \psi_1 B)$.

Dans ce cas, la racine en B du polynôme $(1 + \psi_1 B)$ est $B = -\frac{1}{\psi_1}$ et comme

$|\psi_1| < 1$ alors $|B| > 1$. Cette condition est appelée condition d'inversibilité du processus. Ce résultat peut être généralisé pour le polynôme $\theta(B)$ des processus MA et pour le polynôme $\phi(B)$ des processus AR.

Le théorème¹ suivant peut maintenant être énoncé :

- Un processus AR est toujours inversible. Il est stationnaire lorsque les racines de $\phi_p(B) = 0$ sont à l'extérieur du cercle unité du plan complexe.

$$(1 - \phi_1 B - \dots - \phi_p B^p) = 0 \iff \Phi_p(B) = 0 \iff \prod_{j=1}^p (1 - \mu_j B_j) = 0$$

Ces conditions de stationnarité se ramènent à $|\mu_j| < 1 \forall j$.

- Un processus MA est toujours stationnaire. Il est inversible si les racines de $\theta_q(B) = 0$ sont à l'extérieur du cercle unité du plan complexe.

$$(1 - \theta_1 B - \dots - \theta_q B^q) = 0 \iff \Theta_q(B) = 0 \iff \prod_{j=1}^q (1 - \gamma_j B_j) = 0$$

Ces conditions d'inversibilité se ramènent à $|\gamma_j| < 1 \forall j$.

1. Il s'agit d'une utilisation du théorème de Doob, 1953, p. 577.

• Les conditions de stationnarité et d'inversibilité de l'ARMA sont donc respectivement données par la partie AR et la partie MA de l'ARMA.

Ce résultat, sous forme de théorème, joue un rôle crucial dans l'utilisation ultérieure des processus ARMA car il conditionne l'existence même de ces processus, d'où les différentes techniques de recherche de ces conditions.

2) Recherche des conditions de stationnarité et d'inversibilité des processus ARMA

Deux cas se présentent concrètement pour cette recherche :

– *le processus est connu* (i.e. les paramètres ont déjà été estimés). Dans ce cas, la recherche de la stationnarité et de l'inversibilité permet de vérifier la validité du processus générateur ARMA de la chronique. Le problème mathématique est le calcul des racines de polynômes. Lorsque ces derniers sont de degré peu élevé (≤ 2), le résultat est immédiat. Pour des ordres plus importants, il est indispensable de recourir à des algorithmes de recherche de racines des polynômes : le plus souvent celui de Newton-Raphson ;

– *le processus est inconnu* (dans le sens où il n'a pas encore été estimé). Il convient de déterminer, dans ce cas, les plages de variations des paramètres qui conditionnent l'existence des processus ARMA. Ceci ne peut être effectué que pour des processus où les degrés des polynômes sont faibles.

Pour des polynômes de degré 1, le résultat est simple à établir.

Soit le processus AR(1) :

$$\phi_1(B)x_t = a_t \iff (1 - \phi_1 B)x_t = a_t \iff x_t = \phi_1 x_{t-1} + a_t$$

D'après le théorème précédent (ce processus est toujours inversible) la stationnarité requiert la recherche de la racine $1 - \phi_1 B = 0 \iff B = \frac{1}{\phi_1}$.

Le processus est stationnaire si et seulement si :

$$|B| > 1 \iff \left| \frac{1}{\phi_1} \right| > 1 \iff -1 < \phi_1 < 0 \text{ et } 0 < \phi_1 < 1.$$

Un calcul identique donne les conditions d'inversibilité d'un MA(1).

Pour des polynômes de degrés 2 ou 3, nous pouvons opérer de la manière suivante.

Considérons les racines du polynôme $\theta_q(B)$:

$$\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q = 0$$

qui s'écrit aussi :

$$B^{-q} \theta_q(B) = B^{-q} - \theta_1 B^{-q+1} - \dots - \theta_q = 0$$

soit en posant $F^q = B^{-q}$

$$F^q - \theta_1 F^{q-1} - \dots - \theta_q = 0$$

on obtient ainsi l'équation caractéristique de l'équation de récurrence homogène :

$$a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} = 0.$$

Si les racines de $\theta_q(B)$ sont à l'extérieur du cercle unité, les racines du polynôme en F sont à l'intérieur du cercle unité. Ce problème se ramène donc à la recherche des racines d'une équation caractéristique.

On peut utiliser le théorème de Routh - Samuelson¹.

Soit l'équation homogène suivante :

$$y_t + b_1 y_{t-1} + \dots + b_n y_{t-n} = 0 \quad (b_0 = 1)$$

et son équation caractéristique $b_0 r^n + b_1 r^{n-1} + \dots + b_n = 0$ ($b_0 = 1$), on forme les sommes suivantes :

$$\left\{ \begin{array}{l} \bar{b}_0 = \sum_{i=0}^n b_i \\ \bar{b}_1 = \sum_{i=0}^n b_i (n - 2i) \\ \vdots \\ \bar{b}_j = \sum_{i=0}^n b_i \sum_{k=0}^n C_{j-k}^{n-1} (-1)^k C_k^i \\ \vdots \\ \bar{b}_n = 1 - b_1 + b_2 - \dots + (-1)^{n-1} b_{n-1} + (-1)^n b_n \end{array} \right.$$

avec $C_k^i = 0$ pour $\begin{cases} k < i \\ k > i \end{cases}$

et la matrice M :

$$M = \begin{bmatrix} \bar{b}_1 & \bar{b}_3 & \bar{b}_5 & \dots & \bar{b}_j & \dots \\ \bar{b}_0 & \bar{b}_2 & \bar{b}_4 & \dots & \bar{b}_{j-1} & \dots \\ 0 & \bar{b}_1 & \bar{b}_3 & \dots & & \\ 0 & \bar{b}_0 & \bar{b}_2 & \dots & & \\ 0 & 0 & \bar{b}_0 & \dots & & \\ 0 & 0 & 0 & \dots & & \\ \vdots & \vdots & & & & \\ 0 & 0 & & & \bar{b}_0 & \\ \vdots & \vdots & & & & \end{bmatrix}$$

1. Samuelson, 1971.

Théorème

Une condition nécessaire et suffisante pour que les racines de l'équation caractéristique de l'équation homogène soient en module à l'intérieur du cercle unité est :

$$\left\{ \begin{array}{l} \bar{b}_0 > 0 \\ \Delta_1 > 0 \\ \Delta_2 > 0 \\ \vdots \\ \Delta_i > 0 \\ \vdots \\ \Delta_n > 0 \end{array} \right.$$

où les Δ_i sont les mineurs principaux dans l'ordre naturel de la matrice M .

3) Applications

- Soit le processus AR(1) : $\phi_1(B)x_t = a_t \iff (1 - \phi_1 B)x_t = a_t$

On a le polynôme : $1 - \phi_1 B = 0$

$$B^{-1} - \phi_1 = 0$$

$$r + b_1 = 0$$

$$b_0 = 1 \text{ et } b_1 = -\phi_1 \quad (n = 1)$$

$$\bar{b}_0 = \sum_{i=0}^{i=1} b_i = b_0 + b_1 = 1 - \phi_1$$

$$\begin{aligned} \bar{b}_1 &= \sum_{i=0}^{i=1} b_i(n - 2i) = b_0(n - 0) + b_1(n - 2) = b_0 - b_1 \\ &= 1 - \phi_1(-1) = 1 + \phi_1 \end{aligned}$$

d'où $M = \begin{bmatrix} 1 + \phi_1 \\ 1 - \phi_1 \end{bmatrix}$ et les conditions $\left. \begin{array}{l} \bar{b}_0 > 0 \iff 1 - \phi_1 > 0 \\ \Delta_1 > 0 \iff 1 + \phi_1 > 0 \end{array} \right\} |\phi_1| < 1$

- Typologie des solutions d'un processus AR(1) en fonction de la valeur de ϕ_1

1) $|\phi_1| < 1$

$$\begin{aligned} x_t &= \frac{1}{1 - \phi_1 B} a_t \\ &= (1 + \phi_1 B + \phi_1^2 B^2 + \dots) a_t = \sum_{i=0}^{\infty} \phi_1^i a_{t-i} \end{aligned}$$

Solution stationnaire par construction :

2) $|\phi_1| = 1 \implies x_t$ est non stationnaire (processus de marche au hasard défini au chapitre 5) ;

3) $|\phi_1| > 1 \implies x_t$ est non stationnaire explosif.

• Soit le processus AR (2) : $(1 - \phi_1 B - \phi_2 B^2)x_t = a_t$

Cherchons les conditions pour que :

$$(1 - \phi_1 B - \phi_2 B^2) = 0$$

ou encore :

$$B^{-2} - \phi_1 B^{-1} - \phi_2 = 0$$

soit :

$$r^2 + b_1 r + b_2 = 0 \text{ avec } n = 2$$

$$b_0 = 1 ; \quad b_1 = -\phi_1 ; \quad b_2 = -\phi_2$$

$$\bar{b}_0 = \sum_{i=0}^n b_i = \sum_{i=0}^2 = b_0 + b_1 + b_2 = 1 - \phi_1 - \phi_2$$

$$\bar{b}_1 = \sum_{i=0}^n b_i(n - 2i) = 2b_0 - 2b_2 = 2 + 2\phi_2$$

$$\bar{b}_2 = 1 - b_1 + b_2 = 1 + \phi_1 - \phi_2$$

$$M = \begin{bmatrix} 2 + 2\phi_2 & 0 & 0 & 0 & \dots \\ 1 - \phi_1 - \phi_2 & 1 + \phi_1 - \phi_2 & 0 & 0 & \dots \\ 0 & 2 + 2\phi_2 & 0 & 0 & \dots \\ 0 & 1 - \phi_1 - \phi_2 & 1 + \phi_1 - \phi_2 & 0 & \dots \\ 0 & 0 & 2 + 2\phi_2 & 0 & \dots \\ & & 1 - \phi_1 - \phi_2 & 0 & \dots \end{bmatrix}$$

D'où les conditions :

$$\begin{cases} \bar{b}_0 = 1 - \phi_1 - \phi_2 > 0 \\ \Delta_1 = 2 + 2\phi_2 > 0 \\ \Delta_2 = (2 + 2\phi_2)(1 + \phi_1 - \phi_2) > 0 \end{cases}$$

qui peuvent s'écrire aussi :

$$\begin{cases} 1 - \phi_1 - \phi_2 > 0 \\ 1 + \phi_2 > 0 \\ 1 + \phi_1 - \phi_2 > 0 \end{cases}$$

• Typologie des solutions

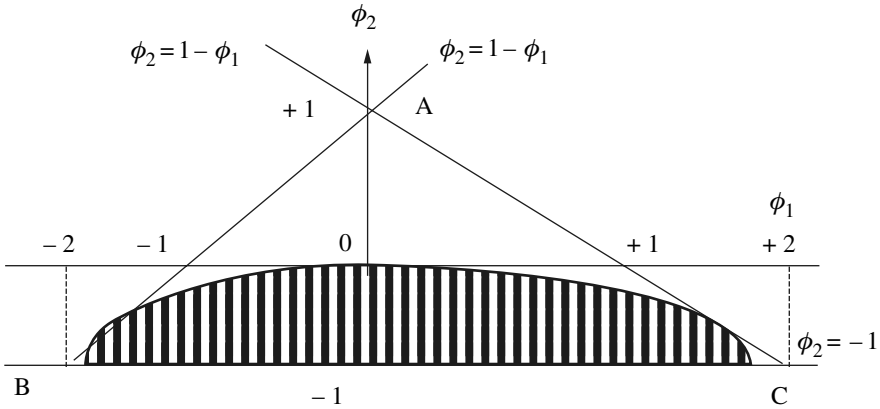


Figure 3.6 – Solution géométrique

La figure 3.6 illustre la résolution graphique de la recherche des conditions de stationnarité d'un AR(2). L'intérieur du triangle A B C représente l'ensemble des couples (ϕ_1, ϕ_2) pour lesquels le processus AR(2) est stationnaire. Il s'agit aussi de coordonnées des racines de l'équation : $(1 - \phi_1 B - \phi_2 B^2) = 0$ qui possède des racines réelles et complexes. La zone intérieure hachurée correspond au cas des racines complexes.

Ce résultat est identique pour les couples $(\theta_1$ et $\theta_2)$ du processus MA(2).

E. Les processus ARMA saisonniers

Il existe deux types de processus ARMA :

- les processus ARMA non saisonniers : ARMA(p, q) définis précédemment
- les processus ARMA saisonniers : ARMA $_{s, s'}(P, Q)$

Pour ces derniers, s et s' représentent des pas de temps constants et différents de un. On leur donne le nom de « période de la saisonnalité » du processus AR stationnaire pour s et de « période de la saisonnalité » du processus MA inversible pour s' .

P est l'ordre du processus AR saisonnier.

Q est l'ordre du processus MA saisonnier.

Par extension de l'écriture des ARMA non saisonniers, les processus ARMA $_{s, s'}(P, Q)$ prennent la forme :

$$x_t = \phi_s x_{t-s} + \phi_{2s} x_{t-2s} + \dots + \phi_{Ps} x_{t-Ps} + a_t - \theta_{s'} a_{t-s'} - \theta_{2s'} a_{t-2s'} - \dots - \theta_{Qs'} a_{t-Qs'}$$

Par exemple :

Soit le processus $\text{ARMA}_{12,6}(1, 2)$, il s'écrit :

$$x_t = \phi_{12}x_{t-12} + a_t - \theta_6a_{t-6} - \theta_{12}a_{t-12}$$

Il s'agit de la même écriture qu'un ARMA non saisonnier, mais l'ordre s'applique uniquement aux valeurs de $t = s$ et à ses multiples. En effet, considérons le processus : $\text{AR}_{12}(1)$. Il s'écrit : $x_t = \phi_{12}x_{t-12} + a_t$

Soit : $x_t = 0 \times x_{t-1} + \dots + 0 \times x_{t-11} + \phi_{12}x_{t-12} + a_t$

Ce processus $\text{AR}_{12}(1)$ est équivalent à un $\text{AR}(12)$ « troué »¹ avec $\phi_1 = \dots \phi_{11} = 0$

Sous forme de polynômes d'opérateurs le processus ARMA saisonnier s'écrit alors :

$$x_t - \phi_s B^s x_t - \dots - \phi_{Ps} B^{Ps} x_t = a_t - \theta_{s'} B^{s'} a_t - \dots - \theta_{Qs'} B^{Qs'} a_t$$

$$(1 - \phi_s B^s - \dots - \phi_{Ps} B^{Ps}) x_t = (1 - \theta_{s'} B^{s'} - \dots - \theta_{Qs'} B^{Qs'}) a_t$$

$$\phi_P(B^s) x_t = \theta_Q(B^{s'}) a_t$$

Ils peuvent s'écrire aussi avec une constante dans le second membre :

$$\phi_P(B^s) x_t = \mu + \theta_Q(B^{s'}) a_t$$

Les conditions d'existence des ARMA purement saisonniers sont identiques à celles déjà indiquées pour les ARMA non saisonniers.

Remarque : les processus ARMA saisonniers sont aussi nommés par certains auteurs processus SARMA.

F. Les processus ARMA non saisonniers et saisonniers à la fois

Il en existe de deux sortes :

- les processus ARMA additifs notés : $\text{ARMA}(p, q) + \text{ARMA}_{ss'}(P, Q)$. Ils s'écrivent : $[\phi_p(B) + \phi_P(B^s) - 1] x_t = [\theta_q(B) + \theta_Q(B^{s'}) - 1] a_t$. Dans la pratique, ils sont rarement employés ;
- les processus ARMA multiplicatifs notés : $\text{ARMA}(p, q) \times \text{ARMA}_{ss'}(P, Q)$. Soit sous forme d'opérateurs : $\phi_p(B)\phi_P(B^s)x_t = \theta_q(B)\theta_Q(B^{s'})a_t$ (avec $s > p$ et $s' > q$). Ils sont équivalents à des processus : $\text{ARMA}(p + Ps, q + Qs')$. Soit :

1. Un processus est qualifié de troué lorsque certains des retards (de type AR ou MA), de rang inférieur à l'ordre du processus, sont affectés d'un coefficient nul.

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - \phi_s B^s - \dots - \phi_{p_s} B^{p_s})x_t \\ = (1 - \theta_1 B - \dots - \theta_q B^q)(1 - \theta_{s'} B^{s'} - \dots - \theta_{q_{s'}} B^{q_{s'}})a_t$$

Ces processus ARMA non saisonniers et saisonniers à la fois, possèdent des conditions d'existence toujours identiques à celles des ARMA non saisonniers et peuvent exister avec une constante dans le second membre.

Exemples d'écriture des processus saisonniers :

$$\text{ARMA}(1,2) \times \text{ARMA}_{6,12}(1,1)$$

$$(1 - \phi_1 B)(1 - \phi_6 B^6)x_t = (1 - \theta_1 B - \theta_2 B^2)(1 - \theta_{12} B^{12})a_t$$

$$(1 - \phi_1 B - \phi_6 B^6 + \phi_1 \phi_6 B^7)x_t \\ = (1 - \theta_1 B - \theta_2 B^2 - \theta_{12} B^{12} + \theta_1 \theta_{12} B^{13} + \theta_2 \theta_{12} B^{14})a_t$$

$$\iff x_t = \phi_1 x_{t-1} + \phi_6 x_{t-6} - \phi_1 \phi_6 x_{t-7} \\ + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_{12} a_{t-12} + \theta_1 \theta_{12} a_{t-13} + \theta_2 \theta_{12} a_{t-14}$$

Remarque : ARMA(1,2) + ARMA_{6,12}(1,1) s'écrit :

$$[(1 - \phi_1 B)(1 - \phi_6 B^6) - 1]x_t = [(1 - \theta_1 B - \theta_2 B^2)(1 - \theta_{12} B^{12}) - 1]a_t$$

Soit :

$$[(1 - \phi_1 B - \phi_6 B^6)]x_t = (1 - \theta_1 B - \theta_2 B^2 - \theta_{12} B^{12})a_t$$

Exercice n° 3.3

Recherche des conditions de stationnarité et d'inversibilité des processus

- MA(1) : $x_t = a_t + 0,8a_{t-1}$

$$x_t = a_t + 0,8a_{t-1} = (1 + 0,8B)a_t$$

$$\text{On a } 1 + 0,8B = 0 \implies B = -1,25 \text{ d'où } |B| > 1$$

Le processus est inversible.

- MA(1) : $x_t = a_t - 0,8a_{t-1}$

$$x_t = a_t - 0,8a_{t-1} = (1 - 0,8B)a_t$$

$$\text{On a } 1 - 0,8B = 0 \implies B = 1,25 \text{ d'où } |B| > 1$$

Le processus est inversible.

- MA(2) : $x_t = a_t + 0,6a_{t-1} - 0,3a_{t-2}$

$$x_t = a_t + 0,6a_{t-1} - 0,3a_{t-2} = (1 + 0,6B - 0,3B^2)a_t$$

$$\text{On a } 1 + 0,6B - 0,3B^2 = 0 \implies B_1 = 3,08 \text{ et } B_2 = -1,08 \text{ d'où } |B_i| > 1$$

Le processus est inversible, nous pouvons vérifier que les paramètres $-0,6$ et $0,3$ sont dans la zone non hachurée (racines réelles) de la figure 3.4.

- AR(1) : $x_t = 0,9x_{t-1} + a_t$

$$(1 - 0,9B)x_t = a_t \implies (1 - 0,9B) = 0 \implies B = 1,11 \text{ d'où } |B| > 1$$

Le processus est stationnaire.

- AR(1) : $x_t = -0,9x_{t-1} + a_t$

$$(1 + 0,9B)x_t = a_t \implies (1 + 0,9B) = 0 \implies B = -1,11 \text{ d'où } |B| > 1$$

Le processus est stationnaire.

- AR(2) : $x_t = 0,9x_{t-1} - 0,7x_{t-2} + a_t$

$$(1 - 0,9B + 0,7B^2)x_t = a_t \implies (1 - 0,9B + 0,7B^2) = 0$$

$$\implies B = \frac{+9 \pm i\sqrt{199}}{14} = +\frac{9}{14} \pm i\frac{\sqrt{199}}{14}$$

$$\rho = \sqrt{\left(\frac{9}{14}\right)^2 + \left(\frac{\sqrt{199}}{14}\right)^2} \approx 1,20 \implies \rho > 1 \text{ ce processus est donc stationnaire,}$$

nous pouvons vérifier que les paramètres 0,9 et -0,7 sont dans la zone hachurée (cas complexe) de la figure 3.4.

- AR(2) : $x_t = x_{t-1} - 0,1x_{t-2} + a_t$

$$(1 - B + 0,1B^2)x_t = a_t \implies (1 - B + 0,1B^2) = 0 \implies B = \frac{+1 \pm \sqrt{0,6}}{0,2}$$

$$\implies B_1 = 8,87 \text{ et } B_2 = 1,13 \text{ d'où } |B_i| > 1.$$

Le processus est stationnaire, les paramètres 1 et -0,1 sont dans la zone non hachurée (racines réelles) de la figure 3.4

- MA(2) : $x_t = +a_t - 0,5a_{t-1} + 0,1a_{t-2}$

$$x_t = (1 - 0,5B + 0,1B^2)a_t \implies (1 - 0,5B + 0,1B^2) = 0$$

$$\implies \Delta = 0,25 - 0,4 = -0,15$$

$$\implies B = \frac{0,5 \pm i\sqrt{0,15}}{0,2} = +\frac{0,5}{0,2} \pm i\frac{\sqrt{0,15}}{0,2}$$

$$\rho = \sqrt{\left(\frac{0,5}{0,2}\right)^2 + \left(\frac{\sqrt{0,15}}{0,2}\right)^2} \approx 3,16 \implies \rho > 1 \text{ ce processus est donc inversible, nous}$$

pouvons vérifier que les paramètres 0,5 et -0,1 sont dans la zone hachurée (racines complexes) de la figure 3.4.

- ARMA(1,1) : $x_t = 0,8x_{t-1} + a_t - 0,7a_{t-1}$

$$(1 - 0,8B)x_t = (1 - 0,7B)a_t$$

$$\implies 1 - 0,8B = 0 \text{ Soit } B = 1,25 \text{ d'où } |B| > 1$$

$$1 - 0,7B = 0 \text{ Soit } B = 1,43 \text{ d'où } |B| > 1$$

Ce processus est stationnaire et inversible.

- ARMA(3,2) : $x_t = 2,5x_{t-1} - 0,5x_{t-2} - x_{t-3} + a_t + a_{t-1} + 2a_{t-2}$

1) Écrire le processus sous forme de polynôme d'opérateurs

$$x_t - 2,5x_{t-1} + 0,5x_{t-2} + x_{t-3} = at + a_{t-1} + 2a_{t-2}$$

$$(1 - 2,5B + 0,5B^2 + B^3)x_t = (1 + B + 2B^2)a_t$$

2) Le processus est-il stationnaire ?

Recherchons les racines de l'équation : $(1 - 2,5B + 0,5B^2 + B^3) = 0$

Il existe une racine évidente $B_1 = 1$, nous pouvons donc factoriser :

$$(1 - B)(1 - 1,5B - B^2) = 0$$

$$(1 - 1,5B - B^2) = 0 \implies \Delta = 6,25 \implies \text{deux racines : } -2 \text{ et } 0,5$$

Nous avons donc : $(1 - B)(1 - 2B)(1 + 0,5B) = 0$

Le processus est donc non stationnaire.

Ou encore :

$$1 - \phi_1 B - \phi_2 B^2 \text{ avec } \phi_1 = 1,5 \text{ et } \phi_2 = 1$$

$$1 - \phi_1 - \phi_2 = 1 - 1,5 - 1 = -1,5 < 0$$

$$1 + \phi_2 = 1 + 1 = 2 > 0$$

$$1 + \phi_1 - \phi_2 = 1 + 1,5 - 1 = 1,5 > 0$$

3) Ce processus est-il inversible ?

$$(1 + B + 2B^2) = 0 \implies \Delta = -7 = 7i^2$$

$$\implies B = \frac{-1 \pm i\sqrt{7}}{4} = -\frac{1}{4} \pm i\frac{\sqrt{7}}{4}$$

$$\rho = \sqrt{\left(-\frac{1}{4}\right)^2 + \left(\frac{\sqrt{7}}{4}\right)^2} = \sqrt{\frac{1}{16} + \frac{7}{16}} = \sqrt{\frac{8}{16}} \implies \rho < 1 \text{ ce processus est donc non inversible.}$$

$$\rho = \sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2} \approx 0,71$$

Ce processus est donc non stationnaire et non inversible.

• ARMA(0,1) \times ARMA_{4,0}(1,0) : $x_t = 0,8x_{t-4} + a_t + 0,8a_{t-1}$

$$(1 - 0,8B^4)x_t = (1 + 0,8B)a_t$$

Le processus est-il stationnaire ?

$$(1 - 0,8B^4) = 0 \implies (1,25 - B^4) = 0 \implies (1,057 - B)(1,057 + B)(1,118 - i^2 B^2) = 0$$

$$\text{Soit : } B_1 = 1,057 ; B_2 = -1,057 ; \rho = \sqrt{(1,58)^2} = 1,58 > 1$$

Le processus est donc stationnaire.

Le processus est-il inversible ?

$$(1 + 0,8B) = 0 \implies B = -1,25 \text{ d'où } |B| > 1$$

Ce processus est donc stationnaire et inversible.

Exercice n° 3.4

Exemples de génération de processus ARMA à l'aide d'Eviews

Soit les processus suivants :

$$\text{MA}(1) : x_t = a_t + 0,8a_{t-1}.$$

$$\text{MA}(2) : x_t = a_t + 0,6a_{t-1} - 0,3a_{t-2}$$

$$\text{AR}(1) : x_t = 0,9x_{t-1} + a_t$$

$$\text{AR}(2) : x_t = 0,9x_{t-1} - 0,7x_{t-2} + a_t$$

$$\text{ARMA}(1,1) : x_t = 0,9x_{t-1} + a_t + 0,8a_{t-1}$$

$$\text{ARMA}(0,1) \times \text{ARMA}_{4,0}(1,0) : (1 - 0,8B^4)x_t = (1 + 0,8B)a_t$$

(a_t est un bruit blanc gaussien de variance 4).

On demande de les générer (sur 500 observations) à partir du logiciel Eviews.

Solution

Les instructions Eviews permettant de générer les processus sont les suivantes (C3EX3.PRG) :

```
CREATE U 500 'Création de l'espace de travail
GENR a = NRND*4
GENR MA1 = a+.8*a(-1)
GENR MA2 = a+.6*a(-1) - 0.3*a(-2)
GENR AR1 = 0
SMPL 2 500
GENR AR1 = .9*AR1(-1) + a
SMPL 1 500
GENR AR2 = 0
SMPL 3 500
GENR AR2 = 0.9*AR2(-1) - 0.7*AR2(-2) + a
SMPL 1 500
GENR ARMA = 0
SMPL 2 500
GENR ARMA = .9*ARMA(-1) + a + 0.8*a(-1)
SMPL 1 500
GENR SARMA = 0
SMPL 5 500
GENR SARMA = 0.8*SARMA(-4) + a + 0.8*a(-1)
```

Exercice n° 3.5

Calculs des caractéristiques de différents processus et études des propriétés d'inversibilité et de stationnarité

Calculer l'espérance, la variance et la covariance des processus suivants et conclure sur leur stationnarité et inversibilité.

$$\varepsilon_t \rightarrow n.i.d.(0, \sigma_\varepsilon^2)$$

a) $x_t = \varepsilon_t - \varepsilon_{t-1}$

b) $x_t = b\varepsilon_t + c\varepsilon_{t-1}$

c) $x_t = \varepsilon_t \varepsilon_{t-1}$

d) $x_t = x_{t-1} + \varepsilon_t$ avec $x_0 = 0$

Étudier les conditions de stationnarité et d'inversibilité des processus suivants filtrés par leurs différences premières.

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

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

g) $x_t = at + b + \varepsilon_t$

Étudier les conditions de stationnarité et d'inversibilité du processus suivant filtré par leurs différences secondes.

h) $x_t = at^2 + bt + c + \varepsilon_t$

Solutions

a) $x_t = \varepsilon_t - \varepsilon_{t-1}$

$$E(x_t) = 0 ; \text{Var}(x_t) = \text{Var}(\varepsilon_t) + \text{Var}(\varepsilon_{t-1}) + \text{Cov}(\varepsilon_t, \varepsilon_{t-1}) = \sigma_\varepsilon^2 + \sigma_\varepsilon^2 = 2\sigma_\varepsilon^2$$

$$E(x_t x_{t-h}) = E[(\varepsilon_t - \varepsilon_{t-1})(\varepsilon_{t-h} - \varepsilon_{t-h-1})] \\ = E(\varepsilon_t \varepsilon_{t-h} - \varepsilon_t \varepsilon_{t-h-1} - \varepsilon_{t-1} \varepsilon_{t-h} + \varepsilon_{t-1} \varepsilon_{t-h-1}) = 0 \quad \forall h, h > 1$$

En effet, pour $h = 1 \implies$

$$E(x_t x_{t+1}) = E[(\varepsilon_t - \varepsilon_{t-1})(\varepsilon_{t+1} - \varepsilon_t)] = E(\varepsilon_t \varepsilon_{t+1} - \varepsilon_t \varepsilon_t - \varepsilon_{t-1} \varepsilon_{t+1} + \varepsilon_{t-1} \varepsilon_t) \\ = 0 + 0 - \sigma_\varepsilon^2 + 0 = -\sigma_\varepsilon^2$$

Ce processus est donc stationnaire.

Remarques : Ce processus s'écrit aussi $x_t = (1 - B)\varepsilon_t$; il s'agit d'un MA(1) stationnaire, par définition, et non inversible puisque le polynôme opérateur de sa partie MA a pour racine $B = 1$.

b) $x_t = b\varepsilon_t + c\varepsilon_{t-1}$

$$E(x_t) = 0 ; \text{Var}(x_t) = (b^2 + c^2)\sigma_\varepsilon^2$$

$$E(x_t x_{t-1}) = E[(b\varepsilon_t + c\varepsilon_{t-1})(b\varepsilon_{t-1} + c\varepsilon_{t-2})] \\ = E(b^2 \varepsilon_t \varepsilon_{t-1} + bc \varepsilon_t \varepsilon_{t-2} + cb \varepsilon_{t-1} \varepsilon_{t-1} + c^2 \varepsilon_{t-1} \varepsilon_{t-2}) \\ = 0 + 0 + cb\sigma_\varepsilon^2 + 0 = cb\sigma_\varepsilon^2$$

$$E(x_t x_{t-h}) = 0 \quad \forall h, h > 1$$

Ce processus est donc stationnaire.

Remarques : Ce processus s'écrit aussi $x_t = (b + cB)\varepsilon_t$; il s'agit d'un MA(1) stationnaire, par définition, et inversible si $|B| = b/c > 1$ soit si $c > 0$ et $b > c$ ou si $c < 0$ et $b < c$.

c) $x_t = \varepsilon_t \varepsilon_{t-1}$

$$E(x_t) = 0 ; \text{Var}(x_t) = \sigma_\varepsilon^2 \times \sigma_\varepsilon^2 = \sigma_\varepsilon^4$$

$$E(x_t x_{t-h}) = E[\varepsilon_t \varepsilon_{t-1} \varepsilon_{t-h} \varepsilon_{t-h-1}] = 0 \quad \forall h \neq 0.$$

Ce processus est donc stationnaire.

d) $x_t = x_{t-1} + \varepsilon_t$ avec $x_0 = 0$

Ce processus s'écrit par substitution : $x_t = x_0 + \sum_{j=1}^t \varepsilon_j = \sum_{j=1}^t \varepsilon_j$

$$E(x_t) = 0 ; \text{Var}(x_t) = \sigma_\varepsilon^2 + \sigma_\varepsilon^2 + \dots + \sigma_\varepsilon^2 = t\sigma_\varepsilon^2$$

$$\begin{aligned}
E(x_t x_{t-h}) &= E \left[\left(\sum_{j=1}^t \varepsilon_j \right) \left(\sum_{j=1}^{t-h} \varepsilon_j \right) \right] \quad \forall h, h > 0 \\
&= E[(\varepsilon_1 + \dots + \varepsilon_{t-1-h} + \varepsilon_{t-h} + \varepsilon_{t+1-h} + \dots + \varepsilon_t)(\varepsilon_1 + \dots + \varepsilon_{t-1-h} + \varepsilon_{t-h})] \\
&= E[\varepsilon_1^2 + \varepsilon_1 \varepsilon_2 + \dots + \varepsilon_1 \varepsilon_{t-h} \\
&\quad + \varepsilon_2^2 + \varepsilon_2 \varepsilon_1 + \dots + \varepsilon_2 \varepsilon_{t-h} \\
&\quad + \dots \\
&\quad + \varepsilon_{t-h}^2 + \varepsilon_{t-h} \varepsilon_1 + \dots \\
&\quad + \varepsilon_{t-h+1} \varepsilon_1 + \dots + \varepsilon_{t-h} \varepsilon_{t+1-h} \\
&\quad + \varepsilon_t \varepsilon_1 + \dots + \varepsilon_t \varepsilon_{t-h}] + \dots \\
&= E[\varepsilon_1^2] + \dots + E[\varepsilon_{t-h}^2] + 0 + \dots \\
&= \sigma_\varepsilon^2 + \sigma_\varepsilon^2 + \dots + \sigma_\varepsilon^2
\end{aligned}$$

$$E(x_t x_{t-h}) = (t-h)\sigma_\varepsilon^2$$

D'où $E(x_t x_{t-h}) = t\sigma_\varepsilon^2 = V(x_t)$ pour $h = 0$

$E(x_t x_{t-h}) = |t-h|\sigma_\varepsilon^2$ pour $h \neq 0$

Ce processus est donc non stationnaire.

Remarques :

- Pour h quelconque, nous pouvons écrire $E(x_t x_{t'}) = \sigma_\varepsilon^2 \text{Min}(t, t')$.
- Ce processus s'écrit aussi $(1-B)x_t = \varepsilon_t$, il s'agit d'un AR(1) inversible, par définition, et non stationnaire puisque le polynôme opérateur de sa partie AR a pour racine $B = 1$ (processus de marche au hasard).

e) $x_t = x_{t-1} + \varepsilon_t$ soit $\Delta x_t = \varepsilon_t$

En effet, $\Delta x_t = \varepsilon_t$ est un bruit blanc stationnaire.

f) $x_t = -x_{t-1} + \varepsilon_t$ soit $(1+B)x_t = \varepsilon_t$

$$T_t = \Delta x_t = (1-B)x_t$$

D'où $x_t = (1-B)^{-1}T_t$. En remplaçant ce résultat dans l'équation de départ, nous obtenons :

$$(1+B)(1-B)^{-1}T_t = \varepsilon_t$$

Soit : $(1+B)T_t = (1-B)\varepsilon_t \Rightarrow T_t = -T_{t-1} + \varepsilon_t - \varepsilon_{t-1}$

Ce processus a pour racine -1 pour son polynôme opérateur de sa partie AR et pour racine 1 pour son polynôme opérateur de sa partie MA. Il est donc non stationnaire et non inversible.

g) $x_t = at + b + \varepsilon_t$

$$v_t = \Delta x_t = (1-B)x_t = [at + b + \varepsilon_t - a(t-1) - b - \varepsilon_{t-1}]$$

$$v_t = a + (1-B)\varepsilon_t$$

Ce processus est un MA(1) avec constante donc stationnaire, par définition, et non inversible puisque le polynôme opérateur de sa partie MA a pour racine $B = 1$.

$$\mathbf{h)} \quad x_t = at^2 + bt + c + \varepsilon_t$$

$$v_t = \Delta^2 x_t = (1 - B)^2 x_t = (1 - B)[(1 - B)x_t] = (1 - B)w_t$$

$$w_t = x_t - x_{t-1} = at^2 + bt + c + \varepsilon_t - a(t-1)^2 - b(t-1) - c - \varepsilon_{t-1}$$

$$w_t = 2at + b - a + \varepsilon_t - \varepsilon_{t-1}$$

$$v_t = (1 - B)w_t = w_t - w_{t-1}$$

$$v_t = 2at + b - a + \varepsilon_t - \varepsilon_{t-1} - 2a(t-1) - b + a - \varepsilon_{t-1} + \varepsilon_{t-2}$$

$$v_t = 2a + \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2}$$

$$v_t = 2a + (1 - 2B + B^2)\varepsilon_t = 2a + (1 - B)^2\varepsilon_t$$

Ce processus est un MA(2) avec constante donc stationnaire, par définition, et non inversible puisque le polynôme opérateur de sa partie MA a pour racine double $B = 1$.

4. Les processus aléatoires dans le domaine des fréquences

Ce chapitre est consacré à l'analyse des séries temporelles dans le domaine des fréquences. Elle est souvent plus riche en terme d'interprétation, mais nécessite un recours à des techniques mathématiques plus complexes. Nous examinons en I. le rôle des filtres et leurs conséquences sur les séries chronologiques analysées. L'interprétation du spectre d'une chronique fait l'objet de la section II. Enfin, nous démontrons, en III., les formules des spectres relatives aux processus ARMA.

I. Filtrage linéaire d'un processus aléatoire

A. Définitions

Soit x_t un processus aléatoire à valeurs dans R ou dans C .

Un filtre est une transformation mathématique notée \mathfrak{F}' , qui associe au processus entrant x_t dans le filtre un processus sortant y_t :

$$x_t \xrightarrow{\text{input}} \mathfrak{F}' \xrightarrow{\text{output}} y_t.$$

Il existe de nombreuses familles de filtres qui diffèrent :

– selon l'échelle du temps (filtres en temps réel, en temps différé...);

- selon la mémoire du filtre : fini, infini ;
- selon le caractère stochastique ou déterministe du filtre ;
- selon sa forme mathématique.

Parmi ces familles de filtre, on utilise les filtres linéaires invariants dans le temps. Ce type de filtre, noté \mathfrak{F} , possède les deux propriétés suivantes :

- la correspondance $x_t \rightarrow y_t$ est linéaire : si x_t^1 et x_t^2 sont deux entrants et α, β deux nombres réels quelconques, alors $\mathfrak{F}(\alpha x_t^1 + \beta x_t^2) = \alpha \mathfrak{F}(x_t^1) + \beta \mathfrak{F}(x_t^2)$;
- la correspondance \mathfrak{F} est invariante par translation sur l'axe des temps (correspondance homogène) : $\mathfrak{F}[x_{t+t_0}] = y_{t+t_0}$.

Soit un processus entrant x_t et le processus sortant y_t d'un filtre \mathfrak{F} . Le pas d'échantillonnage du processus, noté Δ_T , est supposé constant et égal à 1. Soit $\dots \phi_{-\tau'}, \dots \phi_0, \dots \phi_\tau, \dots$ (une suite de coefficients de pondération à valeurs dans R où $\tau' \in N$ et $\tau \in N^*$).

Le filtre linéaire invariant dans le temps s'écrit alors :

$$\mathfrak{F}[x_t] = y_t = \sum_{D_\tau} \phi_\tau x_{t+\tau} \text{ ou } \mathfrak{F}[x_t] = y_t = \sum_{D_\tau} \phi_\tau x_{t-\tau}$$

où D_τ est le domaine de variation de τ . Si par exemple :

- $\tau \in Z$ et que $y_t = \sum_{\tau=-\infty}^{+\infty} \phi_\tau x_{t+\tau}$, le filtre est symétrique et infini ;
- $\tau \in [-q; +q]$, $q \in N$ et que $y_t = \sum_{\tau=-q}^{+q} \phi_\tau x_{t+\tau}$, le filtre est symétrique et fini ;
- $\tau \in N$ et que $y_t = \sum_{\tau=0}^{+\infty} \phi_\tau x_{t+\tau}$, le filtre est à un seul coté asymétrique et infini.

Les filtres de la famille \mathfrak{F} se différencient donc par le domaine de variation de τ et par les valeurs prises par ϕ_τ .

B. La fonction de réponse impulsionnelle et la fonction de réponse en fréquence du filtre

L'ensemble des valeurs prises par les coefficients de pondération, quand τ parcourt son domaine de définition, constitue le graphe de la fonction de réponse impulsionnelle du filtre. On l'appelle ainsi parce qu'elle correspond, formellement, à une entrée qui est une suite d'impulsions ou de fonctions de Dirac.

Soit x_t un processus de Dirac, il s'écrit par exemple :

$$\begin{aligned} x_t &= 1 \text{ pour } t = -1 \\ x_t &= 0 \text{ pour } t \neq -1. \end{aligned}$$

Appliquons-lui le filtre \mathfrak{S} , suivant :

$$\mathfrak{S}[x_t] = \sum_{\tau=1}^p \phi_{\tau} x_{t+\tau} = \phi \sum_{\tau=1}^p x_{t+\tau} \quad (p \in \mathbb{N})$$

On a :

$$\mathfrak{S}(x_t) = \phi \text{ si } t \in \{-p-1, \dots, -2\}$$

$$\mathfrak{S}(x_t) = 0 \text{ ailleurs.}$$

La figure 4.1 illustre la fonction de réponse impulsionnelle de ce filtre.

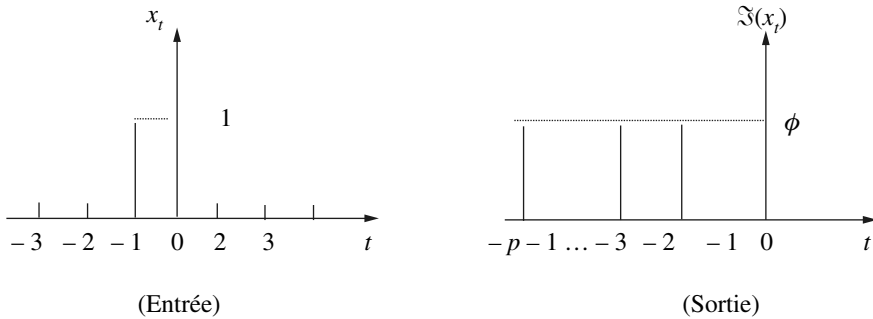


Figure 4.1 – Fonction de réponse impulsionnelle du filtre

La série des ϕ_{τ} (que l'on a supposés constants) correspond à la sortie du filtre lorsqu'une entrée égale à 1 est injectée pendant une unité de temps. C'est donc la réponse à une impulsion unitaire, ce qui justifie l'appellation de la fonction ; p est la longueur du filtre, appelée aussi la mémoire puisqu'après p pas de temps (de $-p-1$ à -2) la sortie devient nulle.

Considérons à présent l'action du filtre linéaire choisi symétrique à mémoire infinie, sur un processus d'entrée $x_t = e^{i\omega t}$

$$\mathfrak{S}[x_t] = \sum_{\tau=-\infty}^{+\infty} \phi_{\tau} e^{i\omega(t+\tau)}$$

On constate que :

$$\mathfrak{S}[x_t] = e^{i\omega t} \left\{ \sum_{\tau=-\infty}^{\infty} \phi_{\tau} e^{i\omega\tau} \right\}$$

$\mathfrak{S}[x_t] = e^{i\omega t} A_\omega$ que l'on peut schématiser :

$$x_t = e^{i\omega t} \xrightarrow{\mathfrak{S}} y_t = e^{i\omega t} A_\omega$$

Ce qui montre que le filtre \mathfrak{S} laisse passer au coefficient A_ω près la fonction $e^{i\omega t}$. Le facteur A_ω porte le nom de fonction de réponse en fréquence du filtre.

Généralisation

Considérons une fonction aléatoire $Z(t)$ continue à valeurs réelles ou complexes. À chacune des valeurs de $Z(t)$ est associée une pondération.

Les ϕ_τ précédents sont alors notés $\phi(\tau)$. Cette fonction $\phi(\tau)$, $\tau \in R$, est à valeur dans R ou dans C .

Supposons que la fonction $\phi(\tau)$ vérifie les conditions d'application du théorème de Plancherel ¹ et en particulier qu'elle soit de carré intégrable.

Dans ces conditions, nous pouvons démontrer que $\phi(\tau)$ admet une transformée de Fourier notée $A(\omega)$ appelée fonction de réponse en fréquence du filtre. Les deux fonctions $\phi(\tau)$ et $A(\omega)$ forment une paire de transformées de Fourier.

$$A(\omega) = \int_{-\infty}^{+\infty} \phi(\tau) e^{-i\omega\tau} d\tau$$

$$\phi(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} A(\omega) e^{+i\omega\tau} d\omega$$

$$= \int_{-\infty}^{+\infty} A(\lambda) e^{i2\pi\lambda\tau} d\lambda \text{ avec } \omega = 2\pi\lambda$$

1. Théorème de Plancherel : Soit $x(t)$ une fonction appartenant à l'espace des classes d'une fonction de carré intégrable sur R . Cet espace est muni de la norme de la convergence en moyenne quadratique : $\|x(t)\| = \left[\int_R [x(t)]^2 dt \right]^{1/2}$. La fonction étant de carré intégrable vérifie :

$\left[\int_R [x(t)]^2 dt \right] < \infty$. Dans ces conditions $\int_{-a}^{+a} x(t) e^{-i2\pi\lambda t} dt$ converge en moyenne quadratique, lorsque $a \rightarrow \infty$, vers une limite $x(\lambda)$ appelée transformée de Fourier de $x(t)$ et $x(\lambda)$ appartient aussi à l'espace des classes de fonctions de carré intégrable.

Réciproquement $\int_{-a}^{+a} x(\lambda) e^{+i2\pi\lambda t} d\lambda$ converge en moyenne quadratique, lorsque $a \rightarrow \infty$ vers $x(t)$ appelée transformée inverse de Fourier.

C. Fonction de transfert, fonction de gain et fonction de phase du filtre

La fonction de réponse en fréquence définie dans le cas continu peut être décomposée au moyen de la relation d'Euler :

$$\begin{aligned} A(\omega) &= \int_{-\infty}^{+\infty} \phi(\tau) e^{-i\omega\tau} d\tau = \int_{-\infty}^{+\infty} \phi(\tau) \cos \omega\tau d\tau - i \int_{-\infty}^{+\infty} \phi(\tau) \sin \omega\tau d\tau \\ &= A_1(\omega) - i A_2(\omega) \end{aligned}$$

$A(\omega)$ est une fonction périodique de période 2π (ou $\frac{2\pi}{\Delta t}$ dans le cas discret).

Sa partie réelle $A_1(\omega)$ est une fonction paire et sa partie imaginaire $A_2(\omega)$ est une fonction impaire.

$A(\omega)$ est une fonction complexe, on peut donc l'écrire sous sa forme polaire :

$$A(\omega) = G(\omega)e^{i\theta(\omega)}$$

$G(\omega)$ le module de la forme polaire s'écrit par définition :

$$G(\omega) = \sqrt{A_1^2(\omega) + A_2^2(\omega)} = \sqrt{|A(\omega)|^2} = \sqrt{T(\omega)}.$$

La fonction $G(\omega)$ est appelée fonction de transfert du filtre et la fonction $T(\omega) = |A(\omega)|^2 = G^2(\omega)$ – produit de $A(\omega)$ par son conjugué – porte le nom de fonction de gain du filtre¹.

L'argument $\theta(\omega)$ est construit à partir de :

$$\cos \theta(\omega) = \frac{A_1(\omega)}{G(\omega)}$$

et de :

$$\sin \theta(\omega) = \frac{A_2(\omega)}{G(\omega)}$$

$$\text{Soit } \tan \theta(\omega) = \frac{A_2(\omega)}{A_1(\omega)} \text{ et } \theta(\omega) = \arctan \frac{A_2(\omega)}{A_1(\omega)}$$

$\theta(\omega)$ porte le nom de fonction de phase du filtre. On peut constater facilement que lorsque le filtre est symétrique alors $\theta(\omega) = 0$. Dans ce cas, $A(\omega) = G(\omega)$.

1. Selon l'appellation de Granger (1969).

Compte tenu de la parité de $A_1(\omega)$ et de $A_2(\omega)$, $G(\omega)$ est une fonction paire et $\theta(\omega)$ une fonction impaire. $T(\omega)$ – carré de $G(\omega)$ – est une fonction paire à valeurs réelles.

Ainsi un filtre est défini dans le temps par sa fonction de réponse impulsionnelle et dans le domaine des fréquences, par sa fonction de réponse en fréquence que l'on peut synthétiser au moyen de la fonction de transfert (ou de gain) et de la fonction de phase. La fonction de gain est la représentation traditionnelle dans le domaine des fréquences du filtre. Elle indique l'action du filtre sur les différentes composantes qui structurent la série brute.

Pour illustrer cette action, considérons le filtre \mathfrak{F} et $x(t)$ une fonction continue, par exemple, l'entrée cosinusoidale : $x(t) = A \cos \omega t$.

$$\begin{aligned} y(t) &= \mathfrak{F}[x(t)] = \mathfrak{F} \left[A \frac{e^{i\omega t} + e^{-i\omega t}}{2} \right] \\ &= \frac{A}{2} \{ \mathfrak{F}(e^{i\omega t}) + \mathfrak{F}(e^{-i\omega t}) \} \end{aligned}$$

Comme : $\mathfrak{F}(e^{-i\omega t}) = \int_{-\infty}^{+\infty} \phi(\tau) e^{-i\omega t(t-\tau)} d\tau = \left[\overline{e^{i\omega t} A(\omega)} \right]$
alors :

$$\begin{aligned} y(t) &= \frac{A}{2} \left\{ A(\omega) e^{i\omega t} + \overline{A(\omega) e^{i\omega t}} \right\} \\ &= AG(\omega) \left\{ \frac{e^{i(\omega t + \theta(\omega))} + e^{-i(\omega t + \theta(\omega))}}{2} \right\} \\ y(t) &= AG(\omega) \cos[\omega t + \theta(\omega)] \end{aligned}$$

L'oscillation de départ a son amplitude augmentée ou réduite par $G(\omega)$ et elle est déphasée – par rapport à l'entrée – de $\theta(\omega)$.

Supposons, par exemple, que l'entrée du filtre soit la superposition de deux oscillations, une de basse fréquence ω_1 (cycle long) et l'autre de haute fréquence ω_2 (cycle court).

Si nous désirons réduire l'oscillation de haute fréquence en conservant au maximum l'oscillation de basse fréquence, il faut choisir un filtre de fonction de

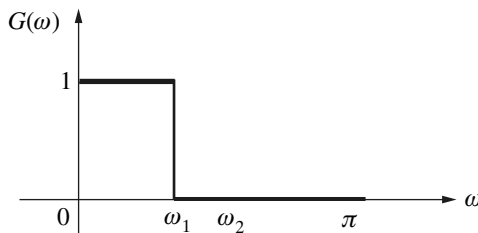


Figure 4.2 – Exemple d'une fonction de transfert

transfert $G(\omega)$ qui a une valeur nulle pour $\omega = \omega_2$ et une valeur proche de 1 pour $\omega = \omega_1$

Cette fonction de transfert (cf. figure 4.2) est une fonction de transfert théorique et il n'est pas toujours possible de construire un filtre qui lui corresponde. En général, ce problème est résolu en choisissant parmi les filtres existants celui qui possède la fonction de transfert la plus proche de celle désirée.

D. Exemples de filtres linéaires

1) Le filtre aux différences d'ordre p non saisonnier

Le filtre aux différences d'ordre p peut se déduire des filtres aux différences d'ordre 1 successifs ; le filtre aux différences d'ordre 1 noté $x_{1,t}$ est donné par :

$$x_{1,t} = x_t - x_{t-1} = x_t - Bx_t = (1 - B)x_t$$

Le filtre aux différences d'ordre 2, noté $x_{2,t}$ s'écrit alors :

$$\begin{aligned} x_{2,t} &= x_{1,t} - x_{1,t-1} = x_t - x_{t-1} - x_{t-1} + x_{t-2} \\ &= x_t - 2x_{t-1} + x_{t-2} = x_t - 2Bx_t + B^2x_t \\ &= (1 - B)^2x_t \end{aligned}$$

Soit :

$$\begin{aligned} x_{2,t} &= (1 - B)x_{1,t} = (1 - B)[(1 - B)x_t] \\ &= (1 - B)^2x_t \end{aligned}$$

En généralisant, l'écriture à l'ordre p est la suivante :

$$\begin{aligned} x_{p,t} &= (1 - B)x_{p-1,t} = (1 - B)(1 - B)^{p-1}x_t \\ &= (1 - B)^p x_t \end{aligned}$$

On développe $(1 - B)^p$ à partir du binôme de Newton :

$$\begin{aligned} x_{p,t} &= (-1)^{2p} C_p^0 x_t + (-1)^{2p-1} C_p^1 x_{t-1} + \dots + (-1)^p C_p^p x_{t-p} \\ x_{p,t} &= \sum_{\tau=0}^p \phi_\tau x_{t-\tau} \end{aligned}$$

avec $\phi_\tau = (-1)^{2p-\tau} C_p^\tau$ et $\sum_{\tau} \phi_\tau = \sum_{\tau} (-1)^{-\tau} C_p^\tau = 0$ (écriture d'un filtre linéaire et invariant).

Ce filtre asymétrique est appelé filtre à un seul côté. À partir de son écriture, la réponse en fréquence peut être déterminée :

$$A_p(\omega) = \sum_{\tau=0}^p \phi_\tau e^{-i\omega\tau}$$

Soit d'après la relation d'Euler :

$$A_p(\omega) = 1 + \sum_{\tau=1}^p (-1)^{2p-\tau} C_p^\tau \cos \omega\tau - i \sum_{\tau=1}^p (-1)^{2p-\tau} C_p^\tau \sin \omega\tau$$

que l'on peut écrire sous sa forme polaire (fonction de transfert) :

$$G_p(\omega) = [2(1 - \cos \omega)]^{p/2} = 2^p \sin^p \frac{\omega}{2}$$

La fonction de gain s'écrit : $T_p(\omega) = G_p^2(\omega) = [2(1 - \cos \omega)]^p$

La fonction de phase de ce filtre est alors : $\theta_p(\omega) = p \left[\frac{\pi}{2} - \frac{\omega}{2} \right]$

Le filtre est ainsi entièrement caractérisé.

Lorsque ω varie entre 0 et π , $G_p(\omega)$ est une fonction croissante de 0 à 2^p . Pour $\omega = 0$ ($p \geq 2$) et $\omega = \pi$, la fonction présente respectivement un minimum et un maximum. Le point de coordonné ($\pi/3$ et 1) indique le renversement d'influence du filtre sur les composantes de la chronique : en-deçà de ce point elles sont réduites et au-delà accrues. Enfin, le déphasage introduit dans les cycles est d'autant plus important aux basses fréquences que l'ordre p de la différence est grand.

De ce filtre général, les caractéristiques du filtre aux différences premières (très utilisé en économétrie) peuvent être déduites. Il s'écrit pour $p = 1$:

$$x_{1,t} = x_t - x_{t-1} = \sum_{\tau=0}^1 \phi_\tau x_{t-\tau}$$

Avec : $\phi_0 = 1, \phi_1 = -1$

Il a pour fonction de réponse en fréquence :

$$A_1(\omega) = \sum_{\tau=0}^1 \phi_\tau e^{-i\omega\tau} = 1 - \cos \omega + i \sin \omega$$

et pour fonction de transfert :

$$G_1(\omega) = \sqrt{2(1 - \cos \omega)} = [2(1 - \cos \omega)]^{1/2} = 2 \sin \frac{\omega}{2}$$

et pour fonction de phase :

$$\theta_1(\omega) = \frac{\pi}{2} - \frac{\omega}{2}$$

Nous pouvons remarquer que $G_2(\omega)$, la fonction de transfert du filtre aux différences secondes est égale à $G_1^2(\omega) = T_1(\omega)$ la fonction de gain du filtre aux différences premières.

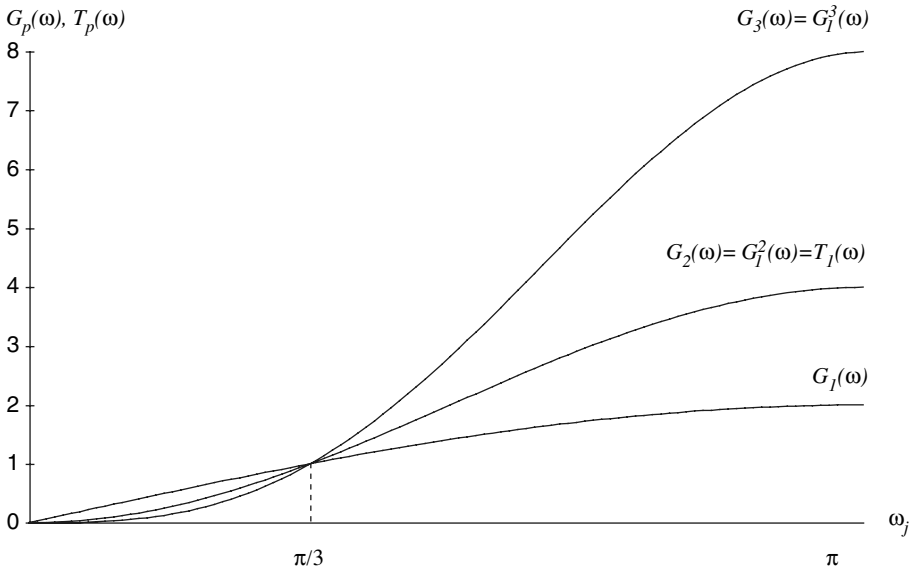


Figure 4.3 – Exemples de fonctions de gain et de transfert

En rendant « discret » l'axe des abscisses ($\omega_j, j = 1, \dots, m$), nous pouvons représenter graphiquement les évolutions des fonctions de transfert et de gain du filtre aux différences.

La figure 4.3 illustre des fonctions de gain et de transfert, on constate que pour une chronique mensuelle, les cyclicités de période 6 mois restent inchangées alors que les fluctuations longues sont amoindries voire éliminées. En revanche, les oscillations de courtes périodes (3 mois par exemple) sont amplifiées.

2) Le filtre aux différences d'ordre p saisonnier

Soit y_t la sortie du filtre aux différences d'ordre p saisonnier de période s et x_t la chronique entrante : $y_t = (1 - B^s)^p x_t$

Ce filtre a pour fonction de gain :

$$T(\omega) = (2 - 2 \cos s\omega)^p = 2^p (1 - \cos s\omega)^p$$

Il s'agit d'une fonction périodique qui possède des maxima :

$$\begin{cases} \omega^{(k)} = \frac{k2\pi}{s} + \frac{\pi}{s} \text{ pour } k = 0, \dots, \frac{s-2}{2} (s \text{ pair}) \text{ ou } \frac{s-1}{2} (s \text{ impair}) \\ T^k(\omega) = 2^{2p} \end{cases}$$

et des minima :

$$\begin{cases} \omega^{(k)} = \frac{k2\pi}{s} \text{ pour } k = 0, \dots, \frac{s}{2} (s \text{ pair}) \text{ ou } \frac{s-1}{2} (s \text{ impair}) \\ T^k(\omega) = 0 \end{cases}$$

Ce filtre élimine la composante de fréquence nulle (tendance) et les composantes cycliques de périodes $\frac{k2\pi}{2}$ (saisonnalité et ses harmoniques) ; en revanche, il introduit ou amplifie les fluctuations de périodes $2s$ (et ses harmoniques).

Nous pouvons déduire de ce filtre général, le filtre aux différences saisonnières d'ordre s utilisé le plus couramment : $y_t = (1 - B^s)x_t$

Considérons le cas d'une chronique mensuelle ($s = 12$), le filtre a pour fonction de gain :

$$T(\omega) = 2(1 - \cos 12\omega)$$

pour fonction de transfert :

$$G(\omega) = \sqrt{T(\omega)} = \sqrt{2^2 \sin^2 \frac{12\omega}{2}} = 2 \sin 6\omega$$

et pour fonction de phase :

$$\theta(\omega) = -\frac{\pi}{2} + 6\omega$$

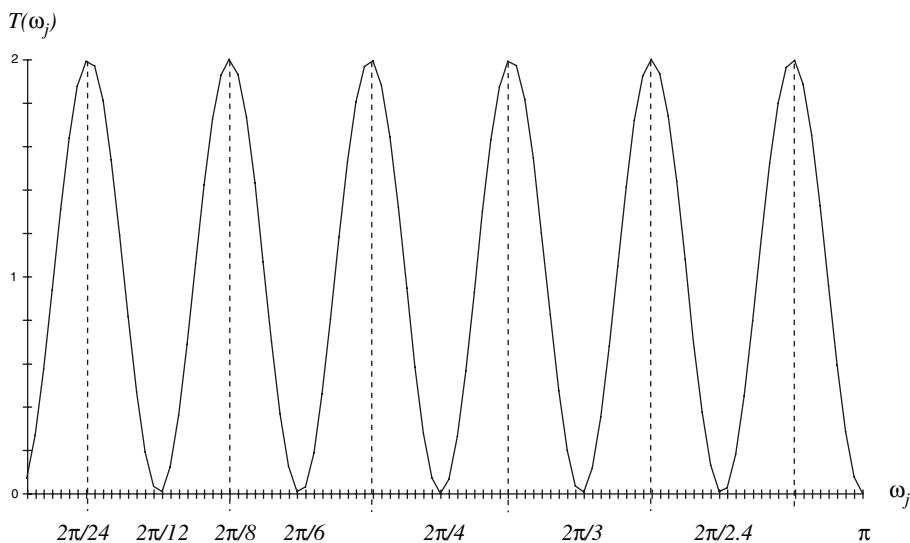


Figure 4.4 – Fonction de gain de filtre $y_t = (1 - B^{12})x_t$ pour une saisonnalité mensuelle

La figure 4.4 illustre la représentation graphique de $T(\omega_j)$ avec $(\omega_j, j = 1, \dots, m)$; le filtre élimine bien une saisonnalité de période $s = 12$ mois, une fluctuation de basse fréquence (tendance) mais introduit dans le « bruit » de la chronique des oscillations de périodes 24 mois, 8 mois, ...

3) Le filtre de la moyenne mobile simple

Soit y_t , la sortie du filtre de la moyenne mobile simple de la chronique entrante x_t , elle s'écrit sous sa forme générale :

$$y_t = \sum_{\tau=-q}^q \phi_{\tau} x_{t-\tau}$$

$$\phi_{\tau}^{(1)} = \frac{1}{2q+1} \quad \forall \tau \in [-q; +q] \quad (\text{moyenne mobile impaire})$$

$$\begin{cases} \phi_{\tau}^{(2)} = \frac{1}{4q} \quad \forall \tau \in \{-q; +q\} \\ \phi_{\tau}^{(2)} = \frac{1}{2q} \quad \forall \tau \in \{-q+1; +q-1\} \end{cases} \quad (\text{moyenne mobile paire})$$

Considérons, par exemple, le cas d'une moyenne mobile impaire. La fonction de réponse en fréquence associée au filtre s'écrit :

$$A_1(\omega) = \sum_{\tau=-q}^q \frac{1}{2q+1} e^{-i\omega\tau}$$

Soit d'après la relation d'Euler :

$$A_1(\omega) = \frac{1}{2q+1} \left[1 + 2 \sum_{\tau=1}^q \cos \omega\tau \right]$$

Après transformation ¹ on obtient :

$$A_1(\omega) = \frac{1}{2q+1} \frac{\sin\left(q + \frac{1}{2}\right)\omega}{\sin\frac{\omega}{2}}$$

La fonction de gain est alors :

1. Terraza, 1979, p. 96.

$$T_1(\omega) = \left[\frac{1}{2q+1} \times \frac{\sin\left(q + \frac{1}{2}\right)\omega}{\sin\frac{\omega}{2}} \right]^2$$

En procédant de manière identique pour la moyenne mobile paire, nous obtenons :

$$T_2(\omega) = \left[\frac{1}{2q} \sin \omega q \cotg \frac{\omega}{2} \right]^2$$

Ces filtres linéaires sont symétriques et ne possèdent pas de ce fait de fonction de phase.

Les fonctions de gain ont des zéros aux fréquences $\omega = \frac{2k\pi}{s}$ avec $k = 1, \dots, q$ et $q = \frac{s}{2}$ dans le cas pair et $q = \frac{s-1}{2}$ dans le cas impair.

Lorsque la longueur de la moyenne mobile est choisie égale à la période de la saisonnalité, celle-ci (et ses harmoniques) est éliminée. Le filtre conserve ou amoindrit les fluctuations de basses fréquences (inférieures à $\frac{2\pi}{s}$) et élimine les cyclicités des hautes fréquences. On appelle « bande passante » l'intervalle de fréquence $\left[0; \frac{2\pi}{s}\right]$ et le filtre porte alors le nom de « filtre passe bande ».

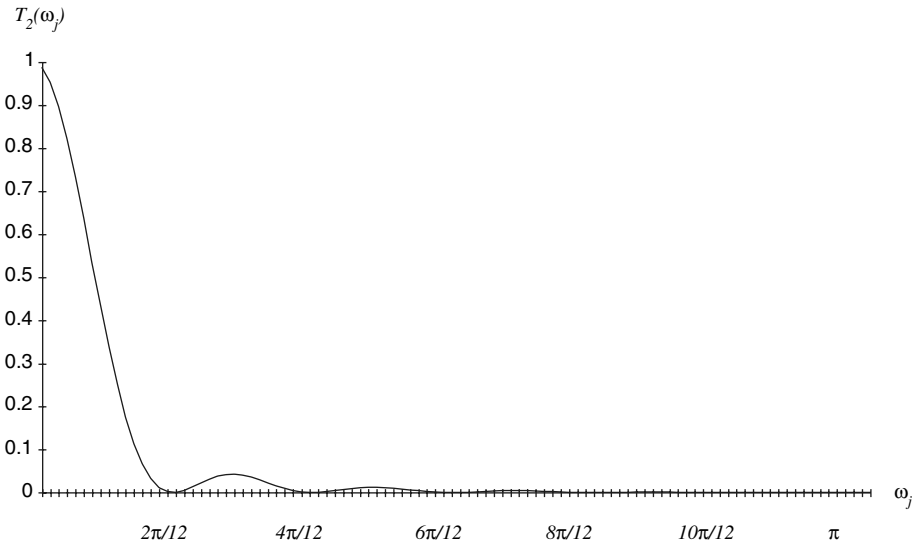


Figure 4.5 – Fonction de gain $T_2(\omega_j)$ pour une série mensuelle ($s = 12$)

Cependant, en ce qui concerne les hautes fréquences, la fonction de gain possède des lobes secondaires qui introduisent dans la chronique des bruits parasites.

Avec $s = 12$ et ω_j pour $j = 1, \dots, m$, la figure 4.5 illustre la représentation graphique de la fonction de gain $T_2(\omega_j)$.

4) Les combinaisons de filtres

Il est possible à partir des filtres décrits précédemment, de construire de nouveaux filtres dont on connaît les caractéristiques. En effet, nous pouvons démontrer que la fonction de gain (ou la fonction de transfert) d'un filtre résultant de la combinaison de plusieurs filtres est égale au produit des fonctions de gain de chacun des filtres.

À titre d'exemple ¹, considérons y_t la série filtrée de x_t après passages successifs dans les filtres aux différences d'ordre d et aux différences saisonnières d'ordre D .

$$\text{Soit : } y_t = (1 - B)^d (1 - B^s)^D x_t$$

La fonction de gain de ce filtre utilisé traditionnellement dans l'algorithme ² de Box et Jenkins est :

$$T(\omega) = 2^d (1 - \cos \omega)^d 2^D (1 - \cos s\omega)^D = 2^{d+D} (1 - \cos \omega)^d (1 - \cos s\omega)^D$$

La figure 4.6 illustre le graphique de la fonction de gain pour $d = D = 1$ et $s = 12$. Ce filtre combine les propriétés des deux filtres :

- élimination de la composante des basses fréquences et saisonnières ;
- amplification ou introduction d'autres fluctuations à des fréquences différentes.

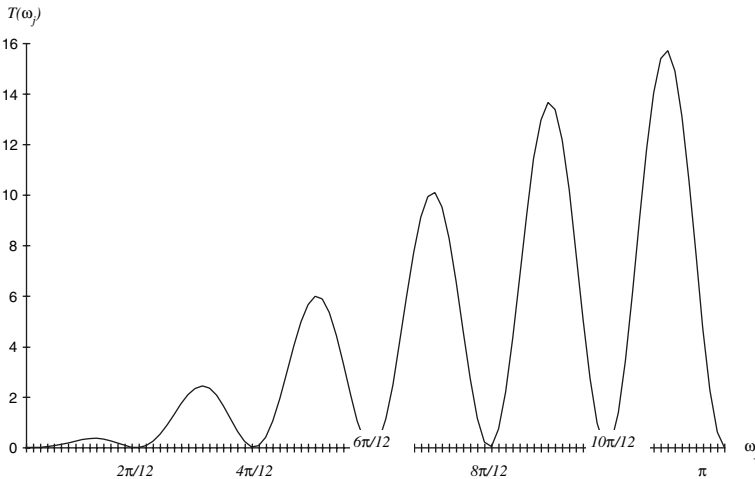


Figure 4.6 – Fonction de gain pour $d = D = 1$ et $s = 12$

1. Terraza, 1981, p. 149.
2. Cf. chapitre 6.

Exercice n° 4.1

Propriétés du filtre du Lissage Exponentiel Simple (LES)

- a) Montrer en utilisant la formule du lissé en t que le LES est un filtre linéaire à un seul côté.
- b) Définir ses fonctions de réponse en fréquence, de transfert, de gain et de phase.
- c) Étudier les propriétés du filtre à partir de l'analyse mathématique de sa fonction de gain.

Solutions

a) Au chapitre 2, nous avons montré que la série temporelle x_t transformée par un lissage exponentiel simple s'écrit :

$$z_{t-1} = \hat{x}_t = \lambda \sum_{i=0}^{\infty} (1-\lambda)^i x_{t-1-i}$$

Il s'agit du lissé en $t-1$ utilisé pour prévoir la chronique en t . Le lissé en t s'écrit donc :

$$z_t = \lambda \sum_{i=0}^{\infty} (1-\lambda)^i x_{t-i}$$

Soit encore en utilisant l'écriture des filtres :

$$z_t = \sum_{\tau=0}^{\infty} \lambda(1-\lambda)^\tau x_{t-\tau} \text{ et dans ce cas } \phi_\tau = \lambda(1-\lambda)^\tau.$$

Le LES est un filtre \mathfrak{N} tel que $\mathfrak{N}[x_t] = z_t$. En effet, si $(\alpha, \beta) \in R^2$, on a :

$$\begin{aligned} \mathfrak{N}[\alpha x_t^1 + \beta x_t^2] &= \sum_{\tau=0}^{\infty} \lambda(1-\lambda)^\tau [\alpha x_{t-\tau}^1 + \beta x_{t-\tau}^2] \\ &= \alpha \left[\sum_{\tau=0}^{\infty} \lambda(1-\lambda)^\tau x_{t-\tau}^1 \right] + \beta \left[\sum_{\tau=0}^{\infty} \lambda(1-\lambda)^\tau x_{t-\tau}^2 \right] \\ &= \alpha z_t^1 + \beta z_t^2 = \alpha \mathfrak{N}[x_t^1] + \beta \mathfrak{N}[x_t^2] \end{aligned}$$

et
$$\mathfrak{N}[x_{t+t_0}] = \sum_{\tau=0}^{\infty} \lambda(1-\lambda)^\tau x_{t+t_0-\tau} = z_{t+t_0}$$

Le LES est un filtre linéaire asymétrique ou filtre à un seul coté.

b) La fonction de réponse en fréquence s'écrit :

$$\begin{aligned} A(\omega) &= \sum_{\tau=0}^{\infty} \lambda(1-\lambda)^\tau e^{-i\omega\tau} = \lambda + \lambda(1-\lambda)e^{-i\omega} + \lambda(1-\lambda)^2 e^{-i2\omega} + \dots \\ &= \lambda[1 + (1-\lambda)e^{-i\omega} + (1-\lambda)^2 e^{-i2\omega} + \dots] = \lambda \frac{1}{1 - (1-\lambda)e^{-i\omega}} \\ &= \lambda \frac{1}{(1 - (1-\lambda)\cos \omega) - i(1-\lambda)\sin \omega} = \frac{\lambda[(1 - (1-\lambda)\cos \omega) + i(1-\lambda)\sin \omega]}{(1 - (1-\lambda)\cos \omega)^2 + i(1-\lambda)^2 \sin^2 \omega} \\ A(\omega) &= \frac{\lambda - \lambda(1-\lambda)\cos \omega}{1 - 2(1-\lambda)\cos \omega + (1-\lambda)^2} + i \frac{\lambda(1-\lambda)\sin \omega}{1 - 2(1-\lambda)\cos \omega + (1-\lambda)^2} \end{aligned}$$

La fonction de transfert du filtre est alors :

$$G(\omega) = \sqrt{|A(\omega)|^2} = \sqrt{T(\omega)}$$

La fonction de gain $T(\omega)$ est donc égale à :

$$T(\omega) = \frac{\lambda^2 \{[(1 - (1 - \lambda)\cos \omega)^2 + (1 - \lambda)^2 \sin^2 \omega]\}}{[(1 - 2(1 - \lambda)\cos \omega) + (1 - \lambda)^2]^2}$$

$$T(\omega) = \frac{\lambda^2 \{1 + (1 - \lambda)^2 - 2(1 - \lambda) \cos \omega\}}{[(1 - 2(1 - \lambda)\cos \omega) + (1 - \lambda)^2]^2}$$

$$T(\omega) = \frac{\lambda^2}{[(1 - 2(1 - \lambda)\cos \omega) + (1 - \lambda)^2]^2}$$

La fonction de phase s'écrit :

$$\theta(\omega) = \text{Arctg} \frac{\lambda(1 - \lambda)\sin \omega}{\lambda[1 - (1 - \lambda)\cos \omega]}$$

$$\theta(\omega) = \text{Arctg} \frac{(1 - \lambda)\sin \omega}{[1 - (1 - \lambda)\cos \omega]}$$

e) Étude de la fonction $T(\omega)$ avec $\omega \in [0 ; \pi]$ et $\lambda \in]0 ; 1[$:

$$\omega = 0 \implies T(0) = \frac{\lambda^2}{[(1 - 2(1 - \lambda)) + (1 - \lambda)^2]} = \frac{\lambda^2}{\lambda^2} = 1$$

$$\omega = \pi \implies T(\pi) = \frac{\lambda^2}{[(1 - 2(1 - \lambda)) + (1 - \lambda)^2]} = \frac{\lambda^2}{(2 - \lambda)^2}$$

La dérivée $T'(\omega)$ est égale à : $T'(\omega) = \frac{-2\lambda^2(1 - \lambda)\sin \omega}{[(1 - 2(1 - \lambda)\cos \omega) + (1 - \lambda)^2]^2}$

$T'(\omega)$ est du signe de $-\sin \omega$ donc négative sur $[0 ; \pi]$. $T(\omega)$ est une fonction décroissante du maximum (0 ; 1) au minimum $\left(\pi ; \frac{\lambda^2}{(2 - \lambda)^2}\right)$. La nature de cette décroissance (cf. tableau 4.1 et figure 4.7) dépend de la valeur de λ (programme Eviews C4EX1.PRG).

Tableau 4.1 – Exemples de décroissance $T(\omega)$ en fonction des valeurs de λ

	$T(\pi/3)$	$T(\pi/2)$	$T(2\pi/3)$	$T(\pi)$
$\lambda = 0,8$	0,77	0,62	0,52	0,44
$\lambda = 0,5$	0,33	0,20	0,14	0,11
$\lambda = 0,2$	0,05	0,02	0,016	0,01

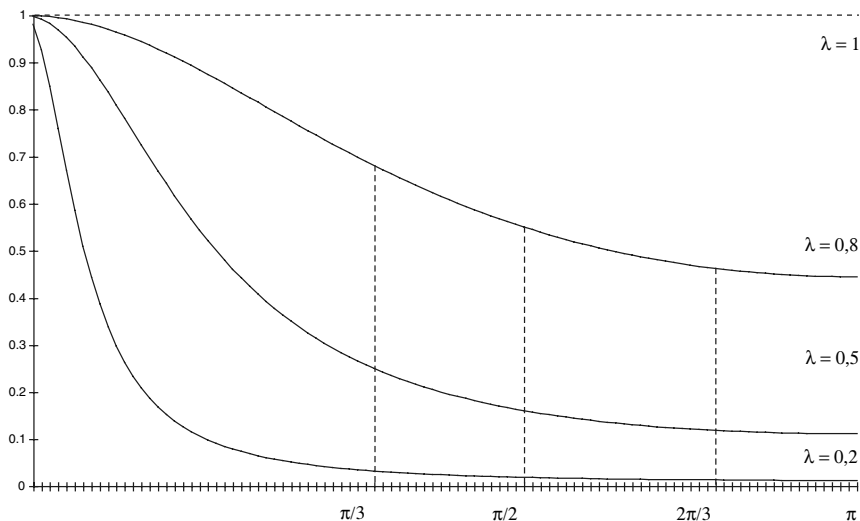


Figure 4.7 – Exemples de fonctions de gain $T(\omega)$ pour différentes valeurs de λ

Si $\lambda = 0 \implies T(\omega) = 0$, la série lissée est nulle.

Si $\lambda = 1 \implies T(\omega) = 1$, la série lissée z_t est égale à la chronique x_t .

Si $\lambda \in]0 ; 1[\implies T(\omega)$ est une fonction décroissante de $\omega = 0$ à $\omega = \pi$ et le lissage est d'autant plus important que λ est proche de 0.

En termes de fréquences, nous constatons que plus λ tend vers zéro et plus le filtre élimine les composantes de hautes fréquences tout en conservant celles de basses fréquences. Le LES peut donc être utilisé pour conserver dans une chronique les cycles longs.

II. Le spectre d'un processus aléatoire

A. Les théorèmes de représentation

L'analyse harmonique d'une fonction aléatoire $Z(t)$ réelle ou complexe est sa décomposition en une somme de fonction de type $a_i z_i(t)$ où les $z_i(t)$ sont des fonctions aléatoires du temps et les coefficients a_i des variables aléatoires réelles ou complexes mais indépendantes du temps. Dans l'analyse harmonique, les fonctions $z_i(t)$ sont choisies complexes : $z_i(t) = e^{i\omega t}$.

Ce choix est guidé par les propriétés d'orthogonalité et d'invariance de ce type de fonction au facteur $A(\omega)$ près, suite à un passage dans un filtre de $z_i(t)$

(cf. précédemment). L'analyse harmonique est donc une généralisation au cas aléatoire de l'analyse de Fourier, elle s'appuie sur deux théorèmes :

- le théorème de Loeve : ce théorème présente la décomposition harmonique de $Z(t)$, selon : $Z(t) = \int_{-\infty}^{+\infty} e^{i2\pi\lambda t} dU_Z(\lambda)$ dans laquelle $dU_Z(\lambda)$ sont des variables aléatoires en général complexes ;
- le théorème de Khintchine : il s'agit d'un théorème équivalent à celui de Loeve mais appliqué à la fonction d'autocovariance $\gamma_Z(\theta)$ de $Z(t)$:

$$\gamma_Z(\theta) = \int_{-\infty}^{+\infty} e^{i2\pi\lambda\theta} E \{ |dU_Z(\lambda)|^2 \}$$

Si la fonction aléatoire $Z(t)$ est stationnaire et continue à l'origine alors : $E \{ |dU_Z(\lambda)|^2 \} = dF_Z(\lambda)$, où les $dF_Z(\lambda)$ sont des accroissements jamais négatifs de somme finie.

La fonction $F_Z(\lambda)$ porte le nom de fonction de répartition spectrale (ou spectre) de la loi de probabilité associée à la fonction aléatoire $Z(t)$.

$F_Z(\lambda)$ est une fonction bornée, monotone non décroissante, continue à gauche, et telle que $\lim_{\lambda \rightarrow \infty} F_Z(\lambda) = 0$.

Grâce à ces propriétés, nous pouvons définir par la fonction F_Z une mesure de probabilité de Lebesgues et construire l'intégrale de même nom qui se décompose en trois intégrales selon la forme unique suivante :

$$F_Z(\lambda) = F_Z^1(\lambda) + F_Z^2(\lambda) + F_Z^3(\lambda)$$

où $F_Z^1(\lambda)$ est une fonction absolument continue.

$F_Z^2(\lambda)$ est une fonction absolument discontinue,

$F_Z^3(\lambda)$ est une fonction singulière.

On retient $F_Z^1(\lambda)$ qui est une fonction continue et dérivable, ce qui permet d'écrire :

$$dF_Z^1(\lambda) = f_Z^1(\lambda)d\lambda$$

Si on utilise cette fonction $f_Z^1(\lambda)$, notée par la suite $f_Z(\lambda)$ fonction continue et absolument intégrale, dans le théorème de Khintchine (appelé aussi théorème de Bochner – Khintchine ou encore de Wiener), on a :

$$\gamma_Z(\theta) = \int_{-\infty}^{+\infty} e^{i2\pi\lambda\theta} f_Z(\lambda)d\lambda = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\omega\theta} f_Z(\omega)d\omega$$

On voit alors qu'il est possible d'appliquer à cette écriture le théorème de Fourier et d'en déduire l'expression de $f_Z(\lambda)$:

$$f_Z(\lambda) = \int_{-\infty}^{+\infty} e^{-i2\pi\lambda\theta} \gamma_Z(\theta) d\theta$$

ou encore :

$$f_Z(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i2\omega\theta} \gamma_Z(\theta) d\theta$$

$f_Z(\lambda)$ porte le nom de fonction de densité spectrale ou encore de spectre de puissance ou enfin de spectre de la fonction aléatoire $Z(t)$.

Ainsi, la fonction d'autocovariance (décomposition temporelle de la fonction aléatoire $Z(t)$) et la fonction de densité spectrale (décomposition fréquentielle de la fonction aléatoire $Z(t)$) forment une paire de transformées de Fourier.

B. Le spectre d'une série temporelle filtrée

Soit une fonction aléatoire $Z(t)$ et $y(t)$ sa filtrée par un filtre \mathfrak{S} . À partir du théorème de Loeve, nous pouvons écrire que :

$$\begin{aligned} \mathfrak{S}[Z(t)] = y(t) &= \int_{-\infty}^{+\infty} A(\lambda) e^{i2\pi\lambda t} dU_Z(\lambda) \\ &= \int_{-\infty}^{+\infty} e^{i2\pi\lambda t} \{A(\lambda) dU_Z(\lambda)\} \end{aligned}$$

Nous pouvons démontrer que, si l'entrant dans un filtre linéaire du type \mathfrak{S} est une fonction aléatoire stationnaire alors, la fonction aléatoire sortante l'est aussi, nous pouvons donc écrire selon le théorème de Loeve :

$$y(t) = \int_{-\infty}^{+\infty} e^{i2\pi\lambda t} dU_Z(\lambda)$$

d'où :

$$dU_y(\lambda) = A(\lambda) dU_Z(\lambda)$$

où $A(\lambda)$ est la fonction de réponse en fréquence du filtre qui s'écrit aussi $A(\lambda) = G(\lambda) e^{i\theta(\lambda)}$

On a alors :

$$E \left\{ |dU_y(\lambda)|^2 \right\} = E \left\{ |A(\lambda)|^2 |dU_Z(\lambda)|^2 \right\} = |A(\lambda)|^2 E \left\{ |dU_Z(\lambda)|^2 \right\}$$

Or d'après le théorème de Khintchine, nous savons que :

$$E \left\{ |dU_y(\lambda)|^2 \right\} = dF_y(\lambda) = f_y(\lambda)d\lambda \quad \text{et} \quad E \left\{ |dU_z(\lambda)|^2 \right\} = dF_z(\lambda) = f_z(\lambda)d\lambda .$$

D'où :

$$f_y(\lambda)d\lambda = |A(\lambda)|^2 f_z(\lambda)d\lambda = G^2(\lambda) f_z(\lambda)d\lambda = T(\lambda) f_z(\lambda)d\lambda$$

Ce résultat important peut être résumé par le schéma suivant :

$$\begin{array}{ccc} Z(t) & \xrightarrow{\text{3}} & y(t) \quad \text{stationnaire} \\ \downarrow & & \downarrow \\ f_z(\lambda) & \longleftarrow & f_y(\lambda)d\lambda \quad = \quad T(\lambda) f_z(\lambda)d\lambda \end{array}$$

Le spectre de puissance d'une fonction aléatoire stationnaire d'ordre 2 filtrée est égal au spectre de puissance de la fonction aléatoire entrante multiplié par la fonction de gain du filtre $T(\lambda)$. Comme nous l'avons montré dans le paragraphe précédent, il est possible de connaître à partir des formes analytiques de ces filtres leurs caractéristiques de sorte que dans le domaine des fréquences l'opération de filtrage est parfaitement maîtrisée ; mais ces caractéristiques montrent les multiples précautions qu'il convient de prendre lorsqu'un filtre est appliqué à une série brute.

C. Le spectre d'une chronique ou l'estimateur spectral

Soit une chronique x_t réalisation d'un processus aléatoire stationnaire d'ordre 2.

L'expression « discrétisée » du spectre s'écrit : $f_x(\omega) = \frac{1}{2\pi} \sum_{\theta=-\infty}^{+\infty} e^{-i\omega\theta} \gamma_x(\theta)$.

Considérons un processus x_t réel. Sa fonction d'autocovariance est paire. Pour un nombre de décalages θ fini et égal à $2m + 1$ avec $m \in N^*$, on a ¹ :

$$f_x(\omega) = \frac{1}{2\pi} \sum_{\theta=-m}^{+m} \gamma_x(\theta) \cos \omega\theta$$

Pour obtenir l'estimateur de $f_x(\omega)$ il faut remplacer l'expression théorique $\gamma_x(\theta)$ par son estimateur noté $R_x(\theta)$ de variance minimale ² qui est égal à :

1. Les formules retenues pour l'estimateur spectral sont celles de Granger (1969).
2. Cf. Jenkins et Watts, 1968.

$R_x(\theta) = \frac{1}{n - \theta} \sum_{t=1}^{n-\theta} (x_t - \bar{x})(x_{t+\theta} - \bar{x})$. L'estimateur spectral s'écrit alors :

$$\hat{f}_x(\omega) = \frac{1}{2\pi} \left[R_x(0) + 2 \sum_{\theta=1}^m \cos \omega\theta R_x(\theta) \right]$$

Nous pouvons démontrer que l'estimateur obtenu est centré mais non convergent. Il ne convient que pour les chroniques qui possèdent des périodicités strictes et sa représentation graphique (périodogramme, équivalent de la fonction d'autocorrélation dans le domaine du temps) possède une forme heurtée qui la rend difficilement interprétable.

Pour contourner cette difficulté de la lecture du graphique, on estime pour chaque fréquence non pas la valeur $\hat{f}_x(\omega)$ mais une valeur moyenne sur des bandes de fréquences égales dont la juxtaposition recouvre l'intervalle donné par ω .

En pratique, on procède à un lissage du spectre. La technique utilisée est assimilable au filtrage précédent de la fonction aléatoire. Nous prenons autour d'un nombre de points équidistants de l'axe des fréquences, une moyenne pondérée sur les fréquences voisines : on « perce » une fenêtre dans le diagramme des fréquences. Le choix d'une de ces fenêtres est délicat car selon sa forme, c'est-à-dire selon les facteurs de pondération choisis, des fuites spectrales aux fréquences adjacentes pour une fréquence donnée sont possibles. Si à une fréquence ω_k du spectre d'une chronique existe un pic de puissance spectrale importante, par rapport aux fréquences adjacentes, les lobes secondaires inhérents à la fenêtre spectrale peuvent générer, à leurs fréquences, des puissances spectrales non négligeables. Il s'avère donc difficile de choisir une bonne fenêtre spectrale, car il faut que la puissance des lobes secondaires soit faible et que les puissances aux fréquences adjacentes ne soient pas corrélées.

Deux fenêtres sont les plus couramment employées en analyse spectrale :

- la fenêtre de Tukey-Hanning,
- la fenêtre de Parzen.

Pour la première de ces fenêtres, nous pouvons démontrer que les pics secondaires n'excèdent pas 2 % du pic principal et que les estimateurs spectraux aux fréquences ω_k et ω_{k+2} ne sont pas corrélés. Elle présente néanmoins l'inconvénient d'engendrer, dans certains cas, des fuites à des puissances spectrales adjacentes. Celles-ci peuvent être alors négatives ce qui est en contradiction avec le fait qu'un spectre, par construction, est toujours positif. Dans ce cas, il est préférable de recourir à l'estimateur de Parzen.

L'estimateur spectral s'écrit alors :

$$\hat{f}_x(\omega) = \frac{1}{2\pi} \left[\lambda_0 R_x(0) + 2 \sum_{\theta=1}^m \lambda_\theta R_x(\theta) \cos \omega\theta \right]$$

Les λ_θ sont les coefficients de pondération issus, ici, de la fenêtre spectrale de Tukey-Hanning :

$$\lambda_\theta = \frac{1}{2} \left[1 + \cos \frac{\pi\theta}{m} \right]$$

$$\lambda_0 = 1$$

$$\lambda_m = 0$$

Dans l'écriture de la fenêtre et du spectre, m est le nombre de points choisis sur l'axe des fréquences du spectre. C'est le degré de précision du spectre. Ainsi m est choisi égal au retard maximum considéré pour le calcul de la fonction d'autocovariance.

Pour $j = 0, \dots, m$ points pris sur l'axe des fréquences du spectre, l'estimateur spectral en un point ω_j s'écrit :

$$\hat{f}_x(\omega_j) = \frac{1}{2\pi} \left[R_x(0) + \sum_{\theta=1}^{m-1} R_x(\theta) \left[1 + \cos \frac{\pi\theta}{m} \right] \cos \omega_j\theta \right]$$

avec $\omega_j = \frac{\pi j}{m}$ et m tel que : $\frac{n}{6} \leq m \leq \frac{n}{3}$.

Il est possible de calculer le biais de l'estimateur, c'est-à-dire la différence entre l'estimateur et sa valeur théorique¹ :

$$B[\omega_j] = E \left[\hat{f}_x(\omega_j) - f_x(\omega_j) \right] = \frac{0,063}{m^2} f_x''(\omega_j) + 0 \times \frac{1}{m^4} = \frac{0,063}{m^2} f_x''(\omega_j)$$

avec $f_x''(\omega_j)$ la dérivée seconde de $f_x(\omega_j)$

La variance de cet estimateur est donnée par :

$$\text{Var} \left[\hat{f}_x(\omega_j) \right] = 0,75 \frac{m}{n} f_x^2(\omega_j) \text{ pour } \omega_j \neq 0.$$

Pour obtenir un spectre estimé stable, il faut minimiser sa variance, or celle-ci s'accroît avec le nombre de bandes m qui découpe l'axe des fréquences et elle diminue avec la longueur n de la série : le rapport $\frac{m}{n}$ joue donc un rôle important. En pratique, n est une donnée, la stabilité du spectre est d'autant meilleure que $\frac{m}{n}$ est petit, ce qui se ramène à m petit. Pour avoir une haute fidélité du spectre, il faut minimiser le biais et pour cela avoir m suffisamment grand. D'ailleurs, nous pouvons constater que la fenêtre de Tukey-Hanning est souvent préférée aux autres à cause d'une valeur faible du biais, pour m donné.

1. Cf. Jenkins et Watts, 1968, p. 247-252.

En définitive, pour une longueur de la chronique n donnée, il y a contradiction entre une recherche de la stabilité qui entraîne m à la baisse et une meilleure fidélité qui fait augmenter m . C'est la raison pour laquelle, grâce à l'expérience, cette valeur de m est choisie entre $\frac{n}{6}$ et $\frac{n}{3}$.

Nous pouvons enfin transformer l'expression précédente de l'estimateur spectral :

$$\begin{aligned}
 \hat{f}_x(\omega_j) &= \frac{1}{2\pi} \left[R_x(0) + \sum_{\theta=1}^{m-1} R_x(\theta) \left[1 + \cos \frac{\pi\theta}{m} \right] \cos \omega_j\theta \right] \\
 &= \frac{1}{2\pi} \left[R_x(0) + \sum_{\theta=1}^{m-1} R_x(\theta) \left[\cos \omega_j\theta + \frac{1}{2} \left[\cos \frac{\pi\theta}{m} (j+1) \right. \right. \right. \\
 &\qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \left. \left. \left. + \cos \frac{\pi\theta}{m} (j-1) \right] \right] \right] \\
 &= \frac{1}{2\pi} \left[R_x(0) + \sum_{\theta=1}^{m-1} R_x(\theta) \cos \omega_j\theta + \frac{1}{2} \sum_{\theta=1}^{m-1} R_x(\theta) \cos \omega_{j+1} \right. \\
 &\qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \left. + \frac{1}{2} \sum_{\theta=1}^{m-1} R_x(\theta) \cos \omega_{j-1} \right] \\
 &= \frac{1}{4\pi} \left[R_x(0) + 2 \sum_{\theta=1}^{m-1} R_x(\theta) \cos \omega_j\theta \right] \\
 &\qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad + \frac{1}{8\pi} \left[R_x(0) + 2 \sum_{\theta=1}^{m-1} R_x(\theta) \cos \omega_{j+1}\theta \right] \\
 &\qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad + \frac{1}{8\pi} \left[R_x(0) + 2 \sum_{\theta=1}^{m-1} R_x(\theta) \cos \omega_{j-1}\theta \right]
 \end{aligned}$$

En posant :

$$L_j = \frac{1}{2\pi} \left[R_x(0) + 2 \sum_{\theta=1}^{m-1} R_x(\theta) \cos \omega_j\theta \right], \text{ l'estimateur spectral s'écrit :}$$

$$\hat{f}_x(\omega_j) = 0,25 L_{j-1} + 0,5 L_j + 0,25 L_{j+1}$$

C'est cette formule, plus simple, qui est utilisée dans les programmes informatiques.

D. La lecture d'un spectre

- Le spectre¹ est représenté graphiquement de la manière suivante : En ordonnée, on porte le plus souvent le logarithme népérien de la valeur spectrale en lieu et place des valeurs brutes. En effet, en économie, les séries brutes sont souvent structurées selon des fluctuations de période longue et de variance importante. Ces fluctuations sont concentrées aux basses fréquences avec des valeurs spectrales importantes. Sur la représentation graphique, elles masquent les valeurs spectrales des fluctuations plus courtes. Les abscisses sont repérées par la période $T \in]-\infty ; 2]$, la fréquence angulaire $\omega \in [0 ; \pi]$ car la fréquence $\lambda \in [0 ; \frac{1}{2}]$.
- Puisqu'un bruit blanc est formé de variables aléatoires indépendantes et distribuées identiquement, sa fonction d'autocovariance est égale à la variance du processus et elle est nulle ailleurs.

$$R_x(\theta) = \sigma_x^2 \text{ pour } \theta = 0$$

$$R_x(\theta) = 0 \text{ ailleurs}$$

En utilisant la formule précédente du spectre, nous pouvons constater que le spectre d'un bruit blanc est :

$$\hat{f}_x(\omega_j) = \frac{1}{2\pi} \sigma_x^2 \quad \forall \omega_j, \omega_j \in [0 ; \pi]$$

Le spectre d'un bruit blanc est une droite parallèle à l'axe des abscisses (cf. figure 4.8).

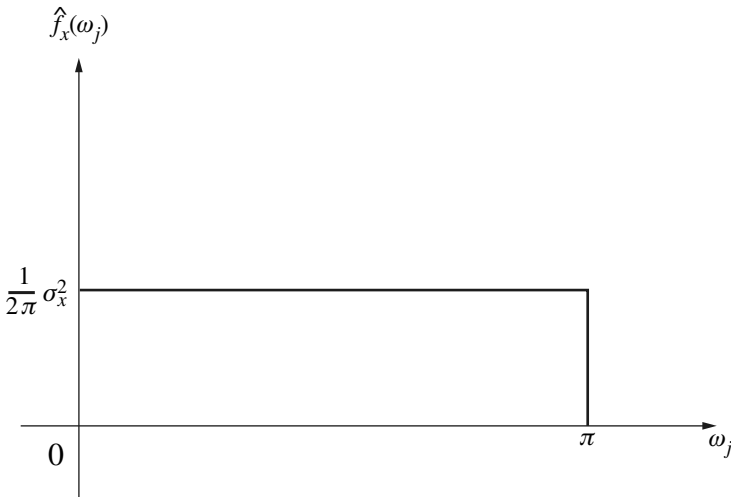


Figure 4.8 – Spectre d'un bruit blanc

1. Cf. Terraza, 1981.

Il semble donc facile *a priori* d'identifier un bruit blanc à partir de son spectre. Cependant, dans la pratique, on a plutôt à faire à des pseudos bruits blancs, c'est-à-dire à des processus pour lesquels la fonction d'autocovariance n'est pas exactement égale à 0 lorsque θ est non nul.

Dans ces cas, les valeurs spectrales sont relativement importantes les unes par rapport aux autres, elles s'enroulent, cependant, autour d'une tendance plate.

- Nous pouvons démontrer que la quantité $\hat{f}_x(\omega_j)$ est une somme pondérée d'estimateurs spectraux empiriques, qui obéissent chacun à un χ^2 à 2 degrés de liberté. En conséquence, la variable aléatoire $\frac{v \hat{f}_x(\omega_j)}{f_x(\omega)}$ obéit à un χ^2 à $\nu = \frac{n}{m}$ degrés de liberté.

D'où :

$$\text{Prob} \left\{ \chi^2(\alpha/2) \leq \frac{v \hat{f}_x(\omega_j)}{f_x(\omega)} \leq \chi^2(1 - \alpha/2) \right\} = 1 - \alpha$$

L'intervalle de confiance à $(1 - \alpha)$ % est égal à :

$$\frac{v \hat{f}_x(\omega_j)}{\chi^2(1 - \alpha/2)} \leq f_x(\omega) \leq \frac{v \hat{f}_x(\omega_j)}{\chi^2(\alpha/2)} \quad \forall j, j = 0, \dots, m$$

ou encore d'après les remarques précédentes :

$$\begin{aligned} \text{Log } \hat{f}_x(\omega_j) + \text{Log} \left(\frac{\nu}{\chi^2_\nu(1 - \alpha/2)} \right) &\leq \text{Log}(f_x(\omega)) \\ &\leq \text{Log } \hat{f}_x(\omega_j) + \text{Log} \left(\frac{\nu}{\chi^2_\nu(\alpha/2)} \right) \end{aligned}$$

Nous constatons que la stabilité de $\hat{f}_x(\omega_j)$ est d'autant meilleure que ν est grand.

Il est donc possible de construire autour du spectre des intervalles de confiance, pour vérifier si un pic spectral est significatif, c'est-à-dire s'il est révélateur de l'existence d'une cyclicité dans la structure de la chronique. Nous pouvons alors procéder graphiquement de la façon suivante : lorsqu'il est possible dans la bande de confiance de tracer une droite parallèle à l'axe des abscisses sans que celle-ci ne rencontre sa borne inférieure ou sa borne supérieure, alors le pic est jugé non significatif et significatif dans le cas contraire.

- L'interprétation des harmoniques permet aussi de conclure sur la signification d'un pic. Lorsque la série brute est un cycle parfait du type $A \cos \omega_0 t$ représenté sur une période T par un nombre de points important, (avec un pas de discrétisation faible) alors le théorème de Fourier indique que la décomposition de cette

série est elle-même. En conséquence, la fonction spectrale théorique de cette sinusoïde parfaite, illustrée par un nombre de points très important, est un point ayant pour abscisse la fréquence de la cosinusoïde et pour ordonnée une valeur proportionnelle à l'amplitude de la cosinusoïde (spectre de raie).

Dans la pratique, on utilise l'estimateur spectral lissé ; les fuites à travers la fenêtre de Tukey-Hanning peuvent engendrer sur le spectre aux fréquences adjacentes de ω_0 , des valeurs spectrales significatives. D'autre part, les cyclicités périodiques de ce type ne se rencontrent pratiquement jamais dans les séries brutes. On dispose en effet, de chroniques dont les cyclicités se répètent plusieurs fois dans l'historique. En supposant que la période soit constante, le cycle peut prendre des formes différentes :

- en dents de scie (c'est celle qui représente le principe de conservation des aires) ;
- en forme de créneaux ;
- en forme de pics ;
- en forme de creux.

Dans tous ces cas, pour approcher la forme de ces cycles, la fonction cosinus précédente s'avère insuffisante. Il est donc nécessaire (conformément au théorème de Fourier) de recourir à des cosinusoïdes de périodes de plus en plus petites : les harmoniques.

Par exemple, pour approcher une cyclicité en dents de scie de période T , il faut employer la cosinusoïde de période T , mais aussi des cosinusoïdes de périodes $T/2$, $T/3$, etc. (les harmoniques) pour approximer la cyclicité de départ.

En économie, les cyclicités de type saisonnalité, qui correspondent au schéma précédent, ne sont pas parfaites car l'amplitude varie et le nombre de points représentant un cycle sur la période est souvent faible. Il en résulte que sur les spectres, les cyclicités de type saisonnières apparaissent en position de fréquence fondamentale et de fréquence harmonique. Les valeurs spectrales aux harmoniques sont d'autant plus accusées que la cyclicité de la chronique est difficile à reconstituer (soit car très fluctuante, soit car peu de points). Ce type de fluctuation cyclique se manifeste donc sur le spectre par des raies d'abscisses harmoniques et dont les ordonnées ont des valeurs spectrales continûment décroissantes.

Dans la pratique, comme on utilise une fenêtre pour l'estimateur spectral, le spectre d'une saisonnalité mensuelle par exemple, est constitué par des pics de valeurs décroissantes aux fréquences :

$$\omega_0 = \frac{2\pi}{12} ; \omega_1 = \frac{2\pi}{6} ; \omega_2 = \frac{2\pi}{4} ; \omega_3 = \frac{2\pi}{3} ; \omega_4 = \frac{2\pi}{2,4}.$$

Nous pouvons nous aider de cette représentation pour déceler dans une série brute l'existence de deux saisonnalités en dents de scie. Supposons, en effet, que la série brute comporte une saisonnalité de 12 mois et de 3 mois. Le spectre met

en évidence cette double saisonnalité par un pic plus important en position de fréquence $\frac{2\pi}{3}$.

- L'effet sur le spectre d'une composante accidentelle

Nous pouvons constater qu'une perturbation ponctuelle unique en un point de la chronique introduit une translation des valeurs spectrales vers les ordonnées les plus fortes. L'introduction de plusieurs perturbations peut créer sur le spectre des cyclicités de fréquences quelconques. Cet exemple montre une nouvelle fois qu'il faut interpréter avec beaucoup d'attention des fluctuations lues sur le spectre et qui n'ont aucun rapport *a priori* avec des fluctuations de type économique (Kondratiev, saisonnalités...).

- La fermeture de la fenêtre spectrale

Cette fermeture est liée au nombre de points retenus pour estimer le spectre de la chronique. Plus le nombre de points est important, plus la variance de l'estimateur de la fenêtre est petit, c'est-à-dire plus la fenêtre est fine. De sorte que, lorsqu'apparaît sur le spectre un pic, celui-ci est robuste, s'il est conservé lors de l'ouverture progressive de la fenêtre spectrale. Dans le cas contraire nous pouvons conclure à une cyclicité de faible importance dans la chronique. Concrètement, une première estimation du spectre est réalisée avec m retards puis avec $\frac{m}{2}$ puis avec $\frac{m}{3}$ à condition que m soit suffisant grand.

- Le choix de m

Le nombre de points spectraux m retenu est choisi entre $\frac{n}{6}$ et $\frac{n}{3}$ pour les raisons déjà évoquées. Il peut arriver – si m n'est pas choisi judicieusement – que la lecture du spectre indique une cyclicité de période *a priori* incompatible avec la chronique de départ. Pour une chronique mensuelle saisonnière, le pic de la saisonnalité correspond à une période de 11,5 mois par exemple (impossible car la période est égale à 12 mois) ; cela tient au fait que m est le nombre de points retenu pour discrétiser le spectre. Comme $\omega_j = \frac{2\pi}{T_j}$ si j est un point de l'axe des

abscisses, on a aussi $\omega_j = \frac{\pi_j}{m}$ avec $j = 0 \dots m$. D'où : $\frac{2\pi}{T_j} = \frac{\pi_j}{m} \implies T_j = \frac{2m}{j}$.

En faisant varier j et m , nous constatons qu'il est possible d'avoir une période qui ne corresponde pas à la période saisonnière.

Exercice n° 4.2

Relation entre la variable m et les valeurs j

Soit une chronique mensuelle de $n = 300$ points, $\frac{n}{6} \leq m \leq \frac{n}{3}$ d'où $50 \leq m \leq 100$.

Supposons que l'on choisisse $m = 50$. Une saisonnalité T_j est au point :

$$T_j = 12 = \frac{2m}{j} \implies j = \frac{100}{12} \implies j = 8,83 \text{ comme } j \text{ est entier, la cyclicité apparaît à}$$

$$j = 8 \text{ et est de période } T_8 = \frac{100}{8} \approx 12,5 \text{ mois.}$$

Si on choisit $m = 60$, la cyclicité de 12 mois apparaît au point $j = 10$.

Dans ce type de cas nous proposons la construction du tableau 4.2 suivant :

Tableau 4.2 – Relation entre les valeurs de j et de m

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
m	6	12	18	24	30	36	42	48	54	60	66	72	78	84	90	96	102	106



(Tous les m grisés donnent une cyclicité $T = 12$ mois et correspondent à la condition $\frac{n}{6} \leq m \leq \frac{n}{3}$.)

Selon le cycle que l'on veut faire apparaître, on peut donc choisir m de telle façon que les périodes de ces cycles correspondent à des valeurs entières de j .

E. Le spectre d'un processus ARMA

D'après le théorème de Wold, un processus stationnaire indéterminable est une combinaison linéaire infinie des valeurs passées d'un bruit blanc, c'est-à-dire un MA infini.

Il s'écrit pour x_t réel :

$$x_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} \text{ avec } \psi_j \text{ réel ; } \psi_0 = 1.$$

On constate avec cette écriture, que le processus x_t peut être considéré comme étant la sortie d'un filtre linéaire à mémoire infinie à un seul côté, où l'entrée est a_t le bruit blanc, la stabilité étant assurée par la condition de convergence en moyenne quadratique sur les ψ_j : $\sum \psi_j^2 < \infty$.

Les ψ_j constituent la fonction de réponse impulsionnelle du filtre.

La fonction de transfert du filtre peut donc être déterminée et sa fonction de gain aussi.

Elle s'écrit : $T_\lambda = |A(\lambda)|^2 = \left| \sum_{\tau=0}^{\infty} \psi_\tau e^{-i2\pi\lambda\tau} \right|^2$ avec $A(\lambda)$ la fonction de réponse en fréquence du filtre et $\psi_0 = 1$; $\lambda = \frac{\omega}{2\pi}$ pour $\lambda = [0 ; 1/2]$.

Par ailleurs, si a_t est l'entrée et x_t la sortie du filtre, il existe la relation suivante entre le spectre des deux processus : $f_{x_t}(\lambda) = T(\lambda) f_{a_t}(\lambda)$.

Or, le spectre du Bruit Blanc a_t est : $f_{a_t}(\lambda) = 2\sigma_{a_t}^2$ sur l'intervalle $\lambda \in \left[0, \frac{1}{2}\right]$.

De sorte que l'on a la relation :

$$f_{x_t}(\lambda) = 2\sigma_{a_t}^2 \left| \sum_{\tau=0}^{\infty} \psi_\tau e^{-i2\pi\lambda\tau} \right|^2$$

qui s'écrit :

$$\begin{aligned} f_{x_t}(\lambda) &= 2\sigma_{a_t}^2 (1 + \psi_1 e^{-i2\pi\lambda} + \psi_2 (e^{-i2\pi\lambda})^2 + \dots)^2 \\ &= 2\psi(e^{-i2\pi\lambda})^2 \sigma_{a_t}^2 \end{aligned}$$

où ψ est un polynôme d'opérateur et $e^{-i2\pi\lambda}$ joue le rôle de l'opérateur de retard B qui intervient dans la définition des processus linéaires finis ARMA(p, q). On en déduit le spectre théorique d'un processus x_t possédant une structure

$$\text{ARMA}(p, q) : x_t = \frac{\theta(B)}{\phi(B)} a_t$$

Il s'écrit :

$$f_{x_t}(\lambda) = 2\sigma_{a_t}^2 \frac{|\theta(e^{-i2\pi\lambda})|^2}{|\phi(e^{-i2\pi\lambda})|^2}$$

Soit :

$$f_{x_t}(\lambda) = 2\sigma_{a_t}^2 \frac{|1 - \theta_1 e^{-i2\pi\lambda} - \theta_2 e^{-i4\pi\lambda} - \dots - \theta_q e^{-i2\pi q\lambda}|^2}{|1 - \phi_1 e^{-i2\pi\lambda} - \phi_2 e^{-i4\pi\lambda} - \dots - \phi_p e^{-i2\pi p\lambda}|^2}$$

Le numérateur et le dénominateur peuvent être transformés en utilisant les relations d'Euler, ce qui conduit à l'écriture pratique suivante :

$$f_{x_t}(\lambda) = 2\sigma_{a_t}^2 \frac{1 + \sum_{j=1}^q \theta_j^2 - 2 \sum_{j=1}^q \theta_j \cos j2\pi\lambda + 2 \sum_{j=1}^{q(j<k)} \sum_{k=1}^{q(j<k)} \theta_j \theta_k \cos(k-j)2\pi\lambda}{1 + \sum_{j=1}^p \phi_j^2 - 2 \sum_{j=1}^p \phi_j \cos j2\pi\lambda + 2 \sum_{j=1}^{p(j<k)} \sum_{k=1}^{p(j<k)} \phi_j \phi_k \cos(k-j)2\pi\lambda}$$

Il est donc possible avec cette formule de connaître la forme théorique des spectres des processus ARMA, non saisonniers, saisonniers, non saisonniers et saisonniers à la fois.

Comme ces formes théoriques sont des formes analytiques, il est possible de les étudier mathématiquement et de faire des simulations graphiques pour en connaître leur représentation.

Enfin, on peut calculer la variance du processus ARMA. Nous pouvons, en effet, démontrer que $\sigma_x^2 = \int_0^{1/2} f_{x_t}(\lambda) d\lambda$.

Ainsi pour un MA(q) nous pouvons vérifier que : $\sigma_{x_t}^2 = \sigma_{a_t}^2 \left(1 + \sum_{j=1}^q \theta_j^2 \right)$ et

pour un AR(1) : $\sigma_{x_t}^2 = \frac{\sigma_{a_t}^2}{1 - \phi_1^2}$.

Exercice n° 4.3

Calcul de la puissance spectrale d'un processus AR(2)

Soit le processus AR(2) : $x_t = 0,9x_{t-1} - 0,4x_{t-2} + a_t$ avec $a_t \rightarrow (i.i.d.(0 ; \sigma_{a_t}^2 = 1))$

- Écrire le processus en utilisant l'opérateur de retard B . Le processus est-il stationnaire ?
- Écrire le spectre de puissance du processus puis étudier mathématiquement ses variations pour $\lambda \in [0 ; 1/2]$.

Solutions

a) Le processus s'écrit : $(1 - 0,9B + 0,4B^2)x_t = a_t$

Le polynôme d'opérateur du second degré a pour discriminant :

$$\Delta = 0,9^2 - 4 \times 0,4 = -0,79 = 0,79i^2 \implies B = \frac{0,9 \pm i\sqrt{0,79}}{0,8} = \frac{9}{8} \pm i\sqrt{\frac{0,79}{0,8}}$$

Ces racines complexes conjuguées ont pour module $\rho = \sqrt{\frac{81}{64} + \frac{0,79}{0,64}} = 1,58$ supérieur à 1. Le processus est donc stationnaire.

b) Le spectre d'un processus AR(2) s'écrit :

$$f_{x_t}(\lambda) = \frac{2\sigma_{a_t}^2}{1 + \phi_1^2 + \phi_2^2 - 2\phi_1(1 - \phi_2) \cos 2\pi\lambda - 2\phi_2 \cos 4\pi\lambda}$$

Soit avec $\phi_1 = 0,9$ et $\phi_2 = -0,4$ et $\sigma_{a_t}^2 = 1$.

$$f_{x_t}(\lambda) = \frac{2}{1,97 - 2,52 \cos 2\pi\lambda + 0,8 \cos 4\pi\lambda}$$

Sa dérivée est :

$$\begin{aligned} f'_{x_t}(\lambda) &= \frac{-2(2,52 \times 2\pi \sin 2\pi\lambda - 0,8 \times 4\pi \sin 4\pi\lambda)}{(1,97 - 2,52 \cos 2\pi\lambda + 0,8 \cos 4\pi\lambda)^2} \\ &= \frac{-4\pi(2,52 \sin 2\pi\lambda - 1,6 \sin 4\pi\lambda)}{(1,97 - 2,52 \cos 2\pi\lambda + 0,8 \cos 4\pi\lambda)^2} \end{aligned}$$

$$\begin{aligned} f'_{x_t}(\lambda) = 0 &\implies 2,52 \sin 2\pi\lambda - 1,6 \sin 4\pi\lambda = 0 \\ &2,52 \sin 2\pi\lambda - 3,2 \sin 2\pi\lambda \cos 2\pi\lambda = 0 \\ &\sin 2\pi\lambda(2,52 - 3,2 \cos 2\pi\lambda) = 0 \\ &\implies \sin 2\pi\lambda = 0 \text{ pour } \lambda = 0 \text{ et } \lambda = 1/2 \end{aligned}$$

$$\cos 2\pi\lambda = 2,52/3,2 \approx 0,7875 \approx \cos 38^\circ \approx \cos 0,6632 \implies \lambda \approx 0,106$$

Soit le tableau de variation suivant :

λ	0		0,106		1/2
$f'_{x_t}(\lambda)$	0	+	0	-	
$f_{x_t}(\lambda)$	8	\nearrow	11,25	\searrow	0,38

$$\lambda = 0,2 \implies f_{x_t}(\lambda) = 3,68 ; \lambda = 0,3 \implies f_{x_t}(\lambda) = 0,95 ; \lambda = 0,4 \implies f_{x_t}(\lambda) = 0,47$$

La figure 4.9 illustre la représentation du spectre du processus AR(2).

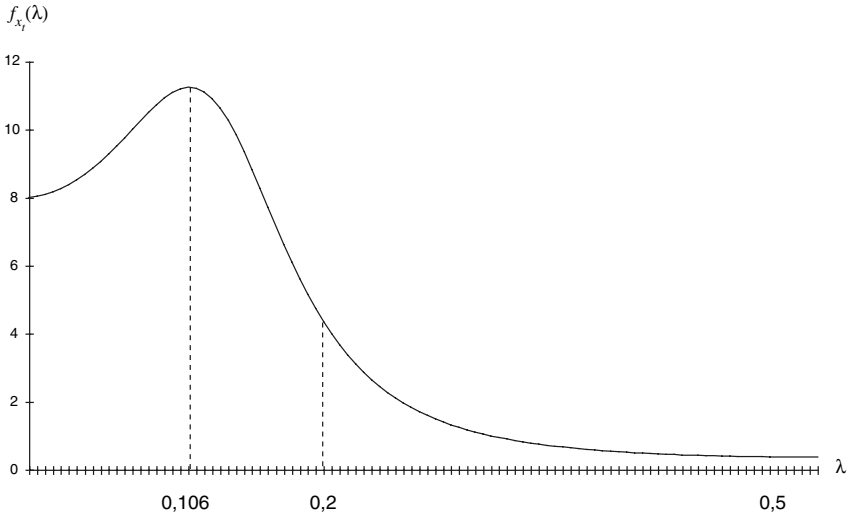


Figure 4.9 – Spectre du processus AR(2) : $x_t = 0,9x_{t-1} - 0,4x_{t-2} + a_t$

Exercice n° 4.4

Simulation graphique du spectre de différents processus ARMA

On demande d'étudier les processus ARMA suivants et de simuler graphiquement leur spectre pour $\lambda \in [0 ; 1/2]$ et avec a_t (i.i.d. $(0 ; \sigma_{a_t}^2 = 1)$).

- a) AR(1) : $x_t = 0,9x_{t-1} + a_t$
- b) AR(1) : $x_t = -0,9x_{t-1} + a_t$
- c) MA(1) : $x_t = a_t - 0,8a_{t-1}$
- d) MA(1) : $x_t = a_t + 0,8a_{t-1}$
- e) ARMA(1, 1) : $x_t = -0,3x_{t-1} + a_t + 0,4a_{t-1}$
- f) AR(2) : $x_t = 0,7x_{t-1} - 0,2x_{t-2} + a_t$
- g) AR(2) : $x_t = -0,7x_{t-1} - 0,4x_{t-2} + a_t$
- h) MA(2) : $x_t = a_t + 0,9a_{t-1} + 0,2a_{t-2}$
- i) MA(2) : $x_t = a_t - 0,5a_{t-1} + 0,8a_{t-2}$
- j) AR₁₂(1) : $x_t = 0,6x_{t-12} + a_t$
- k) MA₁₂(1) : $x_t = a_t + 0,6a_{t-12}$

Solutions

Nous n'étudions les fonctions spectrales que pour les processus AR(1) et MA(1), pour les autres processus nous ne donnons que la forme analytique générale de leur fonction.

- AR(1) : $f_{x_t}(\lambda) = \frac{2\sigma_{a_t}^2}{1 + \phi_1^2 - 2\phi_1 \cos 2\pi\lambda}$

$$f'_{x_t}(\lambda) = 8\pi\sigma_{a_t}^2 \frac{-\phi_1 \sin 2\pi\lambda}{(1 + \phi_1^2 - 2\phi_1 \cos 2\pi\lambda)^2}$$

Sur l'intervalle $\lambda \in [0 ; 1/2]$ quand $-1 < \phi_1 < 0$, $f_{x_t}(\lambda)$ est continûment croissante et quand $0 < \phi_1 < 1$, $f_{x_t}(\lambda)$ est continûment décroissante. Ce spectre admet un point d'inflexion de coordonnées :

$$\left\{ \begin{array}{l} \lambda_1 = \frac{1}{2\pi} \operatorname{Arccos} \left[\frac{-(1 + \phi_1^2) + \sqrt{(1 + \phi_1)^2 + 32\phi_1^2}}{4\phi_1} \right] \\ f_{x_t}(\lambda) = \frac{2\sigma_{a_t}^2}{\left(\frac{3(1 + \phi_1^2)}{2} \right) - \frac{\sqrt{(1 + \phi_1)^2 + 32\phi_1^2}}{2}} \end{array} \right.$$

- MA(1) : $f_{x_t}(\lambda) = 2\sigma_a^2(1 + \theta_1^2 - 2\theta_1 \cos 2\pi\lambda)$

avec $|\theta_1| < 1$ pour que le MA(1) soit inversible et $\lambda \in [0 ; 1/2]$.

$$f'_{x_t}(\lambda) = 2\sigma_{a_t}^2(4\pi\theta_1 \sin 2\pi\lambda) = 8\sigma_{a_t}^2\pi\theta_1 \sin 2\pi\lambda$$

Sur l'intervalle $\lambda \in [0; 1/2]$ quand $0 < \theta_1 < 1$, $f_{x_t}(\lambda)$ est continûment croissante et quand $1 < \theta_1 < 0$, $f_{x_t}(\lambda)$ est continûment décroissante. Ce spectre admet un point d'inflexion de coordonnées :

$$\begin{cases} \lambda_1 = 1/4 \\ f_{x_t}(\lambda) = 2\sigma_{a_t}^2(1 + \theta_1^2) \end{cases}$$

- ARMA(1, 1) : $f_{x_t}(\lambda) = 2\sigma_{a_t}^2 \frac{1 + \theta_1^2 - 2\theta_1 \cos 2\pi\lambda}{1 + \phi_1^2 - 2\phi_1 \cos 2\pi\lambda}$
- AR(2) : $f_{x_t}(\lambda) = \frac{2\sigma_{a_t}^2}{1 + \phi_1^2 + \phi_2^2 - 2\phi_1(1 - \phi_2)\cos 2\pi\lambda - 2\phi_2 \cos 4\pi\lambda}$
- MA(2) : $f_{x_t}(\lambda) = 2\sigma_{a_t}^2(1 + \theta_1^2 + \theta_2^2 - 2\theta_1(1 - \theta_2)\cos 2\pi\lambda - 2\theta_2 \cos 4\pi\lambda)$
- AR₁₂(1) : $f_{x_t}(\lambda) = \frac{2\sigma_{a_t}^2}{1 + \phi_{12}^2 - 2\phi_{12} \cos 12 \times 2\pi\lambda}$
- MA₁₂(1) : $f_{x_t}(\lambda) = 2\sigma_{a_t}^2(1 + \theta_{12}^2 - 2\theta_{12} \cos 12 \times 2\pi\lambda)$

La simulation des spectres à l'aide d'Eviews est effectuée par le programme suivant (C4EX4.PRG) :

```
create U 100 «Création de l'espace de travail
scalar PI = 3.14
gener t = @trend(1)+1
" Processus a
scalar PHI = 0.9
gener AR1 = 2/(1+PHI*PHI-(2*PHI*cos(2*PI*t/200)))
" Processus b
scalar PHI = -0.9
gener AR1b = 2/(1+PHI*PHI-(2*PHI*cos(2*PI*t/200)))
" Processus c
scalar TET = 0.8
gener MA1 = 2*(1+TET*TET-(2*TET*cos(2*PI*t/200)))
" Processus d
scalar TET = -0.8
gener MA1d = 2*(1+TET*TET-(2*TET*cos(2*PI*t/200)))
" Processus e
scalar TET = -0.4
scalar PHI = -0.3
gener num = (1+TET*TET-(2*TET*cos(2*PI*t/200)))
gener den = (1+PHI*PHI-(2*PHI*cos(2*PI*t/200)))
gener ARMA11 = 2*num/den
" Processus f
scalar PHI1 = 0.7
scalar PHI2 = -0.2
gener AR2 = 2/(1+PHI1^2+PHI2^2-(2*PHI1*(1-PHI2)*cos(2*PI*t/200))-(2*PHI2*cos(4*PI*t/200)))
" Processus g
scalar PHI1 = -0.7
scalar PHI2 = -0.4
gener AR2g = 2/(1+PHI1^2+PHI2^2-(2*PHI1*(1-PHI2)*cos(2*PI*t/200))-(2*PHI2*cos(4*PI*t/200)))
" Processus h
scalar TET1 = -0.9
scalar TET2 = -0.2
```

```

gener MA2 = 2*(1+TET1^2+TET2^2-(2*TET1*(1-TET2)*cos(2*PI*t/200))-(2*TET2*cos(4*PI*t/200)))
" Processus i
scalar TET1 = 0.5
scalar TET2 = -0.8
gener MA2i = 2*(1+TET1^2+TET2^2-(2*TET1*(1-TET2)*cos(2*PI*t/200))-(2*TET2*cos(4*PI*t/200)))
" Processus j
scalar PHI = 0.6
gener ARS1 = 2/(1+(PHI^2)-(2*PHI*cos(24*PI*t/200)))
" Processus k
scalar TET = -0.6
gener MAS1 = 2*(1+TET*TET-(2*TET*cos(24*PI*t/200)))

```

La figure 4.10 illustre les représentations des spectres.

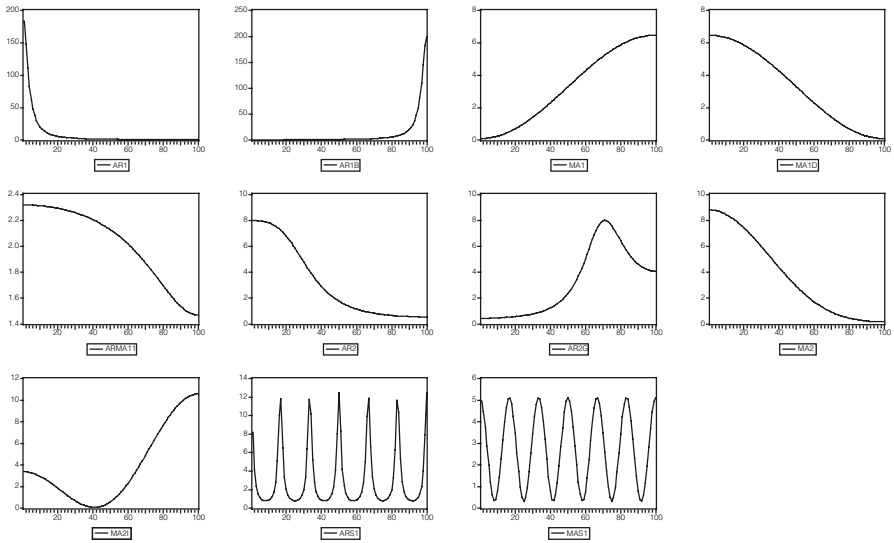


Figure 4.10 – Représentations des spectres des processus ARMA

Exercice n° 4.5

Estimations de spectres de chroniques simulées

On demande d'estimer les spectres des chroniques échantillons des processus suivants avec $a_t \rightarrow i.i.d.(0; 1)$:

- $x_{1t} = a_t (n = 1000)$
- $x_{2t} = 0,5t + 10 + a_t (n = 1000)$
- $x_{3t} = 10 \cos \frac{2\pi t}{12} + a_t (n = 1000)$
- $x_{4t} = 0,5t + 10 + 100 \cos \frac{2\pi t}{12} + a_t (n = 1000)$

$$e) x_{5t} = \begin{cases} 10 \cos \frac{2\pi t}{12} + 10 + a_t \text{ si } \cos \frac{2\pi t}{12} > 0 \\ 10 + a_t \text{ si } \cos \frac{2\pi t}{12} \leq 0 \end{cases} \quad (n = 1000)$$

f) $x_{6t} = a_t$ avec perturbations aux observations $n = 200, 211, 481, 600, 800, 900$ ($n = 1000$).

$$g) x_{7t} = 8 \cos \frac{2\pi t}{12} t + 5 \cos \frac{2\pi t}{3} t + a_t \quad (n = 1000)$$

$$h) x_{8t} = \begin{cases} 8 \cos \frac{2\pi t}{12} t + 5 \cos \frac{2\pi t}{3} t + 10 + a_t \text{ si } \left(8 \cos \frac{2\pi t}{12} t + 5 \cos \frac{2\pi t}{3} t \right) > 0 \\ 10 + a_t \text{ si } \left(8 \cos \frac{2\pi t}{12} t + 5 \cos \frac{2\pi t}{3} t \right) \leq 0 \end{cases} \quad (n = 1000)$$

Solutions

La figure 4.11 présente les spectres des huit processus simulés ; pour des raisons pratiques, seuls les j premiers points des spectres sont représentés graphiquement.

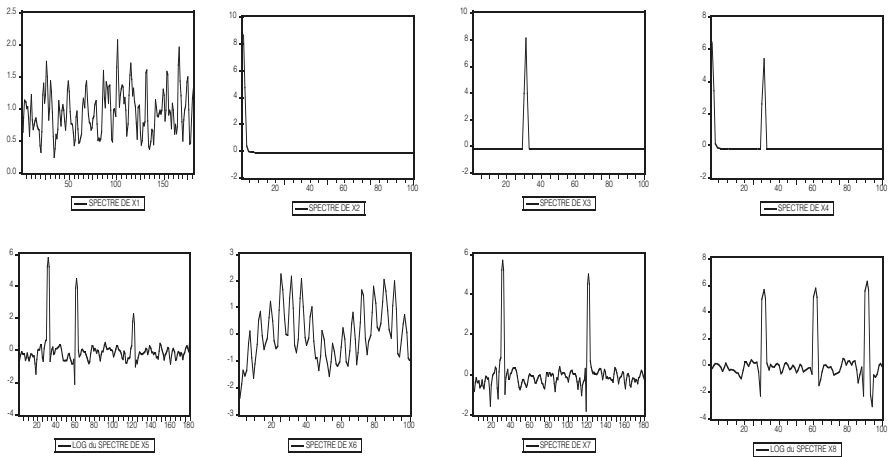


Figure 4.11 – Spectres des processus

Commentaires :

a) $x_{1t} = a_t$: Spectre sans pic significatif qui s'enroule autour d'une tendance plate ($m = 200$).

b) $x_{2t} = 0,5t + 10 + a_t$: Spectre « rouge » typique dans le cas d'une tendance dans une chronique ($m = 200$).

c) $x_{3t} = 10 \cos \frac{2\pi t}{12} + a_t$: Pic de période 12 ($j = 30$ pour $m = 180$) révélateur d'un cycle de même période dans la chronique.

d) $x_{4t} = 0,5t + 10 + 100 \cos \frac{2\pi t}{12} + a_t$: Combinaison du spectre rouge et du pic de période 12 (au point $j = 30$ pour $m = 180$).

$$\text{e) } x_{5t} = \begin{cases} = 10 \cos \frac{2\pi t}{12} + 10 + a_t \text{ si } \cos \frac{2\pi t}{12} > 0 \\ = 10 + a_t \text{ si } \cos \frac{2\pi t}{12} \leq 0 \end{cases} : \text{Apparition d'harmoniques de la}$$

période 12 pour « reconstituer la cyclicités en pics » (périodes 12, 6 et 3 aux points $j = 30, 60$ et 120 pour $m = 300$).

f) $x_{6t} = a_t$ (perturbations aux observations $n = 200, 211, 481, 600, 800, 900$) : Apparition de pics artificiels dus aux perturbations dans le bruit blanc ($m = 300$).

g) $x_{7t} = 8 \cos \frac{2\pi t}{12} + 5 \cos \frac{2\pi t}{3} + a_t$: double cycle de périodes 12 et 3 aux points $j = 30$ et $j = 120$ (pour $m = 300$).

$$\text{h) } x_{8t} = \begin{cases} 8 \cos \frac{2\pi t}{12} + 5 \cos \frac{2\pi t}{3} + 10 + a_t \text{ si } \left(8 \cos \frac{2\pi t}{12} + 5 \cos \frac{2\pi t}{3} \right) > 0 \\ 10 + a_t \text{ si } \left(8 \cos \frac{2\pi t}{12} + 5 \cos \frac{2\pi t}{3} \right) \leq 0 \end{cases}$$

Cyclicités de période 12 au point $j = 30$ (pour $m = 180$). Les harmoniques de ce cycle « masquent » le cycle de période 3 au point $j = 120$ en particulier l'harmonique de période 4 ($j = 90$).

5. Les processus aléatoires non stationnaires

Les chroniques économiques sont rarement des réalisations de processus aléatoires stationnaires. La non-stationnarité des processus peut concerner aussi bien le moment du premier ordre (espérance mathématique) que celui du second ordre (variance et covariance du processus). Dans la procédure de Box et Jenkins, la non-stationnarité est repérée graphiquement (tendance, cycle long, saisonnalité explosive, modification de structure...) ou encore au moyen de la fonction d'auto-corrélation et du spectre de la série brute (fonction d'autocorrélation lentement décroissante, spectre rouge...). Il n'existait pas avant les années 1980 d'analyse rigoureuse, de tests permettant de déceler cette non-stationnarité. En conséquence, les transformations utilisées étaient les filtres aux différences et la formule de Box Cox qui contient comme cas particulier le logarithmique des données. Dans ces conditions, il arrivait fréquemment que la transformation soit mal adaptée aux caractéristiques de la non-stationnarité ce qui avait pour effet d'introduire au sein de la chronique des mouvements parasites (par exemple l'artefact d'Adelman observé sur le spectre d'une série filtrée par les différences premières). Depuis les travaux de Nelson et Plosser (1982), les cas de non-stationnarité les plus fréquents sont analysés à partir de deux types de processus :

- les processus TS (*Trend Stationnary*) qui représentent une non-stationnarité de type déterministe¹ ;
- les processus DS (*Differency Stationnary*) pour les processus non stationnaires aléatoires. Ils constituent, pour des chroniques saisonnières ou non, les modèles de référence pour la construction à l'origine des tests de racine unitaire en l'absence d'hétéroscédasticité.

1. Par définition, un processus est aléatoire d'où l'ambiguïté du terme de processus déterministe.

I. Description des processus TS et DS

Nous n'envisageons dans cette section que le cas de processus non saisonniers.

A. Les processus TS

Un processus TS s'écrit : $x_t = f_t + \varepsilon_t$ où f_t est une fonction polynomiale du temps, linéaire ou non linéaire et ε_t un processus stationnaire de type ARMA.

Le processus TS le plus simple est représenté par une fonction polynomiale de degré 1. Ce processus s'écrit : $x_t = a_0 + a_1t + \varepsilon_t$.

Si ε_t est un bruit blanc (gaussien ou non), les caractéristiques de ce processus sont alors :

$$\begin{aligned}E[x_t] &= a_0 + a_1t + E[\varepsilon_t] = a_0 + a_1t \\V[x_t] &= 0 + V[\varepsilon_t] = \sigma_\varepsilon^2 \\cov[x_t, x_{t'}] &= 0 \text{ pour } t \neq t'\end{aligned}$$

Ce processus TS est non stationnaire car $E[x_t]$ dépend du temps. Comme cette espérance est égale à $a_0 + a_1t$, il s'agit à l'instant t d'un chiffre certain. Dans ce cas, nous pouvons estimer de façon efficace les paramètres a_0 et a_1 de la tendance, en utilisant la méthode des moindres carrés ordinaires (MCO) sur les couples de valeurs (x_t, t) : ces estimateurs possèdent la propriété d'être BLUE¹ et peuvent donc être employés par la suite pour réaliser une prévision de la chronique. Connaissant \hat{a}_0 et \hat{a}_1 , le processus x_t peut être stationnarisé en retranchant de la valeur de x_t en t , la valeur estimée $\hat{a}_0 + \hat{a}_1t$. Dans ce type de modélisation, l'effet produit par un choc (ou par plusieurs chocs aléatoires) à un instant t est transitoire. Le modèle étant déterministe, la chronique retrouvera son mouvement de long terme qui est ici la droite de tendance. Il est possible de généraliser cet exemple à des fonctions polynomiales de degré quelconque mais aussi à des formes f_t non linéaires ou encore à des modèles où ε_t est un processus de type ARMA.

B. Les processus DS

Les processus DS sont des processus que l'on peut rendre stationnaire par l'utilisation d'un filtre aux différences : $(1 - B)^d x_t = \beta + \varepsilon_t$ où ε_t est un processus stationnaire de type ARMA ou encore un bruit blanc, β une constante réelle et d l'ordre du filtre aux différences.

1. *Best Linear Unbiased Estimator*, meilleur estimateur linéaire sans biais cf. Bourbonnais, 2009, p. 55.

Ces processus sont souvent représentés en utilisant le filtre aux différences premières ($d = 1$). Le processus est dit alors processus du premier ordre. Il s'écrit :

$$(1 - B)x_t = \beta + \varepsilon_t$$

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

où ε_t est un processus stationnaire de type bruit blanc (gaussien ou non). L'introduction de la constante β dans le processus DS permet de définir deux processus différents :

- $\beta = 0$: le processus DS est dit sans dérive.

Il s'écrit : $x_t = x_{t-1} + \varepsilon_t$

Il s'agit d'un processus autorégressif d'ordre 1 avec paramètre $\phi_1 = 1$ ou encore d'une équation de récurrence du premier ordre. Puisque ε_t est un bruit blanc, ce processus DS porte le nom de modèle de marche au hasard ou de marche aléatoire (*Random Walk Model*). Un processus de marche au hasard (*random walk model*) est un processus sans mémoire et donc par nature non prévisible ; pour ce type de processus, on parle parfois de la marche de l'homme ivre (*drunken behavior*). Par exemple, les mouvements des cours des actions sont perçus comme semblables à une marche au hasard, c'est-à-dire à un comportement imprévisible. Il est très fréquemment utilisé pour analyser l'efficacité de nombreux marchés comme ceux des produits de base ou financiers. La prévision pour un tel modèle est qualifiée de naïve et sert de référence lors de comparaisons prévisionnelles.

Pour étudier les caractéristiques de ce modèle, écrivons-le sous sa forme moyenne mobile infinie :

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

$$x_{t-1} = x_{t-2} + \varepsilon_{t-1} \implies x_t = x_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

$$x_{t-2} = x_{t-3} + \varepsilon_{t-2} \implies x_t = x_{t-3} + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

etc.

Si le premier terme de la chronique est x_0 , le modèle s'écrit alors :

$$x_t = x_0 + \sum_{i=1}^t \varepsilon_i$$

Les caractéristiques de ce processus sont (en supposant x_0 certain) :

$$E[x_t] = x_0$$

$$V[x_t] = t\sigma_\varepsilon^2$$

$$\text{cov}[x_t, x_{t'}] = \sigma_\varepsilon^2 \times \text{Min}(t, t') \text{ si } t \neq t'$$

Par exemple :

$$\text{cov}(x_4, x_2) = E[(x_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4)(x_0 + \varepsilon_1 + \varepsilon_2)]$$

$$= E(\varepsilon_1 \times \varepsilon_1) + E(\varepsilon_2 \times \varepsilon_2) = 2\sigma_\varepsilon^2$$

En effet, tous les produits $E(\varepsilon_t \times \varepsilon_{t'}) = 0$ si $t \neq t'$.

Ce processus est non stationnaire en variance puisqu'elle dépend du temps. Cette non-stationnarité est aléatoire. Pour stationnariser la marche aléatoire, il suffit d'appliquer au processus le filtre aux différences premières :

$$x_t = x_{t-1} + \varepsilon_t \iff (1 - B)x_t = \varepsilon_t$$

Dans l'écriture développée de la marche au hasard, on constate que $\sum_{i=1}^t \varepsilon_i$ représente en t une accumulation de chocs aléatoires de 1 jusqu'à t . Chacun d'entre eux va donc avoir un effet permanent sur la série brute ce qui a pour conséquence d'éloigner la série de sa valeur initiale x_0 . De sorte que si, à un instant quelconque t , se produit un choc important sur la série, l'effet de cette impulsion éloigne la valeur de la série brute du point d'impact et l'effet est permanent (et donc non transitoire), cette persistance d'un choc est appelée hystérésis.

• $\beta \neq 0$: le processus porte alors le nom de processus DS avec dérive.

Il s'écrit :

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

Comme précédemment, on peut rechercher sa forme équivalente développée :

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

$$x_{t-1} = x_{t-2} + \beta + \varepsilon_{t-1} \implies x_t = x_{t-2} + 2\beta + \varepsilon_{t-1} + \varepsilon_t$$

$$x_{t-2} = x_{t-3} + \beta + \varepsilon_{t-2} \implies x_t = x_{t-3} + 3\beta + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

etc.

Si on suppose la valeur d'origine x_0 connue et déterministe, on a alors :

$$x_t = x_0 + \beta t + \sum_{i=1}^t \varepsilon_i$$

On peut analyser les caractéristiques de ce processus :

$$E[x_t] = x_0 + \beta t$$

$$V[x_t] = t\sigma_\varepsilon^2$$

$$\text{cov}[x_t, x_{t'}] = \sigma_\varepsilon^2 \times \text{Min}(t, t') \text{ si } t \neq t'$$

Le processus est non stationnaire de par son espérance et sa variance.

L'espérance étant de la même forme que celle d'un processus TS, on reconnaît dans ce processus une non-stationnarité déterministe et aléatoire à la fois.

La stationnarisation de ce processus est réalisée en utilisant le filtre aux différences premières :

$$x_t = x_{t-1} + \beta + \varepsilon_t \iff (1 - B)x_t = \beta + \varepsilon_t$$

Soit en employant la forme développée :

$$x_t = x_0 + \beta t + \sum_{i=1}^t \varepsilon_i \text{ et en calculant : } x_{t-1} = x_0 + \beta(t-1) + \sum_{i=1}^{t-1} \varepsilon_i$$

On a :

$$x_t - x_{t-1} = (1 - B)x_t = \beta + \varepsilon_t$$

Pour résumer :

- un processus TS est prévisible par extrapolation de la variable temps ;
- un processus DS est une marche au hasard (la meilleure prévision est la dernière valeur connue) si le corrélogramme des différences premières indique que ses termes sont non significativement différents de 0 ;
- un processus DS a une représentation ARMA en différences premières (un processus générateur existe donc) si le corrélogramme des différences premières indique que quelques termes sont significativement différents de 0.

C. Conséquences d'une mauvaise stationnarisation du processus

Pour stationnariser un processus TS, la bonne méthode est celle des moindres carrés ordinaires ; pour un processus DS il faut employer le filtre aux différences. Dans ce paragraphe, on s'intéresse aux implications d'un mauvais choix de la technique de stationnarisation pour un processus. Nous utilisons pour cela les études de Chan, Hayya et Ord (1977) et de Nelson C.R. et Kang (1981). Ils ont analysé les répercussions sur les résidus d'un mauvais choix de stationnarisation du processus de départ.

1) Conséquence sur un processus TS

Pour un processus TS, la bonne méthode de stationnarisation est celle des moindres carrés ordinaires. Supposons que l'on applique au processus TS du premier ordre un filtre aux différences premières. *A priori*, comme le degré du polynôme est 1, ce filtre peut être considéré comme correct puisqu'un filtre aux différences d'ordre d élimine un polynôme de même degré (cf. chapitre 4).

$$\begin{aligned}
x_t &= \beta_0 + \beta_1 t + \varepsilon_t \text{ avec } \varepsilon_t \longrightarrow i.i.d. (0, \sigma_\varepsilon^2) \\
x_{t-1} &= \beta_0 + \beta_1(t-1) + \varepsilon_{t-1} \\
x_t - x_{t-1} &= (1-B)x_t = \beta_1 + \varepsilon_t - \varepsilon_{t-1} \\
\Delta x_t &= \beta_1 + \eta_t
\end{aligned}$$

Par analogie avec le cas précédent, on peut calculer les caractéristiques de ce nouveau processus :

$$\begin{aligned}
E[\Delta x_t] &= \beta_1 \\
\text{cov}[\Delta x_t, \Delta x_{t-h}] &= \gamma_h \\
&= E\{[\Delta x_t - E(\Delta x_t)][\Delta x_{t-h} - E(\Delta x_{t-h})]\} \\
&= E\{[\eta_t][\eta_{t-h}]\} \\
&= E\{[\varepsilon_t - \varepsilon_{t-1}][\varepsilon_{t-h} - \varepsilon_{t-h-1}]\}
\end{aligned}$$

D'où :

$$\gamma_h = \begin{cases} 2\sigma_\varepsilon^2 & h = 0 \\ -\sigma_\varepsilon^2 & h = \{-1; +1\} \\ 0 & \text{ailleurs} \end{cases}$$

La fonction d'autocorrélation $\rho_h = \frac{\gamma_h}{\gamma_0}$ s'écrit alors ($\gamma_0 = 2\sigma_\varepsilon^2$) :

$$\rho_h = \begin{cases} 1 & h = 0 \\ -\frac{1}{2} & h = \{-1; +1\} \\ 0 & \text{ailleurs} \end{cases}$$

Le processus obtenu Δx_t n'a pas les caractéristiques d'un bruit blanc puisque γ_h (ou ρ_h) n'est pas la fonction d'autocovariance d'un bruit blanc. L'application du filtre aux différences a créé une perturbation artificielle.

Le résultat peut s'analyser d'une façon différente en faisant appel à l'analyse dans le domaine des fréquences. Le processus x_t qui contient une tendance déterministe a un spectre de type rouge, c'est-à-dire fortement relevé aux basses fréquences, de par l'existence de la tendance que l'analyse de Fourier assimile à une période de longueur infinie. Lorsque l'on applique la bonne méthode pour stationnariser la chronique, il reste ε_t qui doit être le spectre d'un bruit blanc (parallèle à l'axe des fréquences, cf. chapitre 4). Avec le filtre aux différences, on obtient $\Delta x_t = \beta_1 + \varepsilon_t - \varepsilon_{t-1} = y_t$ qui est un MA(1) non inversible avec constante.

Comme le spectre théorique d'un MA(1) est :

$$f(\lambda) = 2\sigma_\varepsilon^2(1 + \theta_1 - 2\theta_1 \cos 2\pi\lambda).$$

Pour $\theta_1 = 1$, le polynôme d'opérateur a une racine sur le cercle unité, soit :

$$f(\lambda) = 4\sigma_\varepsilon^2(1 - \cos 2\pi\lambda)$$

Sa dérivée $f'(\omega) = 8\pi\sigma_\varepsilon^2 \sin 2\pi\lambda$ a pour racines $\lambda = 0$ et $\lambda = 1/2$ et est strictement positive pour $\lambda \in [0; 1/2]$. La fonction $f(\lambda)$ est donc croissante ; ce n'est pas le spectre d'un bruit blanc.

2) Conséquence sur un processus DS

La bonne méthode de stationnarisation pour un DS du premier ordre est le filtre aux différences premières.

Supposons que l'on applique la méthode des moindres carrés ordinaires (régression sur le temps) sur les observations d'un échantillon du processus. Le résidu de cette régression doit être un bruit blanc.

Nelson et Kang montrent que, dans ce cas, les résidus sont de moyenne nulle, mais leur covariance dépend de la taille de l'échantillon et du temps. À partir de simulations, ils indiquent que l'élimination d'une tendance linéaire sur un processus de marche aléatoire crée artificiellement une forte autocorrélation des résidus pour les premiers retards et donc introduit un mouvement pseudo-périodique d'ampleur d'autant plus importante que le nombre d'observations de la chronique de départ est élevé. Dans le cas où les ε_t sont autocorrélés positivement (ε_t obéit à un AR(1) dans les simulations des auteurs) la concentration fallacieuse aux basses fréquences du spectre des résidus est renforcée, alors qu'elle est atténuée dans le cas d'une autocorrélation négative.

En conclusion, quand un processus TS linéaire est considéré comme un processus DS, il y a amplification des valeurs spectrales de fréquences hautes qui indique l'introduction dans la chronique traitée d'un mouvement cyclique court. *A contrario*, lorsqu'une série DS est statistiquement traitée comme un TS, le spectre de la série transformée possède une concentration de ses puissances spectrales autour des faibles fréquences. Il y a création, dans la chronique transformée, d'un mouvement cyclique long. Ces résultats montrent sur le plan statistique et en particulier pour calculer une prévision, que le choix d'un processus DS ou TS comme structure de la chronique échantillon n'est pas neutre.

Les conséquences sont également importantes sur le plan économique. Considérons, par exemple, le PIB d'un pays comme la France en valeur réelle¹. Si ce PIB est DS plutôt que TS, il est alors nécessaire de remettre en cause la décomposition traditionnelle (tendance et cycle) et sa justification théorique : l'indépendance des schémas explicatifs. Si le PIB est en effet DS, la croissance et le cycle sont liés et ne peuvent être en conséquence étudiés de façon séparés. Or, d'après les travaux de Nelson et Plosser (1982) sur des chroniques macroéconomiques américaines, la variabilité constatée de la composante conjoncturelle serait due à une structure DS. Comme jusqu'à présent, l'analyse de cette composante s'effectue à partir du résidu d'une régression entre le PIB et une tendance déterministe, cette analyse surestime l'amplitude du cycle et sous-estime l'importance de la tendance.

1. Cf. Ertur, 1992.

Sur ce constat, Beveridge et Nelson (1981) proposent une décomposition des processus selon une tendance stochastique (permanente) qui obéit à une marche aléatoire avec ou sans dérive et une composante stationnaire (transitoire). Par la suite Harvey (1988) utilise les modèles structurels à composantes inobservables (modèle tendance plus cycle et tendance-cycle) représentées sous forme d'un modèle espace d'états estimé par le filtre de Kalman.


Exercice n° 5.1

Conséquence d'une mauvaise stationnarisation d'un processus

Dans cet exercice on demande de générer, sur 200 observations, un processus aléatoire de type bruit blanc gaussien de variance 25, puis de :

- générer une tendance déterministe de paramètres $a_0 = 5$ et $a_1 = 1$,
- générer une tendance stochastique avec dérive de terme constant 5,
- stationnariser ces deux processus selon les deux méthodes adaptées et non adaptées.

Solution

Le corrigé de cet exercice est réalisé à l'aide de Eviews, le fichier est en téléchargement  **Fichier C5EX1.PRГ** :

Le Figure 5.1 illustre l'allure des deux processus générés (YD = processus DS avec dérive, YT = processus TS), nous remarquons la similitude de l'allure de chacun des graphiques : en effet, le processus avec dérive est DS et TS à la fois. Le simple examen du graphique ne permet donc pas de se déterminer sur le type de processus, il faut recourir à des tests statistiques.

L'examen des deux corrélogrammes des processus stationnarisés par les différences premières indique un bruit blanc pour DYD (cf. figure 5.2) car tous les termes appartiennent à l'intervalle de confiance. En revanche, le corrélogramme du processus DYT (cf. figure 5.3) a son premier terme significativement différent de 0, le processus DYT n'est donc pas un bruit blanc.

L'examen des deux corrélogrammes des processus stationnarisés par régression sur le temps indique un bruit blanc pour RES1, résidu du processus YT, (cf. figure 5.4) car tous les termes appartiennent à l'intervalle de confiance. En revanche, le corrélogramme du processus RES2, résidu du processus YD, (cf. figure 5.5) est caractéristique d'un processus non stationnarisé. En effet, les termes du corrélogramme décroissent très faiblement, RES2 n'est donc pas un bruit blanc.

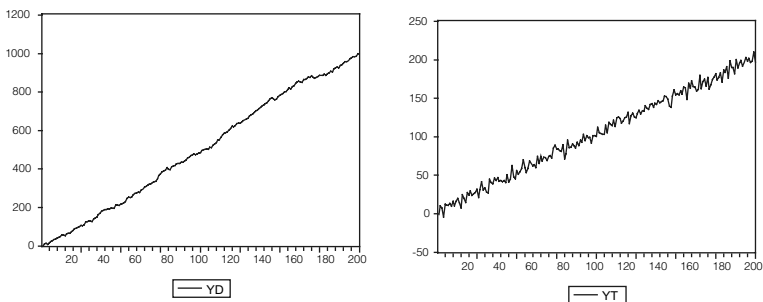


Figure 5.1 – Les deux processus générés

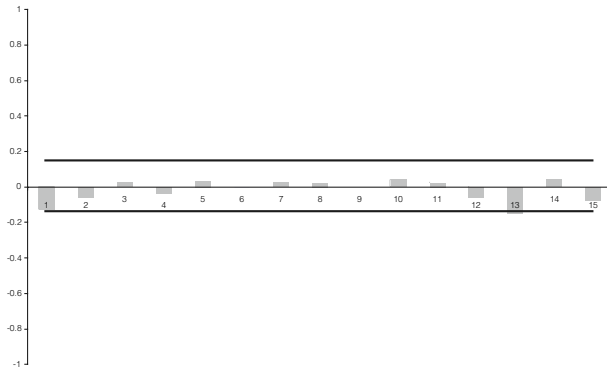


Figure 5.2 – Corrélogramme du processus DYD

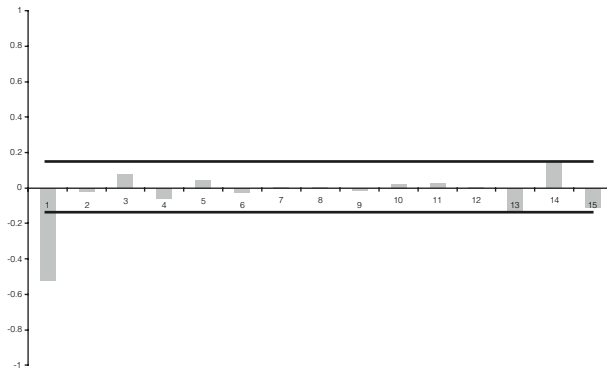


Figure 5.3 – Corrélogramme du processus DYT

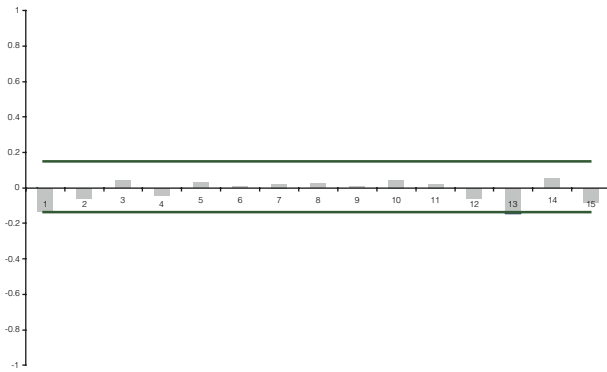


Figure 5.4 – Corrélogramme du processus RES1

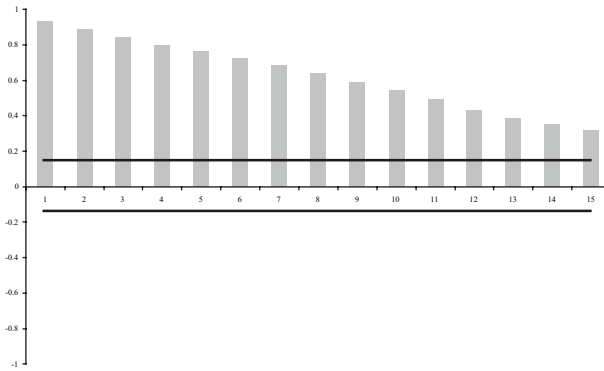


Figure 5.5 – Corrélogramme du processus RES2

II. Tests de racines unitaires non saisonnières

Il apparaît évident, à l'issue de la section précédente, que les structures DS et TS jouent un rôle très important dans le traitement statistique d'une chronique. Comment choisir entre l'une ou l'autre des structures ? Les tests de recherche de racine unitaire dans les processus générateurs tentent de répondre à cette question.

Les tests de racines unitaires sont maintenant très nombreux, ils peuvent être rassemblés selon l'hypothèse testée.

- *Les tests standards* pour lesquels l'hypothèse H_0 concerne la présence d'une racine unitaire contre l'hypothèse H_1 de stationnarité du processus. Ceci concerne les tests de Dickey-Fuller simple (1979), Dickey-Fuller Augmenté (1981) qui sont à l'origine de ce type d'études. Ils ont connu par la suite des extensions comme le test non paramétrique de Phillips-Perron (1988), le test de Seo (1999) qui traite explicitement de l'hétéroscédasticité dans les processus ou encore le test de Dickey et Pantula (1987) pour la recherche de plusieurs racines unitaires.
- *Les tests efficaces*. Dans les tests précédents l'un des modèles de base prend en compte une tendance déterministe. Plusieurs auteurs ont montré que l'élimination de cette composante peut apporter un gain « efficace » dans la recherche des racines unitaires. Les premiers tests de ce type ont été présentés par Sargan et Bharsava (1983) et Schmidt et Phillips (1992). Les plus utilisés sont cependant ceux de Elliot, Rothenberg et Stock (1996) et de Ng et Perron (2001).
- *Les tests avec rupture structurelle dans la chronique échantillon*. Perron (1989) montre que la présence d'un changement structurel de tendance dont le processus générateur est stationnaire autour d'une tendance linéai-

re a pour conséquence d'introduire un biais dans le test de racine unitaire. L'auteur propose donc un test de racine unitaire avec changement structurel (exogène) de tendance. D'autres tests ont, par la suite, été développés en particulier ceux de Zivot et Andrews (1992) avec estimation des dates de changement structurel (endogène) de tendance et de Perron et Rodriguez (2003) avec estimation de plusieurs changements structurels de tendance.

- *Les tests de stationnarité avec hypothèse H_0 la stationnarité.* Ces tests ont donc pour hypothèse H_1 la non stationnarité. Nous pouvons citer le test de Kwiatkowski, Phillips, Schmidt et Shin (1992) souvent utilisé lorsqu'il existe une autocorrélation des résidus des modèles en niveau ou en tendance estimés. Nous pouvons mentionner aussi les tests de Leybourne et Mc Cabe (1994), de Xiao (2001) ou encore celui de Paparoditis et Politis (2003).

Dans cet ouvrage nous présentons tout d'abord en introduction le test de Dickey – Fuller Simple, puis le test de Dickey – Fuller Augmenté qui permettent de mettre en évidence le caractère stationnaire ou non d'une chronique par la détermination d'une tendance déterministe ou stochastique. Nous décrivons par la suite les tests les plus utilisés en pratique : Phillips et Perron puis le test de Dickey et Pantula. Puis les tests de Elliot, Rothenberg et Stock et de Ng et Perron et enfin le test KPSS (Kwiatkowski *et alii.*).

A. Les tests de Dickey-Fuller simples

1) Les modèles de base

Les modèles servant de base à la construction de ces tests sont au nombre de trois :

$$[1] (1 - \phi_1 B)x_t = a_t \quad \text{Modèle autorégressif d'ordre 1 : AR(1).}$$

$$[2] (1 - \phi_1 B)(x_t - \mu) = a_t \quad \text{Modèle AR(1) avec constante, où } E[x_t] = \mu$$

$$[3] (1 - \phi_1 B)(x_t - \alpha - \beta t) = a_t \quad \text{Modèle AR(1) avec tendance}$$

et $a_t \rightarrow i.i.d. (0; \sigma_a^2)$

Le principe des tests est le suivant : si dans l'un de ces trois modèles $\phi_1 = 1$, alors le polynôme d'opérateur du processus contient une racine unitaire (« *Unit Root* ») $B = 1$. Le processus, d'après le théorème de Doob, est alors non stationnaire. Le choix de la spécification dépend des données traitées.

2) Caractéristiques des modèles

- Modèle [1]

$$(1 - \phi_1 B)x_t = a_t \iff x_t = \phi_1 x_{t-1} + a_t$$

Soit l'hypothèse $H_0 : \phi_1 = 1$

Le modèle s'écrit alors sous cette hypothèse :

$$(1 - B)x_t = a_t \iff x_t = x_{t-1} + a_t \iff x_t = x_0 + \sum_{i=1}^t a_i$$

Il s'agit d'un modèle DS de « promenade aléatoire » dont les caractéristiques non stationnaires ont été décrites précédemment :

Si H_0 est acceptée alors le processus stationnarisé est :

$$\Delta x_t = a_t$$

Si l'hypothèse alternative $H_1 : |\phi_1| < 1$ est acceptée, alors x_t est un AR(1) stationnaire.

- Modèle [2]

$$(1 - \phi_1 B)(x_t - \mu) = a_t$$

$$x_t = \phi_1 x_{t-1} + c + a_t \text{ où } c = \mu(1 - \phi_1)$$

Sous $H_0 : \phi_1 = 1$.

Le modèle s'écrit (avec $c = 0$ par conséquent) : $x_t = x_{t-1} + a_t$

C'est un modèle DS de promenade aléatoire (soit le modèle [1]).

On le rend stationnaire par la transformation : $\Delta x_t = a_t$

Sous l'hypothèse alternative : $H_1 : |\phi_1| < 1$ on a un AR(1) stationnaire avec constante.

- Modèle [3]

$$(1 - \phi_1 B)(x_t - \alpha - \beta t) = a_t$$

$$x_t = \phi_1 x_{t-1} + bt + c + a_t$$

avec $c = \alpha(1 - \phi_1) + \phi_1 \beta$ et $b = \beta(1 - \phi_1)$

Sous $H_0 : \phi_1 = 1$

Alors $b = 0$ et $c = \beta$ et le modèle s'écrit : $x_t = x_{t-1} + \beta + a_t$

Il s'agit d'une promenade aléatoire avec dérive (processus DS avec dérive).

Sa solution est de la forme : $x_t = x_0 + \beta t + \sum_{i=1}^t a_i$

Il est non stationnaire, déterministe et aléatoire à la fois (cf. I.B). On le rend stationnaire par le filtre aux différences premières : $\Delta x_t = \beta + a_t$

Avec l'hypothèse alternative $H_1 : |\phi_1| < 1$, le modèle peut s'écrire, en posant $\varepsilon_t = x_t - \alpha - \beta t : (1 - \phi_1 B)\varepsilon_t = a_t$.

Considérons alors le modèle : $x_t = \alpha + \beta t + \varepsilon_t$. Comme $\varepsilon_t = \frac{1}{1 - \phi_1 B} a_t$ est un processus ARMA (partie MA infinie), le processus x_t est un processus TS avec erreur ARMA ; on peut le rendre stationnaire en calculant les écarts par rapport à la tendance estimée par les moindres carrés ordinaires.

En résumé, le tableau 5.1 synthétise la typologie des modèles.

Tableau 5.1 – Typologie des modèles

	Modèle 1	Modèle 2	Modèle 3
H_0	DS sans dérive. Stationnaire à l'ordre 1. Non stationnaire à l'ordre 2.	DS sans dérive. Stationnaire à l'ordre 1. Non stationnaire à l'ordre 2.	DS avec dérive. Non stationnaire à l'ordre 1 et 2.
H_1	AR(1)	AR(1) avec constante	TS avec erreurs ARMA

3) Principe des tests de Dickey-Fuller

Sous l'hypothèse H_0 , le processus x_t n'est pas stationnaire quel que soit le modèle retenu. Les règles habituelles de l'inférence statistique ne peuvent alors pas être appliquées pour tester cette hypothèse, en particulier la distribution de Student du paramètre ϕ_1 . Dickey et Fuller ont étudié la distribution asymptotique de l'estimateur ϕ_1 sous l'hypothèse H_0 . À l'aide de simulations de Monte-Carlo ils ont tabulé les valeurs critiques pour des échantillons de tailles différentes.

En 1976, Fuller considère le modèle [1] : $x_t = \phi_1 x_{t-1} + a_t$ avec $x_0 = 0$ et $a_t \rightarrow i.i.d. (0; \sigma_a^2)$. Il note $\hat{\phi}_1$ l'estimateur du maximum de vraisemblance de ϕ_1 . On peut vérifier que :

$$\hat{\phi}_1 - \phi_1 = \frac{\sum_{t=2}^n x_{t-1} a_t}{\sum_{t=2}^n (x_{t-1})^2}.$$

Ce qui permet à l'auteur d'énoncer le lemme suivant.

Si $\phi_1 = 1$ alors le modèle [1] s'écrit $x_t = x_0 + \sum_{i=1}^t a_i$. Soit $\hat{\phi}_1$ l'estimateur des moindres carrés alors : $(\hat{\phi}_1 - 1) = O_p(n^{-1})$ ou encore : $\forall \varepsilon > 0, \exists M > 0, M \in R$ tel que $P[|\hat{\phi}_1 - 1| \geq M \times n^{-1}] < \varepsilon$

À partir de ce lemme l'auteur démontre la convergence de $(\widehat{\phi}_1 - 1)$, l'estimateur des MCO du modèle : $\Delta x_t = (\phi_1 - 1)x_{t-1} + a_t$. Cet estimateur converge plus rapidement vers sa vraie valeur que dans le cas $|\phi_1| < 1$. Rubin (1950) a montré que $\widehat{\phi}_1$ est un estimateur convergent quelle que soit la valeur de ϕ_1 , mais cette statistique n'obéit pas à une loi normale. C'est pourquoi Fuller a utilisé des simulations, par la méthode de Monte-Carlo, pour obtenir les valeurs critiques. Par la suite Dickey et Fuller ont étendu ces résultats aux modèles [2] et [3] en démontrant la convergence de $(\phi_1^* - 1)$ et $(\widetilde{\phi}_1 - 1)$ estimateurs des MCO de $(\phi_1 - 1)$. Les auteurs ont choisi de tester la valeur $(\widehat{\phi}_1 - 1)$ au lieu de $\widehat{\phi}_1$ pour des raisons purement statistiques. Cela n'est pas gênant pour le test. En effet, $x_t = \phi_1 x_{t-1} + a_t$ s'écrit aussi :

$$x_t - x_{t-1} = \phi_1 x_{t-1} - x_{t-1} + a_t$$

$$\Delta x_t = (\phi_1 - 1)x_{t-1} + a_t$$

Il est donc équivalent de tester comme hypothèse $H_0 : \phi_1 = 1$ ou $\phi_1 - 1 = \rho = 0$, l'hypothèse alternative H_1 est ici $|\phi_1| < 1$ ou encore le processus x_t est stationnaire. Ils construisent alors des « t statistiques » analogues à celles de Student du modèle linéaire général. Le déroulement des tests est alors le suivant :

- estimation par les moindres carrés ordinaires du paramètre $\phi_1 - 1 = \rho$ des trois modèles : $\widehat{\rho}, \rho^*, \widetilde{\rho}$;
- l'estimateur par les moindres carrés ordinaires fournit $t_{\widehat{\rho}}, t_{\rho^*}, t_{\widetilde{\rho}}$ avec par exemple $t_{\widehat{\rho}} = \frac{\widehat{\rho} - \widehat{\phi}_1 - 1}{\widehat{\sigma}_{\widehat{\rho}}}$.

Si $t_{\widehat{\rho}} > t_{\text{tabulé}}$, alors on accepte l'hypothèse H_0 , il existe une racine unité (cf. table 1 en fin d'ouvrage).

Le test se déroule de manière identique pour les modèles de type [2] et [3].

Il convient cependant de noter que dans le cas de l'acceptation de H_0 , on ne peut pas appliquer les tests de Student classiques aux autres coefficients (b et c) mais il faut recourir à des tables statistiques tabulées par simulation (cf. tables 2, 4 et 5 en fin d'ouvrage).

Remarque : les principaux logiciels d'analyse de séries temporelles calculent automatiquement les valeurs critiques de $t_{\widehat{\rho}}$.

Ces tests révèlent l'existence d'une racine unitaire mais restent insuffisants pour discriminer entre les processus TS et DS surtout pour les modèles [2] et [3] qui font intervenir d'autres variables explicatives que x_{t-1} . Ils sont, par conséquent, complétés par les tests d'hypothèses jointes.

4) Les tests d'hypothèses jointes

Les tests d'hypothèses jointes concernent les modèles [2] et [3]. Ils permettent de détecter une racine unitaire et de distinguer les processus DS des processus TS linéaires.

- Modèle [2]

$$x_t = \phi_1 x_{t-1} + c + a_t \text{ où } c = \mu(1 - \phi_1)$$

On effectue le test de recherche d'une racine unitaire, puis le test d'hypothèse jointe noté $H_0^1 : (c; \phi_1) = (0; 1)$ contre l'hypothèse H_1^1 .

Pour effectuer ce test, on calcule la statistique empirique F_1

$$F_1 = \frac{(SCR_c - SCR_2)/2}{SCR_2/(n-2)}$$

avec :

SCR_c = somme des carrés des résidus du modèle [2] contraint sous l'hypothèse H_0^1 , soit

$$SCR_c = \sum_t (x_t - x_{t-1})^2.$$

SCR_2 = somme des carrés des résidus du modèle [2] non contraint estimé par les MCO.

n = nombre d'observations utilisées pour estimer les paramètres du modèle¹.

Le degré de liberté du numérateur ($SCR_c - SCR_2$) est égal à $(n-0) - (n-2)$, puisqu'aucun paramètre n'est à estimer dans le modèle contraint. Dans le cas où l'hypothèse H_0^1 est acceptée la nullité de c peut résulter d'une moyenne μ nulle de la série temporelle. On réalise donc un test de normalité de cette dernière même si les séries économiques (non centrées) sont rarement de moyenne nulle.

Cette statistique F_1 est analogue à une loi de Fisher. Dickey et Fuller ont tabulé les valeurs critiques de sa distribution empirique et asymptotique (il en est de même pour les statistiques F_2 et F_3 qui suivent). Si F_1 est supérieur à la valeur Φ_1 lue dans la table à un seuil α , on rejette l'hypothèse H_0^1 à un seuil α (cf. table 3 en fin d'ouvrage).

1. Certains auteurs considèrent que n est le nombre d'observations disponibles initialement, cependant la régression sur un modèle autorégressif entraîne la perte d'une observation. C'est pourquoi, pour ces auteurs, dans la statistique F_1 le degré de liberté pour SCR_2 est parfois égal à $n-1-2 = n-3$.

- Modèle [3]

Avec :

$$x_t = \phi_1 x_{t-1} + bt + c + a_t$$

$$b = \beta(1 - \phi_1)$$

$$c = \alpha(1 - \phi_1) + \phi_1\beta$$

Après le test de racine unitaire, on effectue pour ce modèle les tests des deux hypothèses jointes suivantes :

$H_0^2 : (c; b; \phi_1) = (0; 0; 1)$ contre l'hypothèse l'un des paramètres est différent.

$H_0^3 : (c; b; \phi_1) = (c; 0; 1)$ contre l'hypothèse l'un des paramètres est différent.

Pour l'hypothèse H_0^2 , on calcule la statistique F_2 :

$$F_2 = \frac{(SCR_c - SCR_3)/3}{SCR_3/(n-3)}$$

Pour l'hypothèse H_0^3 , on calcule la statistique F_3 :

$$F_3 = \frac{(SCR_c^3 - SCR_3)/2}{SCR_3/(n-3)}$$

avec :

SCR_c = somme des carrés des résidus du modèle [3] contraint sous l'hypothèse H_0^2 ,

$SCR_c^3 = \sum_t (x_t - x_{t-1} - \hat{c})^2$ somme des carrés des résidus du modèle [3]

contraint sous l'hypothèse H_0^3 ,



SCR_3 = somme des carrés des résidus du modèle [3] non contraint estimé par les MCO.

n = nombre d'observations utilisées pour estimer les paramètres du modèle.

Les degrés de liberté du numérateur ($SCR_c - SCR_3$) ou ($SCR_c^3 - SCR_3$) sont égaux à $[(n-0) - (n-3)]$ ou $[(n-1) - (n-3)]$, compte tenu du nombre de paramètres à estimer.

Ces statistiques F_2 et F_3 sont analogues à une loi de Fisher, on se réfère donc aux tables de Dickey-Fuller. Si F_2 est supérieur à la valeur lue dans la table Φ_2 (cf. table 6), on rejette l'hypothèse H_0^2 à un seuil α . Si F_3 est supérieur à la valeur lue dans la table Φ_3 , on rejette l'hypothèse H_0^3 à un seuil α (cf. table 7).

Conclusion : Nous constatons que pour réaliser un test de Dickey et Fuller, le résultat n'est pas identique selon l'utilisation de l'un des trois modèles comme processus générateur de la chronique de départ. Les conclusions auxquelles on parvient sont donc différentes et peuvent entraîner des transformations erronées.

C'est la raison pour laquelle Dickey et Fuller et à leur suite d'autres auteurs, ont élaboré des stratégies de tests. Celles-ci sont nombreuses mais ont été néanmoins synthétisées et présentées de façon exhaustive par Ertur (1992). Nous proposons (cf. Figure 5.6) une stratégie de test qui résume notre propos et présente l'avantage de la simplicité. Cette stratégie est facilement automatisable à l'aide des outils de programmation de Rats ou Eviews. Les programmes Eviews permettant de mener les stratégies de test de Dickey-Fuller ou de Dickey-Fuller Augmenté sont en téléchargement :  **C5-STRATEGIETEST-DF.PRG** et  **C5-STRATEGIETEST-DFA.PRG**.

B. Les tests de Dickey et Fuller augmentés

1) Transformations des modèles de base

Dans les modèles précédents, utilisés pour les tests de Dickey-Fuller simples, le processus est, par hypothèse, un bruit blanc. Or il n'y a aucune raison pour que, *a priori*, l'erreur soit non corrélée ; on appelle tests de Dickey-Fuller Augmentés (DFA, 1981) la prise en compte de cette hypothèse. Considérons par exemple le modèle [1], il peut s'écrire dans ce cas :

$$(1 - \phi_1 B)x_t = z_t \text{ où } z_t \text{ est un processus AR}(p - 1).$$

$$z_t = \sum_{i=1}^{p-1} \theta_i z_{t-i} + a_t \text{ où } a_t \longrightarrow i.i.d. (0; \sigma_a^2).$$

z_t s'écrit aussi en utilisant le polynôme d'opérateur : $\theta_{p-1}(B)z_t = a_t$. En multipliant les deux membres du processus AR(1) autocorrélé par $\theta_{p-1}(B)$, on a : $\theta_{p-1}(B)(1 - \phi_1 B)x_t = a_t$ ou encore l'AR(p) :

$$x_t = (\theta_1 + \phi_1)x_{t-1} + \dots + (\theta_{p-1} - \theta_{p-2}\phi_1)x_{t-p+1} - \theta_{p-1}\phi_1 x_{t-p} + a_t \quad [A]$$

Le processus x_t obéissant à un AR(1) à erreurs autocorrélées d'ordre ($p - 1$) est équivalent à un AR(p) à erreurs non autocorrélées, le processus a été blanchi. Les tests de Dickey - Fuller simples peuvent donc lui être appliqués. Cependant l'écriture du modèle Δx_t est plus complexe en raison de la présence des θ_j .

- Soit l'AR (1)

$$x_t = \alpha_1 x_{t-1} + a_t$$

$$x_t - x_{t-1} = \Delta x_t = (\alpha_1 - 1)x_{t-1} + a_t \text{ avec } a_t \longrightarrow i.i.d. (0; \sigma_a^2)$$

- Le processus AR (2)

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + a_t$$

$$x_t - x_{t-1} = \alpha_1 x_{t-1} - x_{t-1} + \alpha_2 x_{t-1} - \alpha_2 x_{t-2} + \alpha_2 x_{t-2} + a_t$$

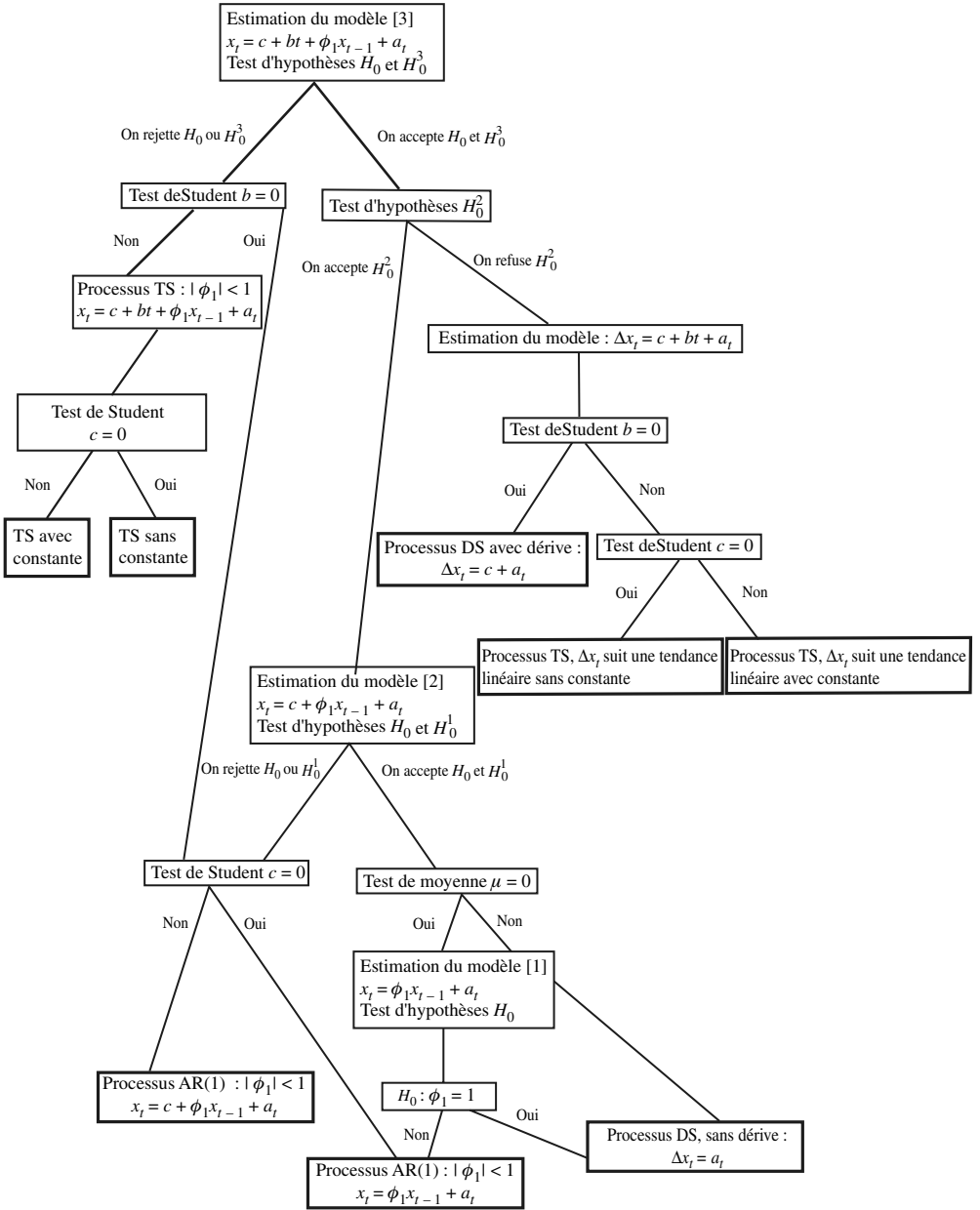


Figure 5.6 – Exemple d'une stratégie des tests de Dickey-Fuller simples

$$\Delta x_t = (\alpha_1 + \alpha_2 - 1)x_{t-1} - \alpha_2(x_{t-1} - x_{t-2}) + a_t$$

$$\Delta x_t = (\alpha_1 + \alpha_2 - 1)x_{t-1} - \alpha_2 \Delta x_{t-1} + a_t$$

• Le processus AR (3)

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \alpha_3 x_{t-3} + a_t$$

$$x_t - x_{t-1} = \alpha_1 x_{t-1} - x_{t-1} + \alpha_2 x_{t-1} - \alpha_2 x_{t-1} + \alpha_2 x_{t-2} + \alpha_3 x_{t-1} - \alpha_3 x_{t-1} \\ + \alpha_3 x_{t-3} + \alpha_3 x_{t-2} - \alpha_3 x_{t-2} + a_t$$

$$\Delta x_t = (\alpha_1 + \alpha_2 + \alpha_3 - 1)x_{t-1} - \alpha_2(x_{t-1} - x_{t-2}) - \alpha_3(x_{t-1} - x_{t-2}) \\ - \alpha_3(x_{t-2} - x_{t-3}) + a_t$$

$$\Delta x_t = (\alpha_1 + \alpha_2 + \alpha_3 - 1)x_{t-1} - (\alpha_2 + \alpha_3)\Delta x_{t-1} - \alpha_3 \Delta x_{t-2} + a_t$$

On en déduit par récurrence l'écriture de l'AR (p) :

$$x_t = \alpha_1 x_{t-1} + \dots + \alpha_p x_{t-p} + a_t$$

$$\Delta x_t = \left(\sum_{k=1}^p \alpha_k - 1 \right) x_{t-1} - \sum_{j=1}^{p-1} \left(\sum_{k=j+1}^p \alpha_k \right) \Delta x_{t-j} + a_t \quad [\text{B}]$$

En utilisant pour $\alpha_1, \alpha_2, \dots, \alpha_p$ les coefficients de l'écriture [A], on a :

$$\alpha_1 = \theta_1 + \phi_1$$

$$\alpha_2 = \theta_2 - \theta_1 \phi_1$$

...

$$\alpha_{p-1} = \theta_{p-1} - \theta_{p-2} \phi_1$$

$$\alpha_p = -\theta_{p-1} \phi_1$$

L'expression [B] s'écrit alors :

$$\Delta x_t = (\theta_1 + \phi_1 + \theta_2 - \theta_1 \phi_1 + \dots + \theta_{p-1} - \theta_{p-2} \phi_1 - \theta_{p-1} \phi_1 - 1)x_{t-1} \\ - \sum_{j=1}^{p-1} \left(\sum_{k=j+1}^p \alpha_k \right) \Delta x_{t-j} + a_t$$

$$\Delta x_t = [\phi_1(1 - \theta_1 - \dots - \theta_{p-1}) + \theta_1 + \theta_2 + \dots + \theta_{p-1} - 1]x_{t-1} \\ - \sum_{j=1}^{p-1} \left(\sum_{k=j+1}^p \alpha_k \right) \Delta x_{t-j} + a_t$$

$$= [(\phi_1 - 1)(1 - \theta_1 - \dots - \theta_{p-1})]x_{t-1} - \sum_{j=1}^{p-1} \left(\sum_{k=j+1}^p (\theta_k - \theta_{k-1} \Phi_j) \right) \Delta x_{t-j} + a_t$$

que l'on peut écrire :

$$\Delta x_t = \rho x_{t-1} - \sum_{j=2}^p \Phi_j \Delta x_{t-j+1} + a_t$$

Tester la présence de la racine unité, revient donc à tester la significativité du coefficient qui se trouve devant x_{t-1} comme pour le test de Dickey-Fuller simple.

D'où l'hypothèse H_0 testée :

$$H_0 : \rho = (\phi_1 - 1)(1 - \theta_1 - \dots - \theta_{p-1}) = 0 \iff \phi_1 - 1 = 0 \text{ ou } \phi_1 = 1$$

Ainsi, quand les erreurs sont corrélées dans le modèle [1], on transforme le modèle de départ selon l'écriture précédente (noté alors modèle [4]) qui met en évidence la racine unité sous H_0 ce qui implique la même procédure de test Dickey-Fuller simple. Les démonstrations précédentes s'étendent aux modèles [2] et [3] appelés respectivement modèles [5] et [6].

En définitive, la conduite du test est identique au cas précédent : recherche de la présence de racine unité, test d'hypothèse jointe et enfin stratégie de test.

2) Principes du test DFA et tests d'hypothèses jointes

Soit : $z_t = \sum_{i=1}^{p-1} \theta_i z_{t-i} + a_t$ avec $a_t \rightarrow i.i.d.(0; \sigma_a^2)$ ou encore $\theta(B) z_t = a_t$ et $\theta(B)$ un polynôme d'opérateur de degré $(p - 1)$.

a) Spécifications des modèles

- Modèle [4]

$$(1 - \phi_1 B)x_t = z_t$$

Sous l'hypothèse nulle d'une racine unitaire $\phi_1 = 1$, le modèle s'écrit alors :

$$x_t = x_{t-1} + z_t \text{ ou encore } \Delta x_t = z_t \text{ soit : } \theta(B) \Delta x_t = a_t$$

Il s'agit d'un processus DS autocorrélé, caractérisé par une non-stationnarité stochastique et stationnaire en différences.

Sous l'hypothèse alternative $|\phi_1| < 1$, le processus x_t est un $AR(p)$; l'écriture du modèle est alors : $A(B) x_t = a_t$ avec $A(B) = \theta(B) (1 - \phi_1 B)$ un polynôme retard d'ordre p . Il s'agit donc d'un processus $AR(p)$ asymptotiquement stationnaire (cf. paragraphe précédent).

- Modèle [5]

$$(1 - \phi_1 B)(x_t - \mu) = z_t$$

qui peut s'écrire :

$$\theta(B)(1 - \phi_1 B)(x_t - \mu) = a_t$$

$$\theta(B)(1 - \phi_1 B)x_t = \theta(B)(1 - \phi_1 B)\mu + a_t$$

$$\theta(B)(1 - \phi_1 B)x_t = c + a_t \text{ avec } c = \theta(B)(1 - \phi_1 B)\mu$$

Sous l'hypothèse nulle d'une racine unitaire $\phi_1 = 1$, il vient :

$$x_t = x_{t-1} + z_t \text{ ou encore } \Delta x_t = z_t \text{ soit : } \theta(B)\Delta x_t = a_t$$

Il s'agit d'un processus DS autocorrélé, caractérisé par une non-stationnarité stochastique et stationnaire en différences avec $c = 0$.

Sous l'hypothèse alternative $|\phi_1| < 1$, le modèle s'écrit :

$$\theta(B)(1 - \phi_1 B)(x_t - \mu) = a_t \text{ ou encore :}$$

$A(B)(x_t - \mu) = a_t$ avec $A(B) = \theta(B)(1 - \phi_1 B)$ un polynôme de retards d'ordre p .

$$A(B)x_t = c + a_t \text{ avec } c = A(B)\mu \neq 0.$$

Il s'agit donc d'un processus AR(p) avec constante asymptotiquement stationnaire.

- Modèle [6]

$$(1 - \phi_1 B)(x_t - \alpha - \beta t) = z_t$$

Soit :

$$\theta(B)(1 - \phi_1 B)(x_t - \alpha - \beta t) = a_t$$

Une démonstration analogue à la précédente permet d'écrire le modèle sous la forme :

$$\theta(B)(1 - \phi_1 B)x_t = c + bt + a_t$$

avec $b = \theta(B)(1 - \phi_1)\beta$ et $c = \theta(B)[\alpha(1 - \phi_1) + \phi_1\beta]$.

Sous l'hypothèse nulle d'une racine unitaire $\phi_1 = 1$, il vient :

$$x_t = x_{t-1} + \beta + z_t \text{ ou encore } \Delta x_t = \beta + z_t. \text{ Soit : } \theta(B)\Delta x_t = c' + a_t \text{ avec } c' = \beta\theta(B)$$

Il s'agit d'un processus DS avec dérive autocorrélé, caractérisé par une non-stationnarité de nature mixte : déterministe et stochastique.

Sous l'hypothèse alternative $|\phi_1| < 1$, le processus est de type TS caractérisé par une tendance purement déterministe ; il s'écrit :

$$(1 - \phi_1 B) = \alpha' + \beta't + z_t$$

ou encore :

$$A(B)x_t = \alpha'\theta(B) + \beta'\theta(B)t + a_t \text{ avec } A(B) = \theta(B)(1 - \phi_1 B).$$

b) Procédures de tests

Les tests ADF sont fondés, sous l'hypothèse alternative $|\phi_1| < 1$, sur l'estimation par les MCO des trois modèles (cf. paragraphe précédent) :

$$\text{Modèle [4] : } \Delta x_t = \rho x_{t-1} - \sum_{j=2}^p \Phi_j \Delta x_{t-j+1} + a_t$$

$$\text{Modèle [5] : } \Delta x_t = \rho x_{t-1} - \sum_{j=2}^p \Phi_j \Delta x_{t-j+1} + c + a_t$$

$$\text{Modèle [6] : } \Delta x_t = \rho x_{t-1} - \sum_{j=2}^p \Phi_j \Delta x_{t-j+1} + c + bt + a_t$$

avec $a_t \rightarrow i.i.d. (0; \sigma_a^2)$.

Le test se déroule de manière similaire aux tests DF simples. Les tables utilisées sont celles récemment tabulées par MacKinnon (1991). En effet, en procédant à un grand nombre de simulations MacKinnon estime pour des échantillons de tailles différentes et pour n'importe quel nombre de variables explicatives les valeurs critiques pour les tests de racine unitaire. Eviews indique directement les valeurs critiques issues de ces tables de MacKinnon.

Détermination du nombre de décalages : La difficulté, pour utiliser ce type de test, réside dans la connaissance *a priori* du degré de l'AR de z_t . Le nombre de retards p doit être choisi pour éliminer l'autocorrélation des z_t . La méthode la plus simple consiste à retenir une valeur de p importante et à procéder par élimination progressive en fonction de la significativité des retards (les tests de Student classiques s'appliquent dans ce cadre). Des auteurs préconisent d'effectuer les tests en prenant trois retards : 4, 8 et 12 et de comparer par la suite les résultats au moyen des critères AIC par exemple (cf. chapitre 7). Enfin, la valeur de p peut être déterminée selon les critères d'information de Akaike ou Schwarz. En partant d'une valeur de p suffisamment élevée, nous estimons les différents modèles jusqu'à $p = 0$ puis nous retenons la valeur de p qui minimise les critères de Akaike ou Schwarz. Si la valeur $p = 0$ est retenue nous utilisons alors le test de Dickey-Fuller simple.

3) Tests d'hypothèses jointes

Après avoir recherché l'existence d'une racine unitaire, on effectue les tests d'hypothèses jointes.

- Premier test

Soit l'hypothèse jointe notée¹ $H_0^4 : (c; \rho) = (0; 0)$ dans le modèle [5] contre l'hypothèse alternative H_1^4 au moins un des éléments est différent.

Pour effectuer ce test, on calcule la statistique empirique F_4

$$F_4 = \frac{(SCR_C - SCR_S)/2}{SCR_S/(n - (p - 1) - 2)}.$$

(Le $p - 1$ du dénominateur correspond au fait que la somme démarre à $j = 2$ et donc si, par exemple, $p = 3$ cela entraîne l'existence de deux variables explicatives).

SCR_C = somme des carrés des résidus du modèle [5] contraint par H_0^4 :

$$\Delta x_t = \sum_{j=2}^p \hat{\Phi}_j \Delta x_{t-j+1} + a_t, \text{ soit } SCR_C = \left(\Delta x_t - \sum_{j=2}^p \hat{\Phi}_j \Delta x_{t-j+1} \right)^2$$

SCR_S = somme des carrés des résidus du modèle [5] estimé par MCO.

1. Cette hypothèse jointe correspond à celle $H_0^1 : (c; \phi_1) = (0; 1)$ du modèle [2].

n = nombre d'observations réellement utilisées pour estimer les paramètres du modèle.

Cette statistique F_4 est analogue à une loi de Fisher, elle a été tabulée par Dickey-Fuller (1981). Si F_4 est supérieur à la valeur lue dans la table $\Phi_{2;n-p-1}^\alpha$, on rejette l'hypothèse H_0^4 à un seuil α .

• Deuxième et troisième tests

Ils correspondent aux hypothèses¹ nulles du modèle [6] :

$$H_0^5 : (c; b; \rho) = (0; 0; 0)$$

$$H_0^6 : (c; b; \rho) = (c; 0; 0)$$

Pour l'hypothèse H_0^5 , on calcule la statistique F_5 :

$$F_5 = \frac{(SCR_C - SCR_6)/3}{SCR_6/(n - (p - 1) - 3)}$$

Si F_5 est supérieur à la valeur lue dans la table $\Phi_{2;n-p-2}^\alpha$, on rejette l'hypothèse H_0^5 à un seuil α .

Pour l'hypothèse H_0^6 , on calcule la statistique F_6 :

$$F_6 = \frac{(SCR_D - SCR_6)/2}{SCR_6/(n - (p - 1) - 3)}$$

SCR_D = somme des carrés des résidus du modèle [6] contraint par H_0^6 :

$$\Delta x_t = \sum_{j=2}^p \Phi_j \Delta x_{t-j+1} + c + a_t, \text{ soit}$$

$$SCR_D = \sum_t \left(\Delta x_t - \sum_{j=2}^p \hat{\Phi}_j \Delta x_{t-j+1} - \hat{c} \right)^2$$

SCR_6 = somme des carrés des résidus du modèle [6] estimé par MCO.

Si F_6 est supérieur à la valeur lue dans la table $\Phi_{3;n-p-2}^\alpha$, on rejette l'hypothèse H_0^6 à un seuil α .

Nous pouvons alors nous référer à la stratégie (cf. figure 5.6) de test présentée précédemment.

1. Ces hypothèses jointes correspondent à celles $H_0^2 : (c; b; \phi_1) = (0; 0; 1)$ et $H_0^3 : (c; b; \phi_1) = (c; 0; 1)$ des modèles [2] et [3].

Exercice n° 5.2

 fichier C5EX2

Exemple d'application des tests DF et DFA aux dépenses en produit pharmaceutiques

On demande d'appliquer la stratégie des tests DF et DFA aux dépenses (CVS) en produits pharmaceutiques en France connues mensuellement sur 14 ans.

Solution

Soit les dépenses (CVS) de produits pharmaceutiques en France dont la Figure 5.7 illustre l'évolution.

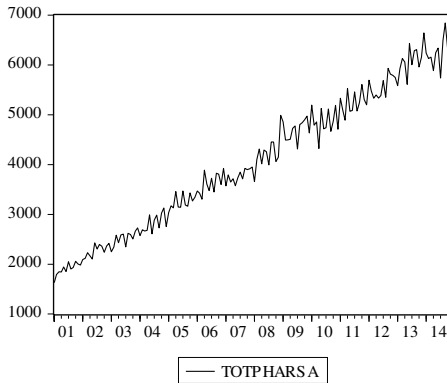


Figure 5.7 – Évolution sur 14 ans des dépenses CVS en produits pharmaceutiques en France

a) Test de Dickey-Fuller simple

Nous appliquons la stratégie de test de la Figure 5.6, l'estimation du modèle [3] $x_t = c + bt + \phi_1 x_{t-1} + a_t$ conduit aux résultats suivants :

Dependent Variable : TOTPHARSA				
Included observations : 167				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1847.351	134.6594	13.71870	0.0000
TOTPHARSA(-1)	-0.084314	0.078141	-1.078994	0.2822
@TREND	30.84197	2.242282	13.75472	0.0000
R-squared	0.980289	Sum squared resid	6315765	

• Test H_0 de racine unitaire :

La statistique du t empirique $= \frac{\tilde{\phi}_1 - 1}{\hat{\sigma}_{\tilde{\phi}_1}} = \frac{-0,0843 - 1}{0,078} = -13,87$ est très inférieure à

la valeur lue dans la table à 1 % (= -4,04), on rejette donc H_0 . L'hypothèse de racine unitaire est, bien entendu, rejetée.

- Test d'hypothèse jointe H_0^3

Soit le calcul de $F_3 = \frac{(SCR_c^3 - SCR_3)/2}{SCR_3/(n-3)}$

avec : $SCR_c^3 = \sum_t (x_t - x_{t-1} - \hat{c})^2 = 13731662,67$ somme des carrés des résidus du

modèle [3] contraint sous l'hypothèse H_0^3 , soit le modèle estimé suivant : $\Delta(\text{TOTPHARSA}) = 30,36 + e_t$

$SCR_3 = 6315765$ (somme des carrés des résidus du modèle [3] non contraint estimé par les MCO).

$n = 167$ (nombre d'observations utilisées pour estimer les paramètres du modèle).

Nous obtenons $F_3 = \frac{(13731662,67 - 6315765)/2}{6315765/(167-3)} = 96,28$ qui est très supérieure à la

valeur lue dans la table à 1 % ($\Phi_3 = 8,73$), on rejette donc H_0^3 .

Nous nous situons donc sur la partie de gauche de la stratégie de test de la Figure 5.6, le coefficient b (coefficient de la pente de la tendance) est significativement différent de 0, nous pouvons conclure à un processus TS.

b) Test de Dickey-Fuller augmenté

Soit l'estimation du modèle [6], d'après la stratégie de test, nous utilisons directement la fonction disponible dans Eviews pour déterminer le nombre de décalages ($p = 2$) par minimisation du critère de Akaike.

La statistique du t empirique ($= -5,40$) est inférieure à la valeur lue dans la table à 1 % ($= -4,01$), on rejette donc H_0 . L'hypothèse de racine unitaire est rejetée.

Nous procédons maintenant au test d'hypothèses jointes conformément à la stratégie de test (avec $p = 3$). Dans le logiciel EVIEWS, il faut directement rentrer la valeur $p - 1 = 2$.

ADF Test Statistic	-5.409548	1 % Critical Value*	-4.0162	
		5 % Critical Value	-3.4377	
		10 % Critical Value	-3.1428	
*MacKinnon critical values for rejection of hypothesis of a unit root.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable : D(TOTPHARSA)				
Included observations : 165				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
TOTPHARSA(-1)	-0.752872	0.139175	-5.409548	0.0000
D(TOTPHARSA(-1))	-0.315487	0.110114	-2.865087	0.0047
D(TOTPHARSA(-2))	-0.347998	0.074468	-4.673124	0.0000
C	1.31E+09	2.32E+08	5.639168	0.0000
@TREND(1981:01)	21453671	3964453.	5.411509	0.0000
R - squared	0.982175	Sum squared resid	5524702	

$$H_0^6 : (c; b; \rho) = (c; 0; 0), \text{ soit le calcul de } F_6 = \frac{(SCR_D - SCR_6)/2}{SCR_6/(n - (p - 1) - 3)}$$

avec $SCR_D = 6537019$ (somme des carrés des résidus du modèle estimé

$$\Delta x_t = \sum_{j=1}^2 a_j \Delta x_{t-j} + c + a_t \text{ soit :}$$

Dependent Variable : DTOTPHARSA				
Included observations : 165				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	69.58873	15.94813	4.363441	0.0000
D(TOTPHARSA(-1))	-0.819789	0.063348	-12.94108	0.0000
D(TOTPHARSA(-2))	-0.596633	0.063348	-9.418286	0.0000
R-squared	0.523292	Sum squared resid	6537019	

$SCR_6 = 5524702$ (somme des carrés des résidus du modèle [6] estimé par MCO) :

$$F_6 = \frac{(6537019 - 5524702)/2}{5524702/(165 - 2 - 3)} = 14,65 \text{ qui est supérieur à la valeur lue dans la table à}$$

1 % ($\Phi_3 = 8,73$), on rejette donc H_0^6 .

Les tests DF et DFA de racine unitaire, nous font rejeter l'hypothèse H_0 , les coefficients de la tendance et du terme constant sont significativement différents de 0. Le processus est de type TS (tendance déterministe). La « bonne méthode de stationnarisation » consiste à calculer le résidu issu de la régression : $TOTPHARSA = \hat{a}_0 + \hat{a}_1 t + e_t$

Un « choc » sur les dépenses pharmaceutiques (déremboursement, promotion des généralistes...) a donc un impact instantané et n'affecte pas la pente de la tendance de long terme.

Le programme Eviews (📄 **fichier C5EX2.PRG**) permettant d'effectuer l'ensemble de ces calculs est en téléchargement.

C. Le test de Phillips et Perron

Le test de Phillips et Perron (1988) est construit sur une correction non paramétrique des statistiques de Dickey-Fuller pour prendre en compte des erreurs hétéroscédastiques et/ou autocorrélées. Il se déroule en quatre étapes :

- estimation par les moindres carrés ordinaires des trois modèles de base des tests de Dickey – Fuller et calcul des statistiques associées, soit e_t le résidu estimé ;
- estimation de la variance dite de court terme des résidus $\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n e_t^2$;
- estimation d'un facteur correctif s_t^2 (appelé variance de long terme) établi à partir de la structure des covariances des résidus des modèles précé-

demment estimés de telle sorte que les transformations réalisées conduisent à des distributions identiques à celles du Dickey-Fuller standard :

$$s_t^2 = \frac{1}{n} \sum_{i=1}^n e_i^2 + 2 \sum_{i=1}^l \left(1 - \frac{i}{l+1}\right) \frac{1}{n} \sum_{t=i+1}^n e_t e_{t-i}.$$

Pour estimer cette variance de long terme, il est nécessaire de définir un nombre de retards l (troncature de Newey-West) estimé en fonction du nombre d'observations n , $l \approx 4(n/100)^{2/9}$;

– calcul de la statistique de *PP* : $t_{\hat{\phi}_1}^* = \sqrt{k} \times \frac{(\hat{\phi}_1 - 1)}{\hat{\sigma}_{\hat{\phi}_1}} + \frac{n(k-1)\hat{\sigma}_{\hat{\phi}_1}}{\sqrt{k}}$ avec

$$k = \frac{\hat{\sigma}^2}{s_t^2}$$

(qui est égal à 1 – de manière asymptotique – si e_t est un bruit blanc). Cette statistique est à comparer aux valeurs critiques de la table de MacKinnon.

Nous remarquons que si $k = 1$ alors $t_{\hat{\phi}_1}^* = \frac{(\hat{\phi}_1 - 1)}{\hat{\sigma}_{\hat{\phi}_1}}$, nous retrouvons la statistique de Dickey-Fuller.

La stratégie de test de la figure 5.6 reste valide. Ce test se ramène dans le cas où les résidus sont homoscédastiques au test de Dickey-Fuller simple.

Il est à noter que les logiciels RATS et Eviews permettent directement l'utilisation de ces tests.

Le test de Hall (1990) est complémentaire du test précédent qui peut ne pas être valide lorsque l'ordre autorégressif devient grand alors que l'ordre de la partie moyenne mobile reste fini. Ce test fait appel pour le calcul à une variable dichotomique.

Le test de Ouliaris, Park et Phillips (1989) est une extension du test de Phillips et Perron lorsqu'il existe une tendance polynomiale de degré k quelconque et le test de Perron (1989) où cette tendance est linéaire par morceaux.

D. Le test de Dickey et Pantula (1987)

Dans le cas de l'existence, dans un processus, d'une racine unitaire détectée par le test de Dickey-Fuller, il s'agit de vérifier l'absence d'autres racines unitaires en réitérant ces tests sur les différences premières du processus. Dickey et Pantula ont montré que cette procédure, qualifiée de séquentielle ascendante (Barthélemy, 1996), peut se révéler fautive car les distributions statistiques sont différentes selon l'existence d'une ou deux racines unitaires.

Les auteurs proposent donc une procédure séquentielle descendante qui permet de tester simultanément l'existence de plusieurs racines unitaires en se référant aux tables de Dickey-Fuller.

Considérons par exemple le test concernant deux racines unitaires, pour cela on spécifie le modèle AR (2) suivant :

$$(1 - \phi_1 B)(1 - \phi_2 B)x_t = a_t \text{ avec } a_t \longrightarrow i.i.d. (0; \sigma_a^2)$$

Soit $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} - \phi_1 \phi_2 x_{t-2} + a_t$

Pour faire apparaître le polynôme $(1 - B)^2$ qui permet de stationnariser le processus, on ajoute aux membres de l'égalité la quantité $-2x_{t-1} + x_{t-2}$ et on obtient :

$$\Delta^2 x_t = (1 - B)^2 x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} - \phi_1 \phi_2 x_{t-2} - 2x_{t-1} + x_{t-2} + a_t$$

Soit : $\Delta^2 x_t = (\phi_1 \phi_2 - 1)\Delta x_{t-1} - (\phi_1 - 1)(\phi_2 - 1)x_{t-1} + a_t$

En posant $\theta_1 = -(\phi_1 - 1)(\phi_2 - 1)$ et $\theta_2 = (\phi_1 \phi_2 - 1)$ on obtient le modèle suivant :

$$\Delta^2 x_t = \theta_2 \Delta x_{t-1} + \theta_1 x_{t-1} + a_t$$

Le test se déroule alors en deux étapes.

- Étape 1 : On teste l'hypothèse nulle de deux racines unitaires contre l'hypothèse alternative d'une seule racine unitaire, soit $H_0 : \theta_1 = \theta_2 = 0$ contre $H_1 : \theta_1 = 0$. Sous l'hypothèse H_1 , on effectue la régression $\Delta^2 x_t = \theta_2 \Delta x_{t-1} + a_t$.
 - si $t_{\theta_2} \leq t^{tab}$ on rejette H_0 et on passe à l'étape 2 ;
 - si $t_{\theta_2} > t^{tab}$ on accepte H_0 , le processus contient deux racines unitaires.
- Étape 2 : On teste l'hypothèse nulle de la présence d'une racine unitaire contre l'hypothèse alternative d'aucune racine unitaire (processus stationnaire) dans $\Delta x_t = (\phi_1 - 1)x_{t-1} + a_t$. Ce test sur $(\phi_1 - 1)$ est équivalent à celui sur θ_1 du modèle $\Delta^2 x_t$.

Ce test de Dickey et Pantula est donc simple à mettre en œuvre en particulier avec les logiciels Eviews ou RATS.

E. Le test KPSS (1992)

Kwiatkowski, Phillips, Schmidt et Shin (1992) proposent un test fondé sur l'hypothèse nulle de stationnarité. Après estimation des modèles [2] ou [3], on

calculer la somme partielle des résidus : $S_t = \sum_{i=1}^t e_i$ et on estime la variance de long terme (s_t^2) comme pour le test de Phillips et Perron.

La statistique est alors $LM = \frac{1}{s_t^2} \frac{\sum_{t=1}^n S_t^2}{n^2}$. On rejette l'hypothèse de stationnarité si cette statistique est supérieure aux valeurs critiques suivantes :

$n = \infty$	1 %	5 %	10 %
Modèle [2]	0,739	0,463	0,347
Modèle [3]	0,216	0,146	0,119

F. Le test de Elliot, Rothenberg et Stock (1996)

Ce test, noté ERS, de point optimal de racine unitaire repose sur une quasi-différenciation de la série x_t , il permet d'améliorer la puissance du test par rapport aux tests standards de Dickey-Fuller.

Soit : $Y_t = x_t - \alpha x_{t-1}$ et $y_t = z_t - \alpha z_{t-1}$

Avec $z_t = 1$ dans le cas sans dérive (modèle [1] du test de Dickey-Fuller), $z_t = 1, \dots, t$ dans le cas avec dérive (modèle [2]), et $\alpha = 1 - c/n$ où $c = 7$ si la série ne semble pas contenir de tendance et $c = 13,5$ si elle semble en contenir une.

Soit le modèle : $Y_t = y_t a + \varepsilon_t$. Nous procédons à l'estimation de a par les Moindres Carrés Généralisés : $Y_t = y_t \hat{a} + e_t$. Ce qui permet de calculer la nouvelle variable : $Z_t = x_t - \hat{a}' z_t$

Enfin, l'estimation du modèle : $\Delta Z_t = \beta_0 Z_{t-1} - \sum_{j=1}^p \beta_j \Delta Z_{t-j} + a_t$ permet de calculer la statistique $ERS = t$ statistique du coefficient β_0 (rapport du coefficient à son écart type) sous l'hypothèse $H_0 : \beta_0 = 0$.

Cette statistique ERS est à comparer aux valeurs tabulées par Elliot *et alii.*, si $ERS > ERS_{tab}$ nous acceptons l'hypothèse H_0 dans le cas contraire elle est rejetée.

G. Le test Ng-Perron (2001)

Ng et Perron proposent quatre tests statistiques fondés sur la méthode des moindres carrés généralisés sur les données hors tendances (*détrendées*). Ces tests sont une version modifiée des tests de Phillips-Perron (1988), Bhargava (1986) et ERS (1996) qui sont robustes aux distorsions de taille lorsque les résidus sont corrélés négativement.

Soit le terme : $k = \frac{\sum_{t=2}^n (x_{t-1}^d)^2}{n^2}$ avec x_t^d la série x_t hors tendance (*détrendée*).

Les quatre statistiques sont définies par :

$$\text{MZa} = \frac{(n^{-1}(x_t^d)^2 - f_0)}{2k}$$

$$\text{MZt} = \text{MZa} \times \text{MSB}$$

$$\text{MSB} = \left(\frac{k}{f_0} \right)^{\frac{1}{2}}$$

$$\text{MPT} = \frac{(c^2k - cn^{-1}(x_t^d)^2)}{f_0} \text{ si } x_t \text{ ne semble pas contenir de tendance.}$$

$$\text{MPT} = \frac{(c^2k - (1 - c)n^{-1}(x_t^d)^2)}{f_0} \text{ si } x_t \text{ semble contenir une tendance.}$$

La valeur de f_0 est une estimation du spectre du résidu pour la fréquence 0 et $c = -7$ si la série ne semble pas contenir de tendance et $c = -13,5$ si elle semble en contenir une.

Ces quatre statistiques sont à comparer à des valeurs critiques tabulées par Ng et Perron.

Exercice n° 5.3

 **fichier C5EX3**

Exemple d'application des tests DF, DFA, Phillips-Perron, KPSS, ERS et Ng-Perron au CAC40

On demande d'appliquer la stratégie des tests DF, DFA, PP, KPSS, ERS et Ng-Perron à l'indice CAC40 (indice représentatif de l'évolution des cours de bourse) sur 1 160 observations quotidiennes.

Solution

La Figure 5.8 illustre l'évolution de l'indice CAC40, et nous suggère un processus de type DS. Nous appliquons la stratégie de test de la Figure 5.6.

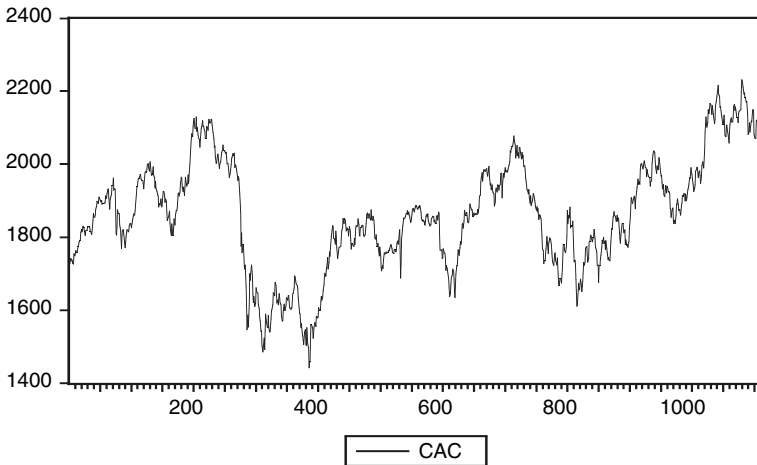


Figure 5.8 – Évolution de l'indice boursier CAC40

- Test Dickey-Fuller simple

La première étape consiste à l'estimation de modèle [3] :

DF Test Statistic	-2.053788	1 % Critical Value*	-3.9710	
		5 % Critical Value	-3.4161	
		10 % Critical Value	-3.1300	
Dependent Variable : D(CAC)				
Sample (adjusted) : 2 1160				
Included observations : 1159				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
CAC(-1)	-0.008447	0.004113	-2.053788	0.0402
C	14.93911	7.437617	2.008589	0.0448
@TREND(1)	0.002103	0.001896	1.108866	0.2677
R-squared	0.003870	Sum squared resid	491818.4	

La valeur empirique $t_{\hat{\phi}_1} = -2,05 >$ aux trois valeurs critiques, on accepte l'hypothèse H_0 .

Nous effectuons maintenant le test de l'hypothèse $H_0^3 : (c; b; \phi_1) = (c; 0; 1)$.

On calcule la statistique F_3 :

$$F_3 = \frac{(SCR_c^3 - SCR_3)/2}{SCR_3/(n-3)}$$

avec : $SCR_c^3 = \sum_t (x_t - x_{t-1} - \hat{c})^2 = 493729,2$ somme des carrés des résidus du modèle

le [3] contraint sous l'hypothèse H_0^3 , soit :

Dependent Variable : DCAC				
Included observations : 1159				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.417204	0.606525	0.687861	0.4917
R-squared	0.000000	Sum squared resid		493729.2

Or $SCR_3 = 491818,4$ (somme des carrés des résidus du modèle [3])

$$D'où : F_3 = \frac{(493729,2 - 491818,4)/2}{491818,4/(1159 - 3)} = 2,24$$

Nous sommes amenés à accepter les hypothèses H_0 et H_0^3 ($F_3 = 2,24 < \Phi_3 \in [6,25; 6,30]$). Nous allons sur la branche de droite de la figure 5.6 et effectuons le test de l'hypothèse H_0^2 , soit le calcul de la statistique

$$F_2 = \frac{(SCR_c - SCR_3)/3}{SCR_3/(n - 3)} \text{ avec :}$$

SCR_c = somme des carrés des résidus du modèle [2] contraint ($c = 0, b = 0, \phi_1 = 1$)

$$\text{soit } \sum_{t=1}^{t=1159} (CAC_t - CAC_{t-1})^2 = 493930,9.$$

D'où : $F_2 = \frac{(493930,9 - 491818,4)/3}{491818,4/(1159 - 3)} = 1,65$. Nous sommes amenés à accepter l'hypothèse H_0^2 ($F_2 = 1,65 < \Phi_2 \in [4,68; 6,09]$), nous estimons donc maintenant le modèle [2] et effectuons le test DF.

DF Test Statistic	-1.805839	1 % Critical Value*	-3.4388	
		5 % Critical Value	-2.8645	
		10 % Critical Value	-2.5683	
Dependent Variable : D(CAC)				
Included observations : 1159				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
CAC(-1)	-0.007093	0.003928	-1.805839	0.0712
C	13.63563	7.344864	1.856485	0.0636
R-squared	0.002811	Sum squared resid		492341.5

La valeur empirique $t_{\phi_1^*} - 1,80 >$ aux trois valeurs critiques, on accepte l'hypothèse H_0 .

Nous effectuons maintenant le test de l'hypothèse H_0^1 , soit le calcul de

$$F_1 = \frac{(SCR_c - SCR_2)/2}{SCR_2/(n - 2)} \text{ avec :}$$

$SCR_c = 493930,9$ (somme des carrés des résidus du modèle [2] contraint sous l'hypothèse H_0^1) et $SCR_2 = 492341,5$ (somme des carrés des résidus du modèle [2]).

$$F_1 = \frac{(493930,9 - 492341,5)/2}{492341,5/(1159 - 2)} = 1,87.$$

Nous sommes amenés à accepter l'hypothèse H_0^1 ($F_1 = 1,87 < \Phi_1 \in [3,38; 3,71]$),

En une dernière étape, nous testons la nullité de la moyenne du CAC40 :

Test of Hypothesis : Mean = 0.000000		
Sample Mean = 1863.760		
Sample Std. Dev. = 154.5946		
<u>Method</u>	<u>Value</u>	<u>Probability</u>
t-statistic	410.6054	0.0000

La moyenne est – largement – significativement différente de 0. Le processus est une marche au hasard sans dérive.

- Test Dickey-Fuller augmenté

Nous appliquons la même stratégie dans le cadre du test DFA.

Soit le modèle [6] à 4 variables décalées, $\rho = 5$ (les décalages 6 et 5 sont non significatifs) :

ADF Test Statistic	-2.278322	1 % Critical Value*	-3.9711	
		5 % Critical Value	-3.4161	
		10 % Critical Value	-3.1300	
Dependent Variable : D(CAC)				
Sample (adjusted) : 6 1160				
Included observations : 1155				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
CAC(-1)	-0.009478	0.004160	-2.278322	0.0229
D(CAC(-1))	0.069846	0.029480	2.369234	0.0180
D(CAC(-2))	-0.005537	0.029536	-0.187468	0.8513
D(CAC(-3))	-0.036930	0.029539	-1.250236	0.2115
D(CAC(-4))	0.069284	0.029507	2.348089	0.0190
C	16.77092	7.525601	2.228516	0.0260
@TREND(1)	0.002180	0.001901	1.146929	0.2516
R-squared	0.014213	Sum squared resid	486574.6	

La valeur empirique $t_{\phi_1}^z = -2,27 >$ aux trois valeurs critiques, on accepte l'hypothèse H_0 .

Nous effectuons maintenant le test de l'hypothèse H_0^6 , soit le calcul de la statistique :

$$F_6 = \frac{(SCR_D - SCR_6)/2}{SCR_6/(n - (p - 1) - 3)}$$

SCR_D = somme des carrés des résidus du modèle estimé :

$$\Delta x_t = \sum_{j=2}^5 a_j \Delta_j \Delta x_{t-j+1} + c + a_t, \text{ soit :}$$

Dependent Variable : DCAC				
Included observations : 1155				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.381017	0.607099	0.627603	0.5304
D(CAC(-1))	0.064750	0.029428	2.200260	0.0280
D(CAC(-2))	-0.010596	0.029488	-0.359321	0.7194
D(CAC(-3))	-0.041931	0.029492	-1.421760	0.1554
D(CAC(-4))	0.064122	0.029454	2.177046	0.0297
R-squared	0.009541	Sum squared resid	488881.0	

$$D' \text{ où } : F_6 = \frac{(488881 - 486574,6)/2}{486574,6/(1155 - 4 - 3)} = 2,72 < \Phi_3 \in [6,25; 6,30]$$

Nous sommes amenés à accepter les hypothèses H_0 et H_0^6 donc nous effectuons le test de l'hypothèse H_0^5 , on calcule la statistique F_5 :

$$F_5 = \frac{(SCR_C - SCR_6)/3}{SCR_6/(n - (p - 1) - 3)}$$

avec SCR_C = somme des carrés des résidus du modèle contraint :

$$\Delta x_t = \sum_{j=2}^5 a_j \Delta x_{t-j+1} + a_t, \text{ soit } SCR_C = 489048,4.$$

Dependent Variable : DCAC				
Included observations : 1155				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(CAC(-1))	0.065111	0.029415	2.213562	0.0271
D(CAC(-2))	-0.010298	0.029476	-0.349365	0.7269
D(CAC(-3))	-0.041620	0.029480	-1.411789	0.1583
D(CAC(-4))	0.064475	0.029441	2.190018	0.0287
R-squared	0.009202	Sum squared resid	489048.4	

Soit $F_5 = \frac{(489048,4 - 486574,6)/3}{486574,6/(1155 - 4 - 3)} = 1,94 < \Phi_2 \in [4,68; 6,09]$, l'hypothèse H_0^5 est donc acceptée.

Nous estimons donc maintenant le modèle [5].

La valeur empirique $t_{\phi_1^*} - 2,03 >$ aux trois valeurs critiques, on accepte l'hypothèse H_0 .

Nous effectuons maintenant le test de l'hypothèse H_0^4 , soit le calcul de :

$$F_4 = \frac{(SCR_C - SCR_5)/2}{SCR_5/(n - (p - 1) - 2)} = \frac{(489048,4 - 487132,1)/2}{487132,1/(1155 - 4 - 2)}$$

$$= 2,26 < \Phi_1 \in [3,38; 3,71]$$

ADF Test Statistic	-2.031024	1 % Critical Value*	-3.4388	
		5 % Critical Value	-2.8645	
		10 % Critical Value	-2.5683	
Dependent Variable : D(CAC)				
Included observations : 1155				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
CAC(-1)	-0.008079	0.003978	-2.031024	0.0425
D(CAC(-1))	0.069552	0.029483	2.359024	0.0185
D(CAC(-2))	-0.005917	0.029538	-0.200314	0.8413
D(CAC(-3))	-0.037283	0.029541	-1.262085	0.2072
D(CAC(-4))	0.068936	0.029509	2.336098	0.0197
C	15.43264	7.435612	2.075503	0.0382
R-squared	0.013084	Sum squared resid	487132.1	

$SCR_C = 489048,4$ (somme des carrés des résidus du modèle estimé sous H_0^4)

$SCR_S = 487132,1$ (somme des carrés des résidus du modèle [5])

Nous acceptons l'hypothèse H_0^4 .

Il est inutile de re-procéder au test de la moyenne, les conclusions sont donc les mêmes : le processus CAC40 est une marche au hasard sans dérive.

Le programme EvIEWS (📁 **fichier C5EX3.PRG**) permettant de calculer les F correspondants aux différents modèles peut être téléchargé du site Web.

• Test de Phillips-Perron

Nous procédons aux tests PP avec un nombre de retards $l = 6$ dans le cadre de trois modèles.

Dans les trois cas, la valeur empirique du t est supérieure aux valeurs critiques ; nous acceptons l'hypothèse H_0 de racine unitaire.

Hypothèse : CAC possède une racine unitaire		
Troncature $l = 6$		
Test de Phillips - Perron	t Statistique ajusté	Probabilité critique
Modèle [1] : modèle sans constante	0,50	0,82
Modèle [2] : modèle avec constante seule	-1,94	0,31
Modèle [3] : modèle avec tendance et constante	-2,20	0,49

Les probabilités critiques sont toutes supérieures à 0,05, nous ne rejetons pas l'hypothèse ; le processus CAC40 possède une racine unitaire.

• Test de KPSS

Hypothèse : CAC ne possède pas une racine unitaire		
Troncature $l = 6$		
Test de Kwiatkowski-Phillips-Schmidt-Shin LM	Statistique	Valeur critique à 5 %
Modèle [2]	2,58	0,46
Modèle [3]	1,31	0,14

La statistique LM est supérieure à la valeur critique (pour un seuil de 5 %) pour les deux spécifications, nous rejetons l'hypothèse H_0 , le processus CAC40 possède donc une racine unitaire.

• Test ERS

Les résultats du test ERS sont les suivants :

Hypothèse H_0 : CAC possède une racine unitaire				
Retard $k = 0$ (critère d'information de Schwarz)				
Elliott-Rothenberg-Stock	P-Statistique	Valeur critique à 10 %	Valeur critique à 5 %	Valeur critique à 1 %
Modèle [2]	7,34	4,48	3,26	1,99
Modèle [3]	9,69	6,89	5,69	3,96

Les statistiques P sont supérieures aux valeurs critiques (aux seuils de 5 % et 10 %) pour les deux spécifications, nous acceptons l'hypothèse H_0 , le processus CAC40 possède donc une racine unitaire.

• Test Ng-Perron

Les résultats du test Ng-Perron sont les suivants :

Hypothèse H_0 : CAC possède une racine unitaire					
Modèle [2] avec constante sans tendance					
Retard $k = 0$ (critère d'information de Schwarz)					
		MZa	MZt	MSB	MPT
Test Ng-Perron		-3,35	-0,95	0,28	7,17
Valeurs critiques	1 %	-13,8	-2,58	0,174	1,78
	5 %	-8,1	-1,98	0,233	3,17
	10 %	-5,7	-1,62	0,275	4,45

Hypothèse H_0 : CAC possède une racine unitaire					
Modèle [3] avec constante et tendance					
Retard $k = 0$ (critère d'information de Schwarz)					
		MZa	MZt	MSB	MPT
Test Ng-Perron		-9,79	-2,13	0,22	9,70
Valeurs critiques	1 %	-23,8	-3,42	0,143	4,03
	5 %	-17,3	-2,91	0,168	5,48
	10 %	-14,2	-2,62	0,185	6,67

Les statistiques sont supérieures aux valeurs critiques (aux seuils de 5 % et 10 %) pour les deux spécifications, nous acceptons l'hypothèse H_0 , le processus CAC40 possède donc une racine unitaire.

Tous les résultats sont convergents, nous pouvons donc conclure que le processus CAC40 n'est pas stationnaire.

S'agit-il d'un processus de marche au hasard sans dérive ? Nous allons le vérifier en calculant le corrélogramme de la série filtrée par les différences premières ceci afin de stationnariser le processus : $DCAC_t = CAC_t - CAC_{t-1}$.

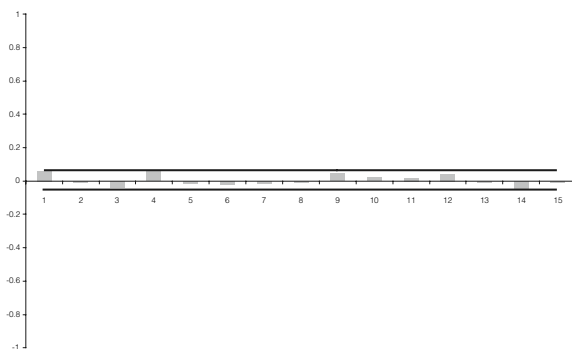


Figure 5.9 – Corrélogramme du CAC40 en différences premières

En première analyse, nous pouvons conclure que le processus CAC40 est une marche au hasard sans dérive car tous les termes de la fonction d'autocorrélation des différences premières appartiennent à l'intervalle de confiance. Cependant même si la série filtrée par les différences premières ne laisse pas apparaître des autocorrélations significativement différentes de 0 (cf. figure 5.9) nous pouvons vérifier la présence d'une forte hétéroscédasticité (cf. chapitre 8) caractéristique des processus à mémoire longue, ce qui rejette donc cette hypothèse de marche au hasard.

Exercice n° 5.4

fichier C5EX4

Exemple d'application des tests DF, DFA et Phillips-Perron et KPSS au cours en dollars du baril de pétrole brut

On demande d'appliquer la stratégie des tests DF, DFA et PP au prix du baril de pétrole brut connu sur 2087 jours.

Solution

La figure 5.10 illustre l'évolution du cours du prix du pétrole brut (PPB).

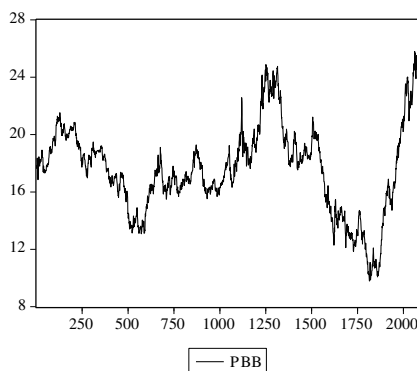


Figure 5.10 – Évolution du cours du pétrole brut

- Test de Dickey-Fuller simple ou augmenté ?

Avant d'appliquer la stratégie de test de la Figure 5.6, nous déterminons si nous recourons au test de Dickey-Fuller simple ou augmenté à partir du critère de Akaike. Pour cela nous partons d'un modèle avec constante et 6 décalages (donc 6 jours) dans le cadre du test DFA et nous diminuons le nombre de décalages jusqu'à 0. Nous retenons le décalage qui minimise le critère d'information de Akaike. Si le décalage 0 est retenu, nous utilisons alors le test DF simple.

Le résultat fournit par Eviews 4 est de 2 retards :

Null Hypothesis : PPB has a unit root
 Exogenous : Constant
 Lag Length : 2 (Automatic based on AIC, MAXLAG = 6)

Nous pouvons appliquer alors la stratégie de test et calculer l'ensemble des statistiques :

<i>Tests d'hypothèses jointes (Dickey-Fuller Augmenté)</i>			
	<i>Calculé</i>	<i>Lu</i>	<i>Conclusion</i>
Modèle 6 :			
H_0	$t^* = -1,33$	-3,41	Acceptée
Modèle 5 :			
H_0	$t^* = -1,43$	-2,86	Acceptée
<i>Tests d'hypothèses jointes</i>			
H_0^6	$F_6 = 1,39$	6,25	Acceptée
H_0^5	$F_5 = 1,03$	4,68	Acceptée
H_0^4	$F_4 = 1,19$	4,59	Acceptée

Nous retenons un processus de type DS, PPB est $I(1)$.

• Test de Phillips-Perron

Nous procédons aux tests PP avec un nombre de retards $l = 6$ dans le cadre de trois modèles. Dans les trois cas, la valeur empirique du t est supérieure aux valeurs critiques ; nous acceptons l'hypothèse H_0 de racine unitaire.

Hypothèse H_0 : PPB possède une racine unitaire		
Troncature $l = 6$		
Test de Phillips-Perron	t Statistique ajusté	Probabilité critique
Modèle [1] : modèle sans constante	0,327	0,78
Modèle [2] : modèle avec constante seule	-1,37	0,59
Modèle [3] : modèle avec tendance et constante	-1,25	0,89

Les probabilités critiques sont toutes supérieures à 0,05, nous ne rejetons pas l'hypothèse H_0 ; le processus PPB possède une racine unitaire.

• Test de KPSS

Enfin, nous procédons aux tests KPSS.

Hypothèse H_0 : PPB ne possède pas une racine unitaire		
Troncature $l = 6$		
Test de Kwiatkowski-Phillips-Schmidt-Shin LM	Statistique	Valeur critique à 5 %
Modèle [2]	1,51	0,46
Modèle [3]	1,21	0,14

La statistique LM est supérieure à la valeur critique (pour un seuil de 5 %) pour les deux spécifications, nous rejetons l'hypothèse H_0 , le processus PPB possède donc une racine unitaire.

Tous les résultats sont convergents, nous pouvons donc conclure que le processus PPB n'est pas stationnaire.

S'agit-il d'un processus de marche au hasard sans dérive ? Nous allons le vérifier en calculant la fonction d'autocorrélation et la statistique de Ljung-Box sur la série filtrée par les différences premières ceci afin de stationnariser le processus : $DPPB_t = PPB_t - PPB_{t-1}$.

Nous pouvons conclure à la lecture du Tableau 5.2 que les probabilités critiques de la Q -Stat sont toutes nulles, il ne s'agit donc pas d'une marche au hasard. Le processus PPB est un processus non stationnaire, mais n'est pas une marche au hasard sans dérive¹.

Tableau 5.2 – Fonction d'autocorrélation du PPB en différences premières

Décalage	AC	Q – Stat	Prob
1	0,163	55,372	0,000
2	– 0,072	66,197	0,000
3	– 0,042	69,890	0,000
4	– 0,036	72,567	0,000
5	– 0,024	73,777	0,000
6	– 0,029	75,549	0,000
7	– 0,003	75,568	0,000
8	0,062	83,713	0,000
9	0,047	88,335	0,000
10	– 0,007	88,452	0,000
11	– 0,042	92,140	0,000
12	– 0,040	95,461	0,000
13	– 0,010	95,681	0,000
14	0,001	95,683	0,000
15	0,033	97,951	0,000

III. Tests de racines unitaires saisonnières

La première application de ces tests dans le cas saisonnier a été réalisée par Hasza et Fuller (1981) puis par Dickey Hasza Fuller (1984). Ils s'adressent à des chroniques pour lesquelles la saisonnalité est de période trimestrielle et portent sur l'ensemble des fréquences correspondant aux racines unitaires. Ils ont été complétés par celui de Hylleberg, Engle, Granger et Yoo (1990), noté HEGY qui concerne la fréquence de chaque racine unitaire.

1. Pour compléter l'étude des tests de racine unitaire non saisonnière, le lecteur peut se référer à Darné, Terraza, 2002.

Franses (1990) présente un test pour les chroniques possédant une saisonnalité mensuelle, Beaulieu et Miron (1993) fournissent une nouvelle version de ce test¹.

A. Les modèles de base

D'une manière générale, une série temporelle peut être représentée par l'un des trois modèles suivants :

- modèle saisonnier purement déterministe : $x_t = \alpha_0 + \beta_t + \sum_{j=1}^{s-1} \alpha_j D_{jt} + a_t$
où $a_t \rightarrow i.i.d. (0; \sigma_a^2)$, s est la période de la saisonnalité et D_{jt} une variable muette relative à la saisonnalité ; ce modèle peut s'étendre à des tendances déterministes polynomiales de degré supérieur à un ;
- modèle saisonnier aléatoire stationnaire : $\varphi(B) x_t = \mu_t + a_t$ où $\varphi(B)$ est un polynôme d'opérateurs dont les racines sont à l'extérieur du cercle unité du plan complexe et μ_t est un terme quelconque entre une constante, une tendance et des variables muettes relatives à la saisonnalité ;
- modèle saisonnier intégré : $\varphi(B) x_t = \mu_t + a_t$ où $\varphi(B)$ est un polynôme d'opérateurs qui possède au moins une racine unitaire (d'où son nom d'intégré).

L'utilisation du filtre aux différences saisonnières proposé par Box et Jenkins sous la forme : $(1 - B^s) = \Delta S(B)$ permet de déterminer les racines unitaires saisonnières contenues dans le filtre moyenne mobile $S(B)$.

Les tests de HEGY et de Franses utilisent cette décomposition pour les polynômes $(1 - B^4)$ et $(1 - B^{12})$, on obtient alors respectivement 4 et 12 racines unitaires, par exemple :

$$(1 - B^4) = (1 - B)(1 + B)(1 - iB)(1 + iB)$$

Tous les termes autres que $(1 - B)$ correspondent à des racines unitaires saisonnières, à savoir (-1) pour la période 2 (fréquence π) et $\pm i$ pour la période 4 (fréquence $\pm\pi/2$).

Pour tester l'hypothèse qu'un polynôme $\varphi(B)$ possède des racines sur le cercle unité du plan complexe on l'écrit en utilisant la proposition de Lagrange².

1. Pour une présentation de ces tests pour des saisonnalités de longueur impaire détaillée se référer à Darné, Litago et Terraza, 2002.

2. Tout polynôme (infini ou rationnel) à valeurs finies distinctes et non nulles pour les différents points $(\theta_1, \dots, \theta_p)$, éventuellement complexes, distincts de zéro, peut s'exprimer selon les termes de polynômes élémentaires et d'un reste selon la forme : $\varphi(B) = \sum_{k=1}^p \lambda_k \Delta(B) / \delta_k(B) + \Delta(B) \varphi^{**}(B)$ où : λ_k est un ensemble de constantes, $\varphi^{**}(B)$ un polynôme d'opérateur en B

B. Le test de Hylleberg, Engle, Granger et Yoo (HEGY)

En utilisant l'expansion de $\varphi(B)$ donnée par la proposition de Lagrange avec $\theta_1 = 1$, $\theta_2 = -1$, $\theta_3 = i$, $\theta_4 = -i$; $\delta_1 = 1 - B$, $\delta_2 = 1 + B$, $\delta_3 = 1 - iB$, $\delta_4 = 1 + iB$ et $\Delta(B) = (1 - B^4)$. On obtient alors :

$$\begin{aligned}\varphi(B) &= \lambda_1 B(1 + B)(1 + B^2) + \lambda_2 (-B)(1 - B)(1 + B^2) \\ &\quad + \lambda_3 (-iB)(1 - B)(1 + B)(1 - iB) \\ &\quad + \lambda_4 (iB)(1 - B)(1 + B)(1 + iB) + (1 - B^4)\varphi^*(B)\end{aligned}$$

On simplifie cette expression en appelant $H(B)$ la somme du deuxième terme :

$\varphi(B) = H(B) + (1 - B^4)\varphi^*(B)$ et en utilisant le processus $\varphi(B) x_t = \mu_t + a_t$, on a :

$$\varphi^*(B)(1 - B^4)x_t = -H(B)x_t + \mu_t + a_t$$

C'est cette expression qui est la base du test mais on ne peut pas estimer les λ_k puisqu'ils sont complexes, on regroupe alors les parties réelles et complexes de l'expression et on pose :

$$\pi_1 = -\lambda_1 ; \quad \pi_2 = -\lambda_2 ; \quad 2\lambda_3 = -\pi_3 + i\pi_4 ; \quad 2\lambda_4 = -\pi_3 - i\pi_4.$$

Ce test, de saisonnalité trimestrielle, est fondé sur le fait que $\varphi(B)$ a une racine unitaire en θ_k si et seulement si $\lambda_k = 0$. Pour tester cette hypothèse, on estime par les MCO une équation de régression appelée « régression auxiliaire » :

$$\varphi^*(B) x_{4t} = \pi_1 x_{1,t-1} + \pi_2 x_{2,t-1} + \pi_3 x_{3,t-2} + \pi_4 x_{3,t-1} + \mu_t + a_t$$

avec :

$$\begin{aligned}x_{1,t} &= (1 + B + B^2 + B^3)x_t ; \quad x_{2,t} = -(1 - B + B^2 - B^3)x_t ; \\ x_{3,t} &= -(1 - B^2)x_t ; \quad x_{4,t} = (1 - B^4)x_t\end{aligned}$$

Dans cette équation $\varphi^*(B)$ est inconnu. Plusieurs solutions ont été proposées, en particulier, celle qui consiste à augmenter le retard du polynôme jusqu'à ce que le résidu soit un bruit blanc ou que l'un des critères de comparaison des

(infini ou rationnel), $\delta_k(B) = 1 - \frac{B}{\theta_k}$, $\Delta(B) = \prod_{k=1}^p \delta_k(B)$ avec p l'ordre de la périodicité.

En ajoutant et en soustrayant $\Delta(B) \sum_{k=1}^p \lambda_k$ dans la proposition, on obtient la forme alternative :

$$\begin{aligned}\varphi(B) &= \sum_{k=1}^p \lambda_k \Delta(B)(1 - \delta_k(B))/\delta_k(B) + \Delta(B)\varphi^{**}(B) \quad \text{avec} \quad \varphi^*(B) = \varphi^{**}(B) + \sum_{k=1}^p \lambda_k \quad \text{et} \\ \varphi(0) &= \varphi^*(0).\end{aligned}$$

modèles soit minimum (AIC ou BIC, par exemple, cf. chapitre 7). Le plus souvent ces tests sont réalisés avec $\varphi^*(B) = 1$.

L'hypothèse nulle du test porte sur la significativité des paramètres de la régression auxiliaire (lesquels sont équivalents aux λ_k), Soit :

$$H_0 : \lambda_k = 0 \text{ contre } H_1 : \lambda_k \neq 0.$$

Supposons, par exemple, que $\lambda_1 = \lambda_2 = 0$ et que $\varphi^*(B) = 1$, en substituant ces valeurs dans : $\varphi^*(B)(1 - B^4)x_t = -H(B)x_t + \mu_t + a_t$, on a :

$$(1 - B^4)x_t = \lambda_3(-iB)(1 - B)(1 + B)(1 - iB) \\ + \lambda_4(iB)(1 - B)(1 + B)(1 + iB) + \mu_t + a_t$$

En introduisant la décomposition de $(1 - B^4)$ on obtient :

$$(1 - B)(1 + B)\left\{(1 - iB)(1 + iB) - \lambda_3(-iB)(1 - iB) \right. \\ \left. - \lambda_4(iB)(1 + B)(1 + iB)\right\}x_t = \mu_t + a_t$$

Cette équation possède deux racines unitaires $B = 1$ et $B = -1$. Ce sont celles qui manquent aux coefficients λ_1 et λ_2 . Le filtre pour rendre stationnaire ce processus est dans ce cas égal à :

$$(1 - B^2) = (1 - B)(1 + B)$$

C. Le test de Franses

Le principe du test de Franses est identique à celui de HEGY. Les formules sont cependant plus complexes du fait à présent des douze racines unitaires. L'expansion de $\varphi(B)$ est toujours réalisée à partir de la proposition de Lagrange. La procédure de Franses pour rendre possible l'estimation des paramètres de $\varphi(B)$ est la suivante :

- Séparation des parties réelles et imaginaires de chaque terme de $\varphi(B)$ puis regroupement des termes communs. On définit alors la série des régresseurs $\varphi_i(B)$:

$$\varphi_1(B) = \left(1 + \sum_{j=1}^{11} B_j\right) ;$$

$$\varphi_2(B) = (1 + B^2)(1 - B)(1 + B^4 + B^8) ;$$

$$\varphi_3(B) = (1 - B^2)(1 + B^4 + B^8)$$

$$\varphi_4(B) = (1 - B^4)(1 - \sqrt{3}B + B^2)(1 + B^2 + B^4)$$

$$\varphi_5(B) = (1 - B^4)(1 + \sqrt{3}B + B^2)(1 + B^2 + B^4) ;$$

$$\varphi_6(B) = (1 - B^4)(1 - B^2 + B^4)(1 - B + B^2) ;$$

$$\varphi_7(B) = (1 - B^4)(1 - B^2 + B^4)(1 + B + B^2)$$

$$\varphi_8(B) = (1 - B^{12})$$

- Élimination des termes complexes de $\varphi(B)$ par l'utilisation de paramètres π_i définis de la manière suivante :

$$\lambda_1 = -\pi_1 ;$$

$$\lambda_2 = -\pi_2 ;$$

$$\lambda_3 = -i(\pi_3 + \pi_4)/2 ;$$

$$\lambda_4 = -(-i\pi_3 + \pi_4)/2 ;$$

$$\lambda_5 = i\pi_5 - ((1 + i\sqrt{3})/2)\pi_6 ;$$

$$\lambda_6 = -i\pi_5 - ((1 + i\sqrt{3})/2)\pi_6 ;$$

$$\lambda_7 = -i\pi_7 - ((1 + i\sqrt{3})/2)\pi_8 ;$$

$$\lambda_8 = i\pi_7 - ((1 - i\sqrt{3})/2)\pi_8 ;$$

$$\lambda_9 = +i\sqrt{3}\pi_9/3 - ((1 + (1/3)i\sqrt{3})/2)\pi_{10} ;$$

$$\lambda_{10} = -i\sqrt{3}\pi_9/3 - ((1 - (1/3)i\sqrt{3})/2)\pi_{10} ;$$

$$\lambda_{11} = -i\sqrt{3}\pi_{11}/3 - ((1 + (1/3)i\sqrt{3})/2)\pi_{11} ;$$

$$\lambda_{12} = i\sqrt{3}\pi_{11}/3 - ((1 - (1/3)i\sqrt{3})/2)\pi_{12} .$$

- Substitution des $\varphi_i(B)$ et des π_i dans $\varphi(B)$ qui est alors une expression linéaire sans terme complexe.

Si on considère, par exemple, que le processus est saisonnier intégré et si on pose :

$x_{it} = \varphi_i(B)x_t$ pour $i = \{1, 8\}$ et $x_{it} = -\varphi_i(B)x_t$ pour $i = \{2, \dots, 7\}$, on obtient la régression auxiliaire des tests :

$$\begin{aligned} \varphi^*(B)x_{8t} = & \pi_1 x_{1,t-1} + \pi_2 x_{2,t-1} + \pi_3 x_{3,t-1} + \pi_4 x_{3,t-2} + \pi_5 x_{4,t-1} + \pi_6 x_{4,t-2} \\ & + \pi_7 x_{5,t-1} + \pi_8 x_{5,t-2} + \pi_9 x_{6,t-1} + \pi_{10} x_{6,t-2} + \pi_{11} x_{7,t-1} + \pi_{12} x_{7,t-2} + \mu_t + a_t \end{aligned}$$

On estime par les MCO cette équation, avec le plus souvent, $\varphi^*(B) = 1$ ou bien $\varphi^*(B)$ avec L retards choisis comme indiqué précédemment. L'hypothèse H_0 des tests est : $\varphi(B) = 0$ contre l'hypothèse $H_1 : \varphi(B) \neq 0$. Ces tests contrôlent la significativité des douze paramètres π_k de la régression auxiliaire (lesquels sont équivalents aux λ_k). Puis on réalise :

– des tests individuels $H_0 : \pi_k = 0$ contre $H_1 : \pi_k \neq 0$ pour $k = 1, \dots, 12$;

– des tests pour les racines complexes conjuguées :

$H_0 : \pi_i = \pi_j = 0$ contre $H_1 : \pi_i \neq \pi_j \neq 0$ pour $j = 3, 5, 7, 9, 11$ et $i = 4, 6, 8, 10, 12$. (On les note $\pi_3 \cap \pi_4, \pi_5 \cap \pi_6$, etc.) ;

– un test pour l'ensemble des coefficients :

$H_0 : \pi_j = \dots \pi_j = 0$ contre $H_1 : \pi_j \neq \dots \pi_j \neq 0$ pour $j = 3, 4, \dots, 12$ (on le note $\pi_3 \cap \dots \cap \pi_{12}$).

Les valeurs critiques de ces tests (t statistiques pour les tests individuels et F statistiques pour les tests conjoints) ont été tabulées par Franses.

Exercice n° 5.5

Application du test de Franses

On demande d'appliquer le test de Franses aux deux processus suivants simulés sur 1092 observations : $x_t = 0,8x_{t-12} + a_t$ (Processus stationnaire autorégressif saisonnier de période douze) et $x_t = x_{t-12} + a_t$ (Processus intégré autorégressif saisonnier de période douze) avec $a_t \rightarrow N(0; 1)$.

Solution

On applique à ces deux processus une recherche de racines unitaires selon la procédure de Franses. Le Tableau 5.3 et le Tableau 5.4 donnent les résultats détaillés des calculs pour les statistiques t et F ainsi que des critères d'évaluation et de comparaison (cf. chapitre 7) des différentes régressions auxiliaires.

Les régressions auxiliaires ont pour notations :

nc = pas de constante ; c = avec constante ; nd = pas de variable *dummy* ;

d = avec variables « dummy » ; nt = pas de tendance linéaire ; t = avec tendance linéaire.

(Ce sont les modèles de base présentés au paragraphe II.A.).

Pour le premier processus (Tableau 5.3) l'introduction de variables déterministes (par exemple c , d , t) dans le modèle purement aléatoire, n'améliore pas significativement la qualité globale de l'ajustement (\bar{R}^2) et n'entraîne pas une variation sensible de l'auto-corrélation des résidus (Q statistique) ainsi que des critères de comparaisons des modèles (AIC, SBIC).

Nous pouvons constater la présence des racines unitaires par les tests t aux seuils de 5 et 10 % sur π_3 , π_5 , π_9 , π_{11} , que l'on ne retrouve pas avec les tests F . En définitive, ce processus ne possède pas de racine unitaire, mais il est difficile de sélectionner le processus générateur de la chronique même si on peut constater que la statistique SBIC est la plus faible pour la régression attendue (nc, nd, nt).

Pour le deuxième processus (Tableau 5.4), l'introduction de variables déterministes entraîne une faible variation à la hausse des statistiques \bar{R}^2 ; elle est plus marquée pour les critères (Q statistique et SBIC). Les tests t et F indiquent que les douze racines sont unitaires pour les deux premières régressions et pour les régressions (c, nd, t) alors que les racines non unitaires apparaissent avec les régressions incorporant des variables dichotomiques saisonnières.

En combinant l'ensemble de ces résultats, on retient sans conteste le premier modèle (nc, nd, nt).

Tableau 5.3 – Tests de Franses de racines unitaires saisonnières du processus

$$x_t = 0,8x_{t-12} + a_t \text{ avec } \varphi^*(B) = 1$$

Statistique t	nc, nd, nt	c, nd, nt	c, d, nt	c, nd, t	c, d, t
π_1	-3,31	-3,36	-3,37	-3,45	-3,46
π_2	-3,78	-3,77	-3,92	-3,78	-3,92
π_3	-1,40	-1,40	-1,68	-1,40	-1,68
π_4	-3,60	-3,60	-4,09	-3,60	-4,09
π_5	-1,91	-1,91	-2,03	-1,92	-2,03
π_6	-4,01	-4,01	-4,30	-4,01	-4,30
π_7	4,74	4,74	4,76	4,73	4,75
π_8	-4,58	-4,57	-4,64	-4,57	-4,63
π_9	-1,86	-1,86	-2,00	-1,87	-2,01
π_{10}	-4,33	-4,33	-4,54	-4,33	-4,55
π_{11}	1,23	1,23	1,30	1,23	1,30
π_{12}	-3,75	-3,75	-4,28	-3,75	-4,28
Statistique F					
$\pi_3 \cap \pi_4$	7,51	7,51	9,83	7,52	9,84
$\pi_5 \cap \pi_6$	13,24	13,23	15,49	13,22	15,47
$\pi_7 \cap \pi_8$	11,71	11,69	11,88	11,68	11,86
$\pi_9 \cap \pi_{10}$	9,46	9,45	10,38	9,46	10,40
$\pi_{11} \cap \pi_{12}$	7,31	7,31	9,66	7,32	9,67
$\pi_3 \cap \dots \cap \pi_{12}$	10,09	10,08	11,79	10,08	11,79
\underline{Q}_2 Stat.	14,3	15,5	13,5	15,3	13,4
\underline{R}^2	0,096	0,09	0,10	0,09	0,10
AIC	2,85	2,86	2,86	2,31	2,86
SBIC	0,076	0,08	0,13	0,17	0,14

Tableau 5.4 – Tests de Franses de racines unitaires saisonnières du processus

$$x_t = x_{t-12} + a_t \text{ avec } \varphi^*(B) = 1$$

Statistique t	nc, nd, nt	c, nd, nt	c, d, nt	c, nd, t	c, d, t
π_1	-1,78	-1,98	-2,00	-2,12	-2,14
π_2	1,42	1,42	-0,03	1,42	-0,03
π_3	-1,04	-1,04	-0,13	-1,04	-0,13
π_4	0,27	0,27	-1,49	0,27	-1,49
π_5	-0,52	-0,52	-1,52	-0,53	-1,52
π_6	-0,90	-0,90	-1,91	-0,90	-1,92
π_7	0,46	0,46	1,09	0,47	1,10
π_8	-1,08	-1,09	-2,02	-1,10	-2,03
π_9	-1,37	-1,36	-2,51	-1,36	-2,51
π_{10}	-1,38	-1,38	-3,32	-1,38	-3,32
π_{11}	0,57	0,57	0,79	0,57	0,79
π_{12}	-0,88	-0,88	-1,78	-0,88	-1,78
Statistique F					
$\pi_3 \cap \pi_4$	0,58	0,58	1,12	0,58	1,13
$\pi_5 \cap \pi_6$	0,54	0,53	1,88	0,54	1,88
$\pi_7 \cap \pi_8$	1,05	1,06	2,93	1,07	2,94
$\pi_9 \cap \pi_{10}$	1,26	1,26	6,03	1,26	6,03
$\pi_{11} \cap \pi_{12}$	0,40	0,40	1,60	0,40	1,60
$\pi_3 \cap \dots \cap \pi_{12}$	0,77	0,77	2,71	0,77	2,72
\underline{Q}_2 Stat.	28,7	29,1	30,3	29,2	30,4
\underline{R}^2	0,002	0,002	0,016	0,001	0,016
AIC	2,87	2,88	2,87	2,88	2,87
SBIC	0,097	0,102	0,14	0,11	0,15

IV. Les processus ARIMA

Les chroniques économiques sont rarement la réalisation de processus stationnaires. C'est la raison pour laquelle – dans l'algorithme de Box et Jenkins – l'étape de la stationnarisation de la chronique est la première. On recourt à trois grands types de transformations :

- si la chronique présente des mouvements de long terme de type non linéaire (logarithme, exponentiel ou puissance) ou des fluctuations importantes autour de ce mouvement, alors il est souvent utile d'utiliser la transformation de Box-Cox qui permet de retrouver des linéarités ou d'amoindrir les fluctuations importantes ou erratiques. Cette transformation s'écrit :

$$y_t = \frac{x_t^\lambda - 1}{\lambda} \quad \text{si } \lambda \neq 0 \\ = \ln x_t \quad \text{si } \lambda = 0$$

La difficulté de son utilisation réside dans le choix de λ ¹. On peut pour cela réaliser plusieurs transformations avec des λ différents et choisir la valeur de λ en fonction de l'aspect de la chronique transformée.

On peut aussi calculer le paramètre λ lors de l'étape d'estimation de l'algorithme de Box et Jenkins.

Les deux autres types de transformations s'opèrent après la précédente et résultent des conclusions issues des tests de racines unitaires non saisonniers et/ou saisonniers :

- si le processus générateur est de type TS, on estime par les moindres carrés ordinaires sa partie déterministe et l'algorithme se poursuit sur la chronique résiduelle.
- si le processus est de type non saisonnier et/ou saisonnier DS, alors on emploie les filtres aux différences pour le stationnariser.

Le recours à ces filtres permet de définir les processus ARMA intégrés, notés ARIMA.

A. Les processus ARIMA non saisonniers

Supposons une chronique x_t échantillon d'un processus DS d'ordre d . Ce processus est stationnarisé par l'utilisation du filtre aux différences d'ordre d (cf. chapitre 4).

$$x_t \xrightarrow{(1-B)^d} y_t = (1-B)^d x_t$$

1. On utilise souvent $\lambda = 0$ afin de rendre additif un schéma structurant multiplicatif.

Le processus x_t contient d racines réelles unité. Il est dit intégré d'ordre d et on l'écrit :

$$x_t \longrightarrow I(d) \text{ avec } d \text{ l'ordre d'intégration.}$$

Si le processus sortant y_t est un ARMA(p, q), alors le processus entrant x_t est un ARIMA(p, d, q). Un ARIMA(p, d, q) non saisonnier s'écrit :

$$\phi_p(B)(1 - B)^d x_t = \theta_q(B)a_t$$

$$(\text{ARIMA}(p, d, q) = \text{ARIMA}(p + d, q) \text{ non stationnaire}).$$

Dans la plupart des chroniques économiques, les tests de Dickey-Fuller indiquent la présence d'une seule racine unité et les chroniques sont dites intégrées d'ordre 1, $I(1)$.

Si la chronique de départ est la réalisation d'un processus stationnaire, elle ne contient pas de racine unité, on dit qu'elle est intégrée d'ordre 0 et elle est notée $I(0)$. En d'autres termes, un ARIMA($p, 0, q$) est un ARMA(p, q) stationnaire.

Parmi les processus ARIMA non saisonniers, on peut citer le processus ARIMA(0, 1, 0) : $(1 - B)x_t = a_t \iff x_t = x_{t-1} + a_t$ qui est le processus de promenade aléatoire « *Random Walk* » (processus DS le plus simple).

B. Les processus ARIMA purement saisonniers (modèles SARIMA)

Considérons une chronique x_t échantillon d'un processus DS d'ordre D de saisonnalité s . Ce processus est stationnarisé après passage dans un filtre saisonnier aux différences D -ièmes.

$$x_t \xrightarrow{(1-B^s)^D} y_t = (1 - B^s)^D x_t$$

Le processus x_t contient D racines réelles ou complexes unitaires. Il est dit intégré saisonnier d'ordre D et on l'écrit $x_t \longrightarrow I_s(D)$ avec D l'ordre d'intégration et s la période de la saisonnalité.

Supposons que le processus sortant du filtre y_t est un processus ARMA $_{s,s'}(P, Q)$ alors le processus entrant x_t est non stationnaire : ARIMA $_{s,s'}(P, D, Q)$, il s'écrit :

$$\phi_p(B^s)(1 - B^s)^D x_t = \theta_Q(B^{s'})a_t$$

On constate qu'un ARIMA $_{s,s'}(P, D, Q)$ est un ARMA $_{s,s'}(P + D, Q)$ non stationnaire et qu'un ARIMA $_{s,s'}(P, 0, Q)$ est un ARMA $_{s,s'}(P, Q)$ stationnaire.

C. Les processus ARIMA non saisonniers et saisonniers à la fois

En réunissant les écritures des modèles ARIMA non saisonniers et ARIMA saisonniers, on peut construire des modèles multiplicatifs ARIMA non saisonniers et saisonniers à la fois. C'est le modèle le plus complet que l'on trouve dans la présentation de Box et Jenkins.

Considérons une chronique x_t échantillon d'un processus à la fois DS d'ordre d et DS saisonniers d'ordre D et de période de saisonnalité s .

Ce processus est stationnarisé par le filtre combinant les différences d'ordre d et les différences saisonnières d'ordre D . Soit :

$$x_t \xrightarrow{(1-B)^d(1-B^s)^D} y_t = (1-B)^d(1-B^s)^D x_t$$

Le processus x_t contient $d + D$ racines unitaires réelles ou complexes (il est dit intégré d'ordre $d + D$), il s'écrit : $x_t \rightarrow I_s(d + D)$.

Si le processus sortant du filtre y_t est un $\text{ARMA}(p, q) \times \text{ARMA}_{s, s'}(P, Q)$ alors le processus entrant x_t est un $\text{ARIMA}(p, d, q) \times \text{ARIMA}_{s, s'}(P, D, Q)$ il s'écrit : $\phi_p(B)\phi_p(B^s)(1-B^s)^D(1-B)^d x_t = \theta_q(B)\theta_Q(B^s)a_t$.

On constate qu'un $\text{ARIMA}(p, d, q) \times \text{ARIMA}_{s, s'}(P, D, Q)$ est un $\text{ARMA}(p + d + Ps + Ds, q + Qs')$ non stationnaire et qu'un $\text{ARIMA}(p, 0, q) \times \text{ARIMA}_{s, s'}(P, 0, Q)$ est un $\text{ARMA}(p, q) \times \text{ARMA}_{s, s'}(P, Q)$ stationnaire.

Exemple d'écriture : $\text{ARIMA}(1, 1, 1) \times \text{ARIMA}_{12, 6}(1, 1, 1)$:

$$(1 - \phi_1 B)(1 - \phi_{12} B^{12})(1 - B)(1 - B^{12})x_t = (1 - \theta_1 B)(1 - \theta_6 B^6)a_t$$

Remarque : ce sont les processus ARIMA que Beveridge, Nelson et Harvey utilisent pour la décomposition « Trend – Cycle » des chroniques.

Exercice n° 5.6

Relation entre lissage exponentiel simple et modèle ARIMA

Démontrer que le LES est un processus ARIMA(0, 1, 1).

Solution

Par construction, le LES est une transformation de Koyck d'un modèle autorégressif (AR) géométrique infini (cf. chapitre 2) et s'écrit :

$$\hat{x}_t = \lambda x_{t-1} + (1 - \lambda)\hat{x}_{t-1} \text{ comme } e_t = x_t - \hat{x}_t \iff \hat{x}_t = x_t - e_t$$

$$\begin{aligned}
 x_t &= e_t + \widehat{x}_t \\
 &= e_t + \lambda x_{t-1} + (1 - \lambda)\widehat{x}_{t-1} \\
 &= e_t + \lambda x_{t-1} + (1 - \lambda)(x_{t-1} - e_{t-1})
 \end{aligned}$$

D'où :

$$x_t = x_{t-1} + e_t - (1 - \lambda)e_{t-1} \text{ soit avec } 1 - \lambda = \theta_1$$

$$x_t = x_{t-1} + e_t - \theta_1 e_{t-1}$$

Comme $0 < \lambda < 1$ alors $0 < \theta_1 < 1$, le LES est donc un processus DS sans dérive avec erreurs autocorrélées ou encore un ARIMA(0, 1, 1). On peut chercher la solution de ce processus DS sans dérive :

$$x_t = x_{t-1} + e_t - \theta_1 e_{t-1}$$

$$x_{t-1} = x_{t-2} + e_{t-1} - \theta_1 e_{t-2}$$

$$x_t = x_{t-2} + e_t - \theta_1 e_{t-1} + e_{t-1} - \theta_1 e_{t-2}$$

$$x_t = x_{t-2} + e_t + (1 - \theta_1)e_{t-1} - \theta_1 e_{t-2}$$

$$x_{t-2} = x_{t-3} + e_{t-2} - \theta_1 e_{t-3}$$

$$x_t = x_{t-3} + e_t + (1 - \theta_1)e_{t-1} + (1 - \theta_1)e_{t-2} - \theta_1 e_{t-3}$$

⋮

$$x_t = x_0 + e_t + (1 - \theta_1)(e_{t-1} + e_{t-2} + \dots + e_1) - \theta_1 e_0$$

$$x_t = x_0 + \sum_{j=0}^t \psi_j e_j$$

$$\text{avec } \psi_0 = -\theta_1 ; \psi_1 = \dots = \psi_{t-1} = (1 - \theta_1) ; \psi_t = 1$$

$$E[x_t] = x_0$$

$$V[x_t] = \sigma_{e_t}^2 + (t - 1)(1 - \theta_1)^2 \sigma_{e_t}^2 + \theta_1^2 \sigma_{e_t}^2 = [2\theta_1 + (1 - \theta_1)^2 t] \sigma_{e_t}^2$$

La variance dépend du temps, le LES est un processus non stationnaire aléatoire.

Remarque : on démontre d'une manière plus complexe que le LED est un ARIMA (0, 2, 2).

$$\text{Soit : } (1 - B)^2 x_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \iff (1 - 2B + B^2)x_t = \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

$$\text{ou encore : } x_t = 2x_{t-1} - x_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

Exercice n° 5.7

fichier C5EX7

Le jeu des quatre erreurs

Soit le prix « spot » d'un produit dérivé connu mensuellement sur 24 ans entre octobre de l'année 1 et avril de l'année 24. Un analyste, pas très doué en économétrie désire calculer une prévision pour ce prix en mai de l'année 24 et utilise donc le logiciel Eviews mais de manière désastreuse. Vous devez retrouver les quatre erreurs commises par cet analyste.

Étape 1 : Graphique de la série brute (cf. figure 5.11).

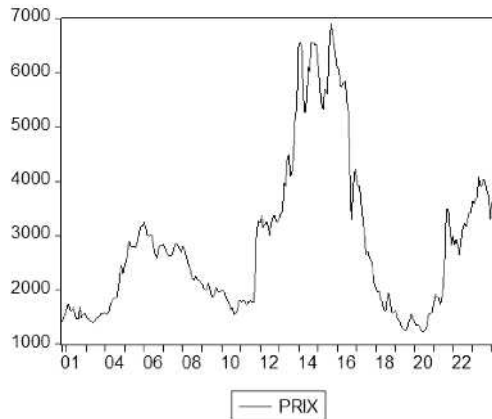


Figure 5.11 – Evolution du prix « spot »

Étape 2 : Analyse préalable, la désaisonnalisation.

La série étant mensuelle, il convient de la désaisonnaliser : soit la série CVS (cf. Tableau 5.5) selon un schéma multiplicatif (notée PRIXSA).

Tableau 5.5 – Coefficients saisonniers multiplicatifs

Coefficients saisonniers :	
1	1,038116
2	1,023395
3	1,007133
4	0,988306
5	0,976205
6	0,966773
7	0,954226
8	0,972208
9	1,007508
10	1,013079
11	1,032391
12	1,024993

Étape 3 : Tests de stationnarité.

De manière classique, notre analyste procède aux tests de racine unitaire de Dickey Fuller Augmenté ($p = 4$ décalages).

Modèle 4 :

Dependent Variable : D(PRIXSA)				
Included observations : 266 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
PRIXSA(-1)	- 0,000965	0,004080	- 0,236444	0,8133
D(PRIXSA(-1))	0,332172	0,061921	5,364427	0,0000
D(PRIXSA(-2))	- 0,119352	0,065357	- 1,826148	0,0690
D(PRIXSA(-3))	- 0,014306	0,065320	- 0,219021	0,8268
D(PRIXSA(-4))	- 0,049941	0,062137	- 0,803710	0,4223

Modèle 5 :

Dependent Variable : D(PRIXSA)				
Included observations : 266 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
PRIXSA(-1)	- 0,013831	0,009243	- 1,496283	0,1358
D(PRIXSA(-1))	0,335353	0,061790	5,427336	0,0000
D(PRIXSA(-2))	- 0,113231	0,065302	- 1,733959	0,0841
D(PRIXSA(-3))	- 0,009980	0,065204	- 0,153063	0,8785
D(PRIXSA(-4))	- 0,044048	0,062088	- 0,709445	0,4787
C	45,91663	29,62006	1,550187	0,1223

Modèle 6 :

Dependent Variable : D(PRIXSA)				
Included observations : 266 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
PRIXSA(-1)	- 0,014226	0,009562	- 1,487721	0,1380
D(PRIXSA(-1))	0,335597	0,061923	5,419591	0,0000
D(PRIXSA(-2))	- 0,112880	0,065458	- 1,724455	0,0858
D(PRIXSA(-3))	- 0,009700	0,065349	- 0,148433	0,8821
D(PRIXSA(-4))	- 0,043532	0,062282	- 0,698949	0,4852
C	43,04537	34,35288	1,253035	0,2113
@TREND(1973 :10)	0,029118	0,175499	0,165917	0,8684

Étape 4 : Stratégie de test.

Puis il applique une stratégie de test :

<i>Tests d'hypothèses jointes de Dickey-Fuller augmenté</i>	
Modèle 6 : $\Delta x_t = \rho x_{t-1} - \sum_{j=2}^p \Phi_j \Delta x_{t-j+1} + c + bt + a_t$	
$H_0^6 : (c; b; \rho) = (c; 0; 0)$	$F_6 = 1,129$
Modèle 5 : $\Delta x_t = \rho x_{t-1} - \sum_{j=2}^p \Phi_j \Delta x_{t-j+1} + c + a_t$	
$H_0^5 : (c; \rho) = (0; 0)$	$F_5 = 1,229$

Comme le coefficient de la tendance dans le modèle [6] a une probabilité critique de 0,8684, il n'est pas significativement différent de 0, il en déduit que le processus est DS sans dérive, processus de marche au hasard.

Étape 5 : Prédiction.

Comme le processus est processus DS sans dérive de marche au hasard, il calcule la prévision de la manière suivante :

PRIX pour mai de l'année 24 = (PRIXSA observé pour avril de l'année 24) \times coefficient saisonnier mai de l'année 24.

Solution

Première erreur : on ne désaisonnalise pas systématiquement une série mensuelle (les prix sont rarement saisonniers), visiblement le graphique n'indique aucune saisonnalité.

Deuxième erreur : le nombre de décalage dans les tests ADF est déterminé par les critères d'information de Akaïke et/ou Schwarz et non choisi de manière arbitraire.

Troisième erreur : Dans le modèle 6 comme l'hypothèse H_0 est acceptée, les probabilités critiques des coefficients (donc celui de la tendance) ne doivent pas être calculées à partir d'une loi de Student mais doivent être lues dans les tables de Dickey-Fuller. Donc les probabilités critiques indiquées par Eviews ne sont pas utilisables.

Quatrième erreur : Il s'agit bien d'un processus DS sans dérive mais il faut calculer le corrélogramme de la série en différences premières pour savoir s'il s'agit d'une marche au hasard.

La correction de cet exercice est effectuée au chapitre 8.

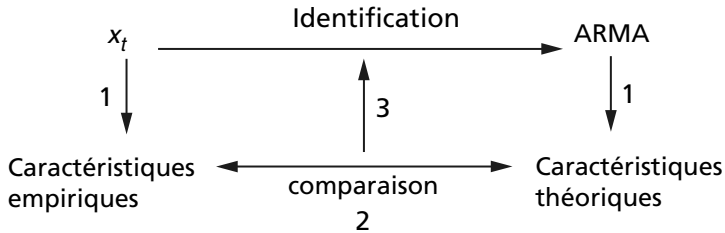
6. L'identification des processus ARMA

Nous avons traité, au chapitre précédent, de la manière de rendre stationnaire un échantillon d'un processus aléatoire non stationnaire. En toute rigueur, il faudrait tester la linéarité du processus aléatoire générateur. Si l'on suppose ce test réalisé, il est alors possible de chercher quel est le processus générateur de la chronique dans la classe des processus ARMA linéaires et stationnaires.

Cette recherche porte le nom, dans l'algorithme de Box-Jenkins, d'étape d'identification. C'est une étape délicate qui conditionne la prévision de la chronique. C'est la raison pour laquelle de nombreux travaux ont conduit à différentes procédures d'identification. Identifier un processus ARMA, c'est trouver – dans la classe des processus précédemment décrits – celui qui est susceptible de s'adapter le mieux aux données empiriques qui constituent la chronique. Cette identification pourrait être réalisée de différentes façons :

- par constitution d'un grand fichier informatique contenant une multitude de processus ARMA simulés avec des coefficients différents de telle sorte que les processus soient stationnaires et inversibles. Il suffirait alors de comparer les chroniques à ces simulations. Cette façon de procéder est irréalisable. L'ordre des processus est trop élevé et les processus sont des équations de récurrence dont la valeur de départ est inconnue, de plus la variance des processus ne serait jamais égale à celle de l'échantillon ;
- par constitution d'un « panier » de processus ARMA théoriques sensé représenter la plupart des chroniques économiques (AR (1), MA (1)...ARMA (1, 1)...). Connaissant chacun des processus théoriques du panier et l'estimation des paramètres de la chronique échantillon, on retiendrait alors la représentation qui fournirait la plus petite somme des carrés des résidus. Cette méthode est aussi à rejeter car elle ne peut pas garantir que le processus générateur sous-jacent à la chronique se trouve bien dans le panier de processus ARMA sélectionnés.

L'identification du processus ARMA s'opère, donc, selon la méthodologie préconisée par Box et Jenkins en comparant les mêmes caractéristiques empiriques de la chronique et théoriques des processus ARMA.



Un processus est caractérisé par ses moments. Les plus importants, parmi ceux-ci, sont les fonctions d'autocorrélation puisqu'il s'agit de processus du second ordre.

Dans leur ouvrage, Box et Jenkins (1976) proposent de recourir à la fonction d'autocorrélation (FAC) et à la fonction d'autocorrélation partielle (FAP) de la chronique.

I. Les caractéristiques des processus AR(p)

Soit un processus AR (p) centré¹ :

$$\phi_p(B)x_t = a_t \text{ non saisonnier}$$

$$x_t = \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + a_t$$

A. Caractéristiques de la FAC d'un AR(p)

Pour $k > 0$, calculons la fonction d'autocovariance :

$$\gamma_k = E[x_t x_{t-k}]$$

$$\gamma_k = \phi_1 E[x_{t-1} x_{t-k}] + \dots + \phi_p E[x_{t-p} x_{t-k}] + E[a_t x_{t-k}]$$

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p} + 0$$

avec $E[a_t x_{t-k}] = 0$ puisque x_{t-k} ne dépend des a_j que pour $j < t - k$.

1. Si le processus est non centré : $\phi_p(B)x_t = a_t + \delta$ alors $E(x_t) = \mu = \frac{\delta}{\left(1 - \sum_{j=1}^p \phi_j\right)}$

(cf. chapitre 7 section III.). Les autres moments restent inchangés.

En divisant les deux membres de l'équation par :

$$\gamma_0 = E[x_t^2] = E[\phi_1 x_t x_{t-1} + \dots + \phi_p x_t x_{t-p} + x_t a_t] = \phi_1 \gamma_1 + \dots + \phi_p \gamma_p + \sigma_a^2$$

on obtient¹ :

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1 \rho_{k-1} + \dots + \phi_p \rho_{k-p}$$

$$\rho_k - \phi_1 \rho_{k-1} - \dots - \phi_p \rho_{k-p} = 0 \text{ (équation de Yule-Walker)}$$

$$\rho_k (1 - \phi_1 B - \dots - \phi_p B^p) = 0$$

$$\phi_p(B) \rho_k = 0$$

Les conditions de stationnarité vont impliquer la convergence vers zéro de la suite des termes ρ_k . Soit les racines en B de ce polynôme opérateur : $B_1 = \frac{1}{\lambda_1}$, ..., $B_p = \frac{1}{\lambda_p}$, $\phi_p(B) = 0$.

C'est une équation aux différences homogènes et elle a pour solution générale : $\rho_k = A_1 \lambda_1^k + \dots + A_p \lambda_p^k$ où $A_1 \dots A_p$ sont des constantes déterminées à partir de conditions initiales. Cette solution montre que l'on rencontre en général (*i.e.* en dehors des racines doubles) deux types de situation :

- une racine λ_i réelle, le produit $A_i \lambda_i^k$ décroît et tend vers 0 quand k augmente (exponentielle amortie par exemple si $\lambda_i > 0$) ;
- un couple de racines λ_i et λ_j complexes, elles interviennent sous la forme $A_i \lambda_i^k + A_j \lambda_j^k$, soit sous la forme $A \lambda^k \sin(\omega k + \phi)$, c'est-à-dire selon l'évolution d'une sinusoïde amortie.

En conclusion, la FAC théorique d'un AR (p) se présente selon une exponentielle amortie et/ou une sinusoïde amortie donc qui tendent rapidement vers 0.

B. Caractéristique de la FAP d'un AR(p)

Considérons l'équation de Yule-Walker d'un AR(p) : $\phi_p(B)x_t = a_t$ et faisons varier k de 1 à p dans cette équation. On obtient le système d'équations linéaires de Yule-Walker :

$$\begin{bmatrix} \rho_1 \\ \vdots \\ \rho_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ & \ddots & & \vdots \\ & & & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_p \end{bmatrix}$$

et on le note : $\rho = P\Phi$.

1. La seule valeur corrélée avec a_t pour $k = 0$ est a_t elle-même d'où $E(a_t x_{t-k}) = E(a_t x_t) = E(a_t^2) = \sigma_a^2$.

Si on remplace les « ρ » par leurs valeurs empiriques calculées à partir de la chronique échantillon, on obtient : $\Phi = P^{-1}r$. Il est donc possible de pré-estimer les paramètres du processus AR à partir de ce système d'équations de Yule-Walker, c'est ce qui est réalisé dans les logiciels informatiques avant leur estimation définitive.

Utilisons le système d'équation de Yule-Walker pour calculer les premiers paramètres des processus AR(1) et AR(2).

Processus AR(1) : $\phi_1 = \rho_1 = \rho_{11}$

Processus AR(2) :

$$\begin{aligned} \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} &= \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} \\ \Leftrightarrow \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} &= \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \\ \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} &= \begin{bmatrix} 1 & -\rho_1 \\ -\rho_1 & 1 \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \frac{1}{1 - \rho_1^2} \\ \phi_1 &= \frac{\rho_1(1 - \rho_2)}{1 - \rho_1^2} \\ \phi_2 &= \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \rho_{22} \end{aligned}$$

Si on procède de la même façon pour un AR(p), on vérifie que : $\phi_p = \rho_{pp}$.

Ce résultat permet de constater, mais on peut le démontrer, que :

- le dernier coefficient ϕ_p d'un AR(p) est égal à ρ_{pp} le coefficient d'auto-corrélation partiel de même rang ; c'est d'ailleurs ainsi qu'est calculée la FAP empirique d'une chronique-échantillon ;
- la FAP d'un AR(p) a ses p premières valeurs différentes de 0 et les autres sont nulles.

La variance du processus est calculée à partir des résultats précédents. Nous avons vu que :

$$\gamma_0 = \sigma_x^2 = \phi_1 \gamma_1 + \dots + \phi_p \gamma_p + \sigma_a^2.$$

D'où :

$$1 = \phi_1 \rho_1 + \dots + \phi_p \rho_p \frac{\sigma_a^2}{\sigma_x^2}.$$

Soit

$$\sigma_x^2 = \frac{\sigma_a^2}{1 - \phi_1 \rho_1 - \dots - \phi_p \rho_p} \quad \text{et} \quad \sigma_a^2 = \sigma_x^2 (1 - \phi_1 \rho_1 - \dots - \phi_p \rho_p)$$

En connaissant σ_x^2 la variance de la chronique, r_1, \dots, r_p (les estimations de ρ_1, \dots, ρ_p) et ϕ_1, \dots, ϕ_p , on obtient une pré-estimation de la variance du bruit blanc qui est aussi une inconnue dans les modèles lors des estimations.

C. Exemple d'application

Soit le processus AR(1) :

$$x_t = \phi_1 x_{t-1} + a_t \quad |\phi_1| < 1 \quad (\text{condition de stationnarité}).$$

Ce processus centré a pour fonction d'autocovariance :

$$E[x_t x_{t-k}] = \phi_1 E[x_{t-1} x_{t-k}] + E[a_t x_{t-k}]$$

$$\gamma_k = \phi_1 \gamma_{k-1}$$

et pour fonction d'autocorrélation :

$$\rho_k = \phi_1 \rho_{k-1} \iff \rho_k - \phi_1 \rho_{k-1} = 0$$

L'équation de Yule-Walker est une équation aux différences du premier ordre de solution $\rho_k = A_1 \lambda_1^k$ avec $\frac{1}{\lambda_1}$ la racine du polynôme $\phi_1(B) = 0$ qui s'écrit aussi :

$$(1 - \phi_1 B) = 0 \iff B = \frac{1}{\phi_1}$$

D'où

$$B = \frac{1}{\lambda_1} = \frac{1}{\phi_1} \iff \lambda_1 = \phi_1$$

Soit $\rho_k = A_1 \phi_1^k$. Pour $k = 0$ on a $\rho_0 = 1 = A_1 \phi_1^0 = A_1$

$$\rho_k = \phi_1^k$$

Lorsque $0 < \phi_1 < 1$ la fonction d'autocorrélation est une exponentielle amortie et lorsque $-1 < \phi_1 < 0$, une sinusoïde amortie. Comme $p = 1$, nous avons $\phi_1 = \rho_{11}$ et $\rho_{kk} = 0$ ($\forall k, k > 1$).

Enfin

$$\sigma_x^2 = \frac{\sigma_a^2}{1 - \phi_1 \rho_1} = \frac{\sigma_a^2}{1 - \rho_1^2}.$$

La figure 6.1 illustre le comportement théorique de ces deux fonctions.

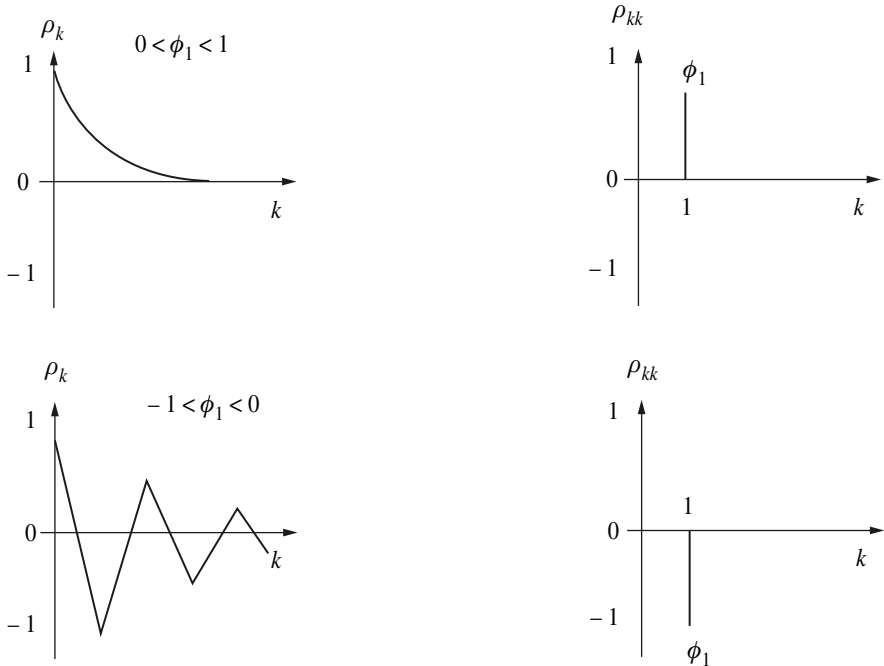


Figure 6.1 – Fonctions d'autocorrélation simple (FAC) et partielle (FAP) d'un AR(1)

Dans le cas des processus AR saisonniers, les résultats précédents s'appliquent mais au retard s de la période de la saisonnalité et à ses multiples ; dans les cas non saisonniers et saisonniers à la fois, une combinaison de ces comportements donne la plupart des allures des FAC et des FAP.

En résumé, un processus AR(p) a les caractéristiques suivantes :

- la FAC est une exponentielle et/ou une sinusoïde amortie ;
- seuls les p premiers termes de la FAP sont significativement différents de 0.

Exercice n° 6.1

Génération de processus AR et analyse des corrélogrammes

Afin de se familiariser avec les FAC et FAP des processus AR, on demande de générer à l'aide de Eviews sur 200 périodes les processus suivants et d'en étudier les FAC et les FAP.

1) AR(1) : $(1 - 0,8B)x_t = a_t$

2) AR(1) : $(1 + 0,8B)x_t = a_t$

3) AR(2) : $(1 - 0,5B - 0,4B^2)x_t = a_t$

4) AR(2) : $(1 + 0,5B + 0,4B^2)x_t = a_t$


5) AR₆(1) : $(1 - 0,6B^6)x_t = a_t$

6) AR₆(1) : $(1 + 0,6B^6)x_t = a_t$

7) AR₆(1) × AR(1) : $(1 + 0,2B^6)(1 - 0,8B)x_t = a_t$

8) AR₆(1) × AR(1) : $(1 + 0,2B^6)(1 + 0,8B)x_t = a_t$

Solution

Le programme Eviews ( fichier **C6EX3.PRG**) permettant de générer les processus est en téléchargement.

Nous ne présentons pas, dans cet ouvrage, les graphiques des FAC et des FAP ; en effet, ceux-ci peuvent être directement obtenus dans tous les logiciels d'analyse de séries temporelles.

II. Les caractéristiques des processus MA(q)

Soit un processus MA(q) centré et non saisonnier¹ :

$$x_t = \theta_q(B)a_t$$

ou encore :

$$x_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

1. Si le processus est non centré : $x_t = \theta_q(B)a_t + \mu$ alors $E(x_t) = \mu$. Les autres moments restent inchangés.

A. Caractéristiques de la FAC d'un MA(q)

Pour $k > 0$, calculons $\gamma_k = E[x_t x_{t-k}]$:

$$\begin{aligned} \gamma_k &= E \left[(a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q})(a_{t-k} - \theta_1 a_{t-k-1} - \dots - \theta_q a_{t-k-q}) \right] \\ &= E \left[a_t a_{t-k} - \theta_1 a_t a_{t-k-1} - \dots - \theta_q a_t a_{t-k-q} \right. \\ &\quad \left. - \theta_1 a_{t-1} a_{t-k} + \theta_1^2 a_{t-1} a_{t-k-1} + \dots + \theta_1 \theta_q a_{t-1} a_{t-k-q} \right. \\ &\quad \vdots \\ &\quad \left. - \theta_q a_{t-q} a_{t-k} + \theta_1 \theta_q a_{t-q} a_{t-k-1} + \dots + \theta_q^2 a_{t-q} a_{t-k-q} \right] \end{aligned}$$

or

$$E[a_t a_{t-k}] = 0 \quad k > q$$

$$E[a_t a_{t-k}] = \sigma_a^2 \quad k \leq q$$

- Pour $k = 0$

$$\gamma_0 = \sigma_x^2 = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_a^2$$

- Lorsque $0 < k \leq q$

$$\gamma_k = (-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q) \sigma_a^2$$

- Et pour $k > q$

$$\gamma_k = 0$$

On en déduit l'écriture de la fonction d'autocorrélation théorique :

$$\begin{cases} \rho_k = \frac{\gamma_k}{\gamma_0} = \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} & 0 < k \leq q \\ \rho_k = 0 & k > q \end{cases}$$

La FAC d'un MA(q) est donc nulle lorsque $k > q$. On constate qu'il s'agit d'un résultat similaire à celui obtenu pour la FAP d'un AR(p).

On peut, comme pour l'AR(p) faire varier k de 1 à q , mais on obtient un système d'équations non linéaires qui ne peut pas être résolu par la méthode de Cramer. On fait alors appel à un algorithme d'optimisation non linéaire qui per-

met de trouver les valeurs de $\theta_1 \dots \theta_q$ en fonction des valeurs empiriques connues de la FAC de la chronique-échantillon. Dans les logiciels, l'algorithme employé pour cette pré-estimation des paramètres du MA(q) est celui de Newton-Raphson.

B. Caractéristiques de la FAP d'un MA(q)

La FAP d'un MA(q) possède une expression analytique relativement complexe. On peut l'approcher en généralisant le résultat obtenu à partir du cas particulier MA(1). Soit un processus MA(1) : $x_t = a_t - \theta_1 a_{t-1}$.

Calculons sa FAC à partir de la formule précédente :

$$\rho_k = \begin{cases} \frac{-\theta_1}{1 + \theta_1^2} & k = 1 \\ 0 & k \geq 2 \end{cases}$$

On a démontré précédemment que les trois premières autocorrélations partielles d'un processus quelconque s'écrivent

$$\rho_{11} = \rho_1, \quad \rho_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \quad \text{et} \quad \rho_{33} = \frac{\rho_1^3 - \rho_1 \rho_2 (2 - \rho_2) + \rho_3 (1 - \rho_1^2)}{1 - \rho_2^2 - 2\rho_1^2 (1 - \rho_2)}$$

Remplaçons ρ_1 , ρ_2 et ρ_3 par les valeurs du processus MA(q), c'est-à-dire :

$$\rho_1 = -\frac{\theta_1}{1 + \theta_1^2} \quad ; \quad \rho_2 = 0 \quad ; \quad \rho_3 = 0$$

D'où

$$\rho_{11} = -\frac{\theta_1}{1 + \theta_1^2} ;$$

$$\rho_{22} = \frac{-\frac{\theta_1^2}{(1 + \theta_1^2)^2}}{1 - \frac{\theta_1^2}{(1 + \theta_1^2)^2}} = -\frac{\theta_1^2}{(1 + \theta_1^2)^2 - \theta_1^2} = -\frac{\theta_1^2}{1 + \theta_1^2 + \theta_1^4}$$

De la même façon, on peut calculer ρ_{33} : $\rho_{33} = -\frac{\theta_1^3}{(1 + \theta_1^4)(1 + \theta_1^2)}$.

Multiplions les numérateurs et les dénominateurs de ces trois autocorrélations partielles par $(1 - \theta_1^2)$:

$$\rho_{11} = -\frac{\theta_1(1 - \theta_1^2)}{(1 - \theta_1^4)} \quad ; \quad \rho_{22} = \frac{-\theta_1^2(1 - \theta_1^2)}{(1 - \theta_1^6)} \quad ; \quad \rho_{33} = \frac{-\theta_1^3(1 - \theta_1^2)}{(1 - \theta_1^8)}$$

On fait apparaître ainsi la relation de récurrence :

$$\rho_{kk} = \frac{-\theta_1^k(1 - \theta_1^2)}{1 - \theta_1^{2(k+1)}}.$$

Dans cette relation, le rapport $\frac{1 - \theta_1^2}{1 - \theta_1^{2(k+1)}}$ est inférieur à 1. La valeur absolue des autocorrélations partielles successives est donc inférieure à θ_1^k . Or, le paramètre θ_1 est tel que $|\theta_1| < 1$ (condition d'inversibilité). Le terme θ_1^k est donc soit une exponentielle amortie ($\theta_1 > 0$ et $\rho_1 < 0$ et les autocorrélations partielles sont négatives) soit une sinusoïde amortie ($\theta_1 < 0$ et $\rho_1 > 0$ et les autocorrélations partielles sont de signes alternés).

Cette démonstration peut être réitérée pour un processus MA(2), mais l'expression de ρ_{kk} est plus complexe. Elle se comporte comme la somme de deux exponentielles si les racines de $\theta_2(B) = 0$ sont réelles, ou comme une sinusoïde si ces mêmes racines sont complexes. On conclut alors que, d'une façon générale, la fonction d'autocorrélation partielle (FAP) d'un MA(q) se comporte, soit comme une exponentielle ou une sinusoïde amortie, soit comme la résultante des deux.

Il s'agit d'un résultat similaire à celui de la FAC d'un AR(p).

Dans le cas des processus MA purement saisonniers, les résultats précédents s'appliquent mais au retard de s' et à ses multiples. Dans le cas non saisonnier et saisonnier à la fois, on observe une combinaison des cas précédents.

D'après la formule précédente de la fonction d'autocovariance, il est également possible d'obtenir une estimation préliminaire de la variance du bruit blanc de ces processus :

$$\sigma_a^2 = \frac{\sigma_x^2}{1 + \theta_1^2 + \dots + \theta_q^2}$$

En résumé, un processus MA(q) a les caractéristiques suivantes :

- seuls les q premiers termes de la FAC sont significativement différents de 0 ;
- la FAP est une exponentielle et/ou une sinusoïde amortie qui tendent donc rapidement vers 0.

Nous pouvons constater la symétrie parfaite des comportements de la FAC d'un AR(p) et de la FAP d'un MA(q) d'une part, et de la FAP d'un AR(p) et de la FAC d'un MA(q) d'autre part. Il suffit d'inverser la lecture de la FAC et de la FAP des graphiques de la figure 6.1.

Exercice n° 6.2

Génération de processus MA et analyse des corrélogrammes

Afin de se familiariser avec les FAC et les FAP des processus MA, on demande de générer à l'aide de Eviews sur 200 périodes les processus suivants et d'en étudier les FAC et les FAP.

1) MA(1) : $x_t = (1 + 0,8B)a_t$

2) MA(1) : $x_t = (1 - 0,8B)a_t$

3) MA(2) : $x_t = (1 + 0,5B + 0,4B^2)a_t$

4) MA(2) : $x_t = (1 - 0,5B - 0,4B^2)a_t$

5) MA₆(1) : $x_t = (1 + 0,6B^6)a_t$

6) MA₆(1) : $x_t = (1 - 0,6B^6)a_t$

7) MA₆(1) × MA (1) : $x_t = (1 + 0,3B^6)(1 + 0,9B)a_t$

8) MA₆(1) × MA (1) : $x_t = (1 + 0,3B^6)(1 - 0,9B)a_t$

Solution

Le programme Eviews ( fichier C6EX4.PRG) permettant de générer les processus est en téléchargement.

Nous ne présentons pas, dans cet ouvrage, les graphiques des FAC et des FAP ; en effet, ceux-ci peuvent être directement obtenus dans tous les logiciels d'analyse de séries temporelles.

III. Les caractéristiques des processus ARMA(p, q)

Soit un processus ARMA(p, q) centré¹ : $\phi_p(B)x_t = \theta_q(B)a_t$

$$x_t = \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q(B)a_t$$

1. Si le processus est non centré : $\phi_p(B)x_t = \theta_q(B)a_t + \delta$ alors $E(x_t) = \frac{\delta}{\left(1 - \sum_{j=1}^p \phi_j\right)}$. Les

autres moments restent inchangés.

A. Caractéristiques de la FAC d'un ARMA(p, q)

Pour $k > 0$ calculons la fonction d'autocovariance :

$$\gamma_k = E[x_t x_{t-k}] = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p} + E[a_t x_{t-k}] - \theta_1 E[a_{t-1} a_{t-k}] - \dots - \theta_q [a_{t-q} x_{t-k}]$$

Les termes $E[a_{t-j} x_{t-k}]$ sont différents de 0 puisque par définition x_{t-k} ne dépend que de $a_{t-k}, \dots, a_{t-k-q}$. On peut alors écrire que :

$$\begin{cases} \text{si } k > 0 & E[a_t x_{t-k}] = 0 \\ \text{si } k \leq 0 & E[a_t x_{t-k}] \neq 0 \end{cases}$$

Ainsi pour $k \geq q + 1$, on a par exemple $E[a_t x_{t-k-q}] = 0$ et dans ce cas :

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p}$$

La fonction d'autocorrélation théorique est alors :

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1 \rho_{k-1} + \dots + \phi_p \rho_{k-p}$$

Soit l'équation de Yule-Walker :

$$\phi_p(B) \rho_k = 0$$

Ce résultat montre que pour établir une règle générale du comportement de la FAC d'un ARMA(p, q), il est nécessaire d'envisager deux cas :

- $q - p < 0$ c'est-à-dire l'ordre de la partie MA est inférieur à celui de la partie AR, dans ce cas la FAC se comporte comme un mélange de fonctions exponentielle - sinusoïdale amorties dont l'allure dépend de la résolution de l'équation de Yule-Walker : $\phi_p(B) \rho_k = 0$, mais aussi des valeurs initiales des coefficients $\rho_q, \rho_{q-1}, \dots, \rho_{q-p+1}$;
- $q - p \geq 0$ c'est-à-dire l'ordre de la partie MA est alors supérieur à l'ordre de AR. Dans ce cas, on obtient $q - p + 1$ valeurs initiales $(\rho_0, \rho_1, \dots, \rho_{q-p})$ qui n'obéiront pas au comportement général précédent, on aura une FAC qui tendra vers 0 quand k augmente, mais son allure est inconnue pour $k \leq q - p$.

La variance du processus est donnée par :

$$\begin{aligned}\gamma_0 &= E[x_t^2] = E[(\phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q})^2] \\ &= \phi_1^2 \gamma_0 + \dots + \phi_p^2 \gamma_0 + \sigma_a^2 + \theta_1^2 \sigma_a^2 + \dots + \theta_q^2 \sigma_a^2 \\ &\quad - 2\phi_1 \theta_1 \sigma_a^2 - \dots - 2\phi_h \theta_h \sigma_a^2\end{aligned}$$

avec $h = \min(p, q)$

D'où :

$$\gamma_0 = \frac{\sigma_a^2(1 + \theta_1^2 + \dots + \theta_q^2 - 2\phi_1 \theta_1 - \dots - 2\phi_h \theta_h)}{1 - \phi_1^2 - \dots - \phi_p^2}$$

B. Caractéristiques de la FAP d'un ARMA(p, q)

En ce qui concerne la FAP, par symétrie avec la FAC, cette fonction possède en général p premières valeurs particulières puis se comporte comme une fonction combinant exponentielles et sinusoïdales amorties.

Pour illustrer les comportements théoriques des FAC et des FAP considérons le cas simple d'un processus ARMA(1, 1) :

$$x_t = \phi_1 x_{t-1} + a_t - \theta_1 a_{t-1}.$$

La fonction d'autocovariance s'écrit :

$$\left\{ \begin{array}{l} \gamma_0 = \phi_1 \gamma_{-1} + E[a_t x_t] - \theta_1 E[a_{t-1} x_t] \\ \gamma_1 = \phi_1 \gamma_0 + E[a_t x_{t-1}] - \theta_1 E[a_{t-1} x_{t-1}] \\ \text{or } E[a_t x_{t-1}] = 0 \text{ car } x_{t-1} \text{ dépend de } a_{t-1} \\ \gamma_1 = \phi_1 \gamma_0 + 0 - \theta_1 E[a_{t-1} x_{t-1}] \\ \gamma_2 = \phi_1 \gamma_1 \\ \vdots \\ \gamma_k = \phi_1 \gamma_{k-1} \end{array} \right.$$

Comme $a_t x_t = \phi_1 a_t x_{t-1} + a_t^2 - \theta_1 a_t a_{t-1}$.

On a $E[a_t x_t] = \sigma_a^2$ et $E[a_{t-j} x_{t-j}] = \sigma_a^2, \forall j$.

De même :

$$\begin{aligned}
 E[a_{t-1}x_t] &= \phi_1 E[a_{t-1}x_{t-1}] + E[a_{t-1}a_t] - \theta_1 E[a_{t-1}^2] \\
 &= \phi_1 \sigma_a^2 - \theta_1 \sigma_a^2 \\
 &= (\phi_1 - \theta_1) \sigma_a^2
 \end{aligned}$$

D'où :

$$\begin{cases}
 \gamma_0 = \phi_1 \gamma_{-1} + \sigma_a^2 - \theta_1 (\phi_1 - \theta_1) \sigma_a^2 \\
 \gamma_1 = \phi_1 \gamma_0 - \theta_1 \sigma_a^2 \\
 \gamma_2 = \phi_1 \gamma_1 \\
 \vdots \\
 \gamma_k = \phi_1 \gamma_{k-1}
 \end{cases}$$

En remplaçant $\gamma_1 = \gamma_{-1}$, on en déduit la fonction d'autocovariance du processus pour $k \geq 0$:

$$\begin{cases}
 \gamma_0 = \frac{1 + \theta_1^2 - 2\theta_1 \phi_1}{1 - \phi_1^2} \sigma_a^2 \\
 \gamma_1 = \frac{(1 - \theta_1 \phi_1)(\phi_1 - \theta_1)}{1 - \phi_1^2} \sigma_a^2 \\
 \vdots \\
 \gamma_k = \phi_1 \gamma_{k-1}
 \end{cases}$$

et la fonction d'autocorrélation théorique :

$$\begin{cases}
 \rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(1 - \theta_1 \phi_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 - 2\theta_1 \phi_1} \\
 \rho_2 = \phi_1 \rho_1 \\
 \vdots \\
 \rho_k = \phi_1 \rho_{k-1}
 \end{cases}$$

La FAC de l'ARMA(1, 1) se comporte comme celle d'un AR(1) : si ϕ_1 est positif nous avons une exponentielle, si ϕ_1 est négatif nous avons une sinussoïde. La présence de la partie MA n'affecte que la première valeur de l'autocorrélation.

La FAP est telle que $\rho_{11} = \rho_1$ puis elle présente l'allure d'une exponentielle et/ou d'une sinusoïdale amortie. Enfin, il est clair qu'avec le système précédent, il est possible de calculer une première estimation des paramètres de l'ARMA(1, 1) en remplaçant les r_i par leurs estimations ρ_i : $\phi_1 = \frac{\rho_2}{\rho_1}$ et θ_1 est solution ($\in [-1; +1]$) de l'équation suivante :

$$\theta_1^2(\rho_1 - \phi_1) + \theta_1(\phi_1^2 - 2\rho_1\phi_1 + 1) + \rho_1 - \phi_1 = 0$$

C. Synthèse

Le tableau 6.1 présente de manière synthétique les propriétés des fonctions d'autocorrélation simple et partielle des processus AR, MA et ARMA.

Tableau 6.1 – Résumé des propriétés des fonctions d'autocorrélation simple et partielle

Processus	FAC	FAP
AR(1)	Décroissance exponentielle ($\phi_1 > 0$) ou sinusoïdale amortie ($\phi_1 < 0$)	Pic significatif pour le premier retard : Positif si $\phi_1 > 0$ et négatif si $\phi_1 < 0$, les autres coefficients nuls pour des retards > 1
AR(2)	Décroissance exponentielle ou sinusoïdale selon les signes de ϕ_1 et ϕ_2	Pics significatifs pour le premier et second retards, les autres coefficients sont nuls pour des retards > 2
AR(p)	Décroissance exponentielle et/ou sinusoïdale rapide	Pics significatifs pour les p premiers retards, les autres coefficients sont nuls pour des retards $> p$
MA(1)	Pic significatif pour le premier retard : positif si $\theta_1 < 0$ et négatif si $\theta_1 > 0$. Les autres coefficients sont nuls pour des retards > 1	Décroissance exponentielle ($\theta_1 > 0$) ou sinusoïdale amortie ($\theta_1 < 0$)
MA(2)	Pics significatifs pour le premier et second retards, les autres coefficients sont nuls pour des retards > 2	Décroissance exponentielle ou sinusoïdale selon les signes de θ_1 et θ_2
MA(q)	Pics significatifs pour les q premiers retards, les autres coefficients sont nuls pour des retards $> q$	Décroissance exponentielle et/ou sinusoïdale
ARMA(1,1)	Décroissance géométrique à partir du premier retard, le signe est déterminé par $\phi_1 - \theta_1$	Décroissance exponentielle ($\theta_1 > 0$) ou sinusoïdale amortie ($\theta_1 < 0$)
ARMA(p,q)	Décroissance exponentielle ou sinusoïdale amortie tronquée après $(q - p)$ retards	Décroissance exponentielle ou sinusoïdale amortie tronquée après $p - q$ retards

Dans les cas saisonniers, les comportements de ces fonctions sont identiques aux périodes de la saisonnalité et de ses harmoniques. Les cas non saisonniers et saisonniers à la fois sont plus complexes à déterminer.

IV. Simulations et exercices

A. Limite à l'utilisation des fonctions d'autocorrélation

On génère un processus AR(1) : $x_t = 2 + 0,9 \times x_{t-1}$ pur sur 1 000 observations, puis on analyse les fonctions d'autocorrélation à partir d'un échantillon de 20, 50, 100 et 1 000 observations.

Les instructions Eviews sont les suivantes :

```
CREATE U 1000 'Création de l'espace de travail
GENR AR1=0
SMPL 2 1000
GENR AR1 = 2 +.9*AR1(-1)
SMPL 1 20
IDENT(8,P) AR1
SMPL 1 50
IDENT(8,P) AR1
SMPL 1 100
IDENT(8,P) AR1
SMPL 1 1000
IDENT(8,P) AR1
```

Pour un échantillon de vingt observations, l'estimation est très médiocre (il convient de rappeler qu'il n'y a pas de perturbations), elle s'améliore avec l'augmentation de la taille de l'échantillon et converge vers les valeurs vraies des fonctions d'autocorrélation.

Cela est à rapprocher du nombre de valeurs, souvent faible, disponible dans les chroniques étudiées et donc de la mauvaise qualité des estimateurs de la FAC et de la FAP.

Résultats

Sample: 120 Included observations: 20						Sample: 150 Included observations: 70							
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob		
		1	0.827	0.827	15.847	0.000			1	0.869	0.898	41.857	0.000
		2	0.655	-0.060	26.664	0.000			2	0.767	-0.017	75.463	0.000
		3	0.514	0.063	33.499	0.000			3	0.855	-0.012	102.117	0.000
		4	0.373	-0.065	37.327	0.000			4	0.611	-0.013	123.24	0.000
		5	0.243	-0.067	39.657	0.000			5	0.534	-0.013	133.69	0.000
		6	0.123	-0.069	39.536	0.000			6	0.463	-0.014	152.37	0.000
		7	0.015	-0.070	39.544	0.000			7	0.399	-0.014	161.59	0.000
		8	-0.062	-0.071	39.789	0.000			8	0.340	-0.015	169.14	0.000

Sample: 100 Included observations: 100						Sample: 1000 Included observations: 1000							
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob		
		1	0.898	0.898	83.020	0.000			1	0.900	0.900	812.36	0.000
		2	0.805	-0.002	150.52	0.000			2	0.810	-0.003	1471.0	0.000
		3	0.722	-0.003	206.33	0.000			3	0.729	-0.003	2034.9	0.000
		4	0.647	-0.003	249.77	0.000			4	0.655	-0.003	2437.8	0.000
		5	0.575	-0.003	285.74	0.000			5	0.559	-0.003	2738.7	0.000
		6	0.462	-0.003	314.78	0.000			6	0.521	-0.003	3075.1	0.000
		7	0.462	-0.003	336.18	0.000			7	0.479	-0.003	3333.7	0.000
		8	0.412	-0.003	356.97	0.000			8	0.459	-0.003	3490.6	0.000

B. Exercices

Exercice n° 6.3

Calcul des termes de la FAC d'un processus MA(2) théorique

$$x_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} \text{ avec } a_t \rightarrow i.i.d.(0, \sigma_a^2)$$

Solution

$$x_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} \text{ avec } a_t \rightarrow i.i.d.(0, \sigma_a^2)$$

$$x_{t-k} = a_{t-k} + \theta_1 a_{t-1-k} + \theta_2 a_{t-2-k}$$

$$\gamma_0 = \sigma_x^2 = E(x_t x_t) = E[(a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2})(a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2})]$$

$$\gamma_0 = E(a_t^2) + \theta_1 E(a_{t-1}^2) + \theta_2 E(a_{t-2}^2) = (1 + \theta_1 + \theta_2)\sigma_a^2$$

$$\gamma_1 = E(x_t x_{t-1}) = E[(a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2})(a_{t-1} + \theta_1 a_{t-2} + \theta_2 a_{t-3})]$$

$$\gamma_1 = \theta_1 E(a_{t-1}^2) + \theta_1 \theta_2 E(a_{t-2}^2) = (\theta_1 + \theta_1 \theta_2)\sigma_a^2$$

$$\gamma_2 = E(x_t x_{t-2}) = E[(a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2})(a_{t-2} + \theta_1 a_{t-3} + \theta_2 a_{t-4})]$$

$$\gamma_2 = \theta_2 E(a_{t-2}^2) = \theta_2 \sigma_a^2$$

$\gamma_k = 0$ pour $k > 2$ car tous les $E(x_t x_{t-k})$ sont nuls.

Exercice n° 6.4

Calcul des termes de la FAC et de la FAP de processus ARMA

Soit les processus suivants :

- | | |
|---|--|
| 1) $x_t = 1 + a_t + 0,5a_{t-1}$ | $a_t \rightarrow i.i.d.(0, \sigma_a^2 = 1)$ |
| 2) $x_t = 0,4x_{t-1} - 0,2x_{t-2} + 40 + a_t$ | $a_t \rightarrow i.i.d.(0, \sigma_a^2 = 12,8)$ |
| 3) $x_t = 1 + a_t - 0,5a_{t-1} - 0,2a_{t-2}$ | $a_t \rightarrow i.i.d.(0, \sigma_a^2 = 1)$ |
| 4) $x_t = 2 + 0,8x_{t-1} + a_t$ | $a_t \rightarrow i.i.d.(0, \sigma_a^2 = 4)$ |
| 5) $x_t = 0,5x_{t-1} + a_t + 0,8a_{t-1}$ | $a_t \rightarrow i.i.d.(0, \sigma_a^2 = 1)$ |

a) Vérifier que le processus est stationnaire et inversible.

b) Calculer $E[x_t]$.

c) Calculer la variance et les trois premières valeurs des autocovariances et des autocorrélations.

d) Calculer les trois premières valeurs des autocorrélations partielles.

Solution

1) Soit le processus : $x_t = 1 + a_t + 0,5a_{t-1}$ avec $a_t \rightarrow i.i.d.(0, \sigma_a^2 = 1)$

a) Inversibilité : $(1 - 0,5B) = 0 \rightarrow B = 2 > 1$. Le processus est inversible.

Un processus MA est toujours stationnaire.

b) $E[x_t] = 1 + E[a_t] + 0,5E[a_{t-1}] = 1$

c) Variance et autocovariances

$$\gamma_k = E[(x_t - E(x_t))(x_{t-k} - E(x_{t-k}))] = E[(x_t - 1)(x_{t-k} - 1)]$$

$$\gamma_0 = \sigma_x^2 = E[(a_t + 0,5a_{t-1})(a_t + 0,5a_{t-1})] = E(a_t a_t) + 0,5^2 E(a_{t-1} a_{t-1})$$

$$\gamma_0 = \sigma_x^2 = \sigma_a^2 + 0,5^2 \sigma_a^2 = (1 + 0,5^2) \sigma_a^2 = 1,25$$

$$\gamma_1 = E[(a_t + 0,5a_{t-1})(a_{t-1} + 0,5a_{t-2})] = 0,5E(a_{t-1} a_{t-1}) = 0,5 \sigma_a^2 = 0,5$$

$$\gamma_2 = E[(a_t + 0,5a_{t-1})(a_{t-2} + 0,5a_{t-3})] = 0$$

$$\gamma_k = 0 \quad \forall k > 1$$

$$\rho_0 = 1 ; \rho_1 = \gamma_1 / \gamma_0 = 0,5 / 1,25 = 0,4 ; \rho_k = 0 \quad \forall k > 1$$

d) Fonction d'autocorrélation partielle

Soit la relation de récurrence : $\rho_{kk} = \frac{-\theta_1^k (1 - \theta_1^2)}{1 - \theta_1^{2(k+1)}} (\theta_1 = -0,5)$ (cf. II. B.).

$$\rho_{11} = \rho_1 = -\frac{\theta_1}{1 - \theta_1^2} = 0,4$$

$$\rho_{22} = -\frac{\theta_1^2}{1 + \theta_1^2 + \theta_1^4} = -\frac{0,25}{1 + 0,25 + 0,0625} = -0,19$$

$$\rho_{33} = -\frac{\theta_1^3}{(1 + \theta_1^4)(1 + \theta_1^2)} = -\frac{-0,125}{(1 + 0,0625)(1 + 0,25)} = 0,09$$

Remarque : nous pouvons aussi utiliser les formules présentées au chapitre 3 (cf. résolution de l'exercice suivant).

2) $x_t = 0,4x_{t-1} - 0,2x_{t-2} + 40 + a_t \quad a_t \rightarrow i.i.d.(0, \sigma_a^2 = 12,8)$

a) Le processus peut s'écrire $(1 - 0,4B + 0,2B^2)x_t = 40 + a_t$

Le processus est stationnaire si les racines du polynôme d'opérateurs sont à l'extérieur du cercle unité du plan complexe :

$$1 - 0,4B + 0,2B^2 = 0 \Rightarrow B = \frac{0,4 \pm i\sqrt{0,64}}{0,4} = \begin{cases} B' = 1 + i \times 2 \\ B'' = 1 - i \times 2 \end{cases}$$

Le module des nombres complexes B' et B'' est $\rho = \sqrt{1 + 2^2} \approx 2,24 > 1$, le processus est stationnaire.

Nous pouvons aussi vérifier directement à l'aide des conditions suivantes¹ :

$$\begin{cases} \phi_1 + \phi_2 < 1 \\ -\phi_1 + \phi_2 < 1 \\ \phi_2 > -1 \end{cases} \Rightarrow \begin{cases} 0,4 - 0,2 = 0,2 < 1 \\ -0,4 - 0,2 = -0,6 < 1 \\ -0,2 > -1 \end{cases} \text{ Les conditions sont vérifiées, le processus est stationnaire.}$$

b) Le processus est un AR(2) avec constante, il n'est donc pas centré.

$$\begin{aligned} E[x_t] &= 0,4E[x_{t-1}] - 0,2E[x_{t-2}] + 40 + 0 \Rightarrow m_x - 0,4m_x + 0,2m_x = 40 \\ &\Rightarrow m_x = 40/0,8 = 50 \end{aligned}$$

c) Afin de simplifier les calculs des autocovariances nous centrons le processus :

$$x_t - E[x_t] = x_t - 50 = X_t.$$

Le processus centré s'écrit alors : $X_t = 0,4X_{t-1} - 0,2X_{t-2} + a_t$ avec $E[X_t] = 0$.

La fonction d'autocovariance s'écrit :

$$\begin{aligned} E[X_t X_{t-k}] &= 0,4E[X_{t-1} X_{t-k}] - 0,2E[X_{t-2} X_{t-k}] + E[a_t X_{t-k}] \text{ Soit} \\ \gamma_k &= 0,4\gamma_{k-1} - 0,2\gamma_{k-2} + 0 \end{aligned}$$

$$\text{Comme } \gamma_{-1} = \gamma_1 \text{ on a : } \begin{cases} \gamma_1 = 0,4\gamma_0 - 0,2\gamma_1(1) \\ \gamma_2 = 0,4\gamma_1 - 0,2\gamma_0(2) \end{cases}$$

$$\text{De (1) on tire } \gamma_1 = 0,4/1,2\gamma_0 = 1/3\gamma_0$$

$$\text{De plus } \gamma_0 = E[X_t X_t] = V[X_t] = 0,4E[X_{t-1} X_t] - 0,2E[X_{t-2} X_t] + E[a_t X_t]$$

$$\gamma_0 = 0,4\gamma_1 - 0,2\gamma_2 + E[a_t X_t]$$

$$\text{Or } a_t X_t = 0,4X_{t-1}a_t - 0,2X_{t-2}a_t + a_t^2$$

1. Cf. chapitre 3.

D'où $E[a_t X_t] = 0 - 0 + \sigma_a^2$ et $\gamma_0 = 0,4\gamma_1 - 0,2\gamma_2 + \sigma_a^2$

À partir de l'équation (2) on a : $\gamma_2 = 0,4 \times 1/3\gamma_0 - 0,2\gamma_0 = -0,2/3\gamma_0$

D'où $\gamma_0 = 0,4 \times 1/3\gamma_0 - 0,2(-0,2/3)\gamma_0 + \sigma_a^2$

Or $\sigma_a^2 = 12,8 \Rightarrow \gamma_0 = \frac{12,8 \times 3}{2,56} = 15$. D'où :

$$\begin{aligned} \gamma_1 &= 1/3\gamma_0 = 5 ; \gamma_2 = -0,2/3\gamma_0 = -1 ; \gamma_3 = 0,4 \times -1 - 0,2 \times 5 = -1,4 ; \\ \gamma_4 &= 0,4 \times -1,4 - 0,2 \times -1 = -0,36 ; \gamma_5 = 0,4 \times -0,36 - 0,2 \times -1,4 = 0,136 \end{aligned}$$

L'équation de Yule-Walker du processus s'écrit : $\gamma_k = 0,4\gamma_{k-1} - 0,2\gamma_{k-2}$

Soit $\rho_k = \gamma_k/\gamma_0 = 0,4\rho_{k-1} - 0,2\rho_{k-2}$

D'où : $\rho_0 = 1 ; \rho_1 = 0,4\rho_0 - 0,2\rho_{-1} = 0,4 - 0,2\rho_1 \Rightarrow \rho_1 = 0,4/1,2 = 0,33$

$$\begin{aligned} \rho_2 &= 0,4\rho_1 - 0,2\rho_0 = -0,067 ; \rho_3 = 0,4\rho_2 - 0,2\rho_1 = -0,093 ; \\ \rho_4 &= 0,4\rho_3 - 0,2\rho_2 = -0,024 ; \rho_5 = 0,4\rho_4 - 0,2\rho_3 = 0,009 \end{aligned}$$

Nous pouvons vérifier ces résultats à partir des calculs des autocovariances :

$$\begin{aligned} \rho_0 &= 1 ; \rho_1 = \gamma_1/\gamma_0 = 5/15 = 0,33 ; \rho_2 = \gamma_2/\gamma_0 = -1/15 = -0,068 ; \\ \rho_3 &= \gamma_3/\gamma_0 = -1,4/15 = -0,093 ; \rho_4 = \gamma_4/\gamma_0 = -0,36/15 = -0,024 ; \\ \rho_5 &= \gamma_5/\gamma_0 = 0,136/15 = 0,009 \end{aligned}$$

d) Les autocorrélations partielles se calculent à partir des formules suivantes (cf. chapitre 3) :

$$\rho_{kk} = \frac{|P_k^*|}{|P_k|} \text{ . D'où : } \rho_{11} = \frac{|P_1^*|}{|P_1|} = 0,33 \quad P_1 = [1] \quad P_1^* = [\rho_1]$$

La première valeur de l'autocorrélation partielle est égale à la première valeur de l'autocorrélation.

$$\rho_{22} = \frac{|P_2^*|}{|P_2|} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \frac{-0,067 - (0,33)^2}{1 - (0,33)^2} = -0,199$$

$$\rho_{33} = \frac{P_3^*}{P_3} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = \frac{\begin{vmatrix} 1 & 0,33 & 0,33 \\ 0,33 & 1 & -0,067 \\ -0,067 & 0,33 & -0,093 \end{vmatrix}}{\begin{vmatrix} 1 & 0,33 & -0,067 \\ 0,33 & 1 & 0,33 \\ -0,067 & 0,33 & 1 \end{vmatrix}} = -0,00069$$

La FAC décroît de manière exponentielle et ρ_{33} n'est pas significativement différent de 0, ce qui vérifie la nature AR(2) du processus.

3) Soit le processus : $x_t = 1 + a_t - 0,5a_{t-1} - 0,2a_{t-2}$ $a_t \rightarrow i.i.d.(0, \sigma_a^2 = 1)$

a) Inversibilité : $(1 - 0,5B - 0,2B^2) = 0 \rightarrow B = \frac{0,5 \pm \sqrt{1,05}}{-0,4}$
 $B' = -3,8$ et $B'' = 1,31$. Le processus est inversible.

Un processus MA est toujours stationnaire.

b) $E[x_t] = 1 + E[a_t] - 0,5E[a_{t-1}] - 0,2E[a_{t-2}] = 1$

c) Variance et autocovariances

$$\gamma_k = E[(x_t - E(x_t))(x_{t-k} - E(x_{t-k}))] = E[(x_t - 1)(x_{t-k} - 1)]$$

$$\begin{aligned} \gamma_0 &= \sigma_x^2 = E[(a_t - 0,5a_{t-1} - 0,2a_{t-2})(a_t + 0,5a_{t-1} - 0,2a_{t-2})] \\ &= E(a_t a_t) + 0,5^2 E(a_{t-1} a_{t-1}) + 0,2^2 E(a_{t-2} a_{t-2}) = (1 + 0,5^2 + 0,2^2) \sigma_a^2 = 1,29 \end{aligned}$$

$$\begin{aligned} \gamma_1 &= E[(a_t - 0,5a_{t-1} - 0,2a_{t-2})(a_{t-1} - 0,5a_{t-2} - 0,2a_{t-3})] \\ &= -0,5E(a_{t-1} a_{t-1}) - 0,2 \times -0,5E(a_{t-2} a_{t-2}) = (-0,5 + 0,1) \sigma_a^2 = -0,4 \end{aligned}$$

$$\begin{aligned} \gamma_2 &= E[(a_t - 0,5a_{t-1} - 0,2a_{t-2})(a_{t-2} - 0,5a_{t-3} - 0,2a_{t-4})] \\ &= -0,2E(a_{t-2} a_{t-2}) = -0,2 \sigma_a^2 = -0,2 \end{aligned}$$

$$\gamma_k = 0 \quad \forall k > 2$$

$$\rho_0 = 1 ; \rho_1 = \gamma_1 / \gamma_0 = -0,31 ; \rho_2 = -0,155 ; \rho_k = 0 \quad \forall k > 2$$

d) Fonction d'autocorrélation partielle

$$\rho_{11} = \rho_1 = -0,31$$

$$\rho_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = -0,28$$

$$\rho_{33} = \frac{\rho_3^3 - \rho_1 \rho_2 (2 - \rho_2) + \rho_3 (1 - \rho_1^2)}{1 - \rho_2^2 - 2\rho_1^2 (1 - \rho_2)} = -0,177$$

4) Soit le processus : $x_t = 2 + 0,8x_{t-1} + a_t$ $a_t \rightarrow i.i.d.(0, \sigma_a^2 = 4)$

a) Stationnarité : $(1 - 0,8B) = 0 \rightarrow B = 1,25$. Le processus est stationnaire.

Un processus AR est toujours inversible.

b) $E[x_t] = 2 + 0,8E[x_{t-1}] + E[a_t] \rightarrow m_x = 2 + 0,8m_x + 0$

$$E[x_t] = m_x = 10$$

c) Variance et autocovariances

$$\gamma_k = E[(x_t - E(x_t))(x_{t-k} - E(x_{t-k}))] = E[(x_t - 10)(x_{t-k} - 10)]$$

$$\gamma_0 = \sigma_x^2 = E[(0,8x_{t-1} + a_t)(0,8x_{t-1} + a_t)] = 0,8^2 E(x_{t-1}^2) + E(a_t^2) + 2 \times 0,8E(x_{t-1} a_t)$$

$$\text{Or } E(x_{t-1}^2) = \gamma_0 \text{ et } E(x_{t-1} a_t) = 0$$

$$\gamma_0 = 0,8^2\gamma_0 + \sigma_a^2 \rightarrow \gamma_0 = \frac{\sigma_a^2}{1 - 0,8^2} = 11,11$$

Nous avons : $\gamma_k = \phi_1\gamma_{k-1} = \phi_1^k\gamma_0$

$$\gamma_1 = \phi_1\gamma_0 = 0,8 \times 11,11 = 8,8 ; \gamma_2 = \phi_1\gamma_1 = 0,8 \times 8,8 = 7,11 ;$$

$$\gamma_3 = \phi_1\gamma_2 = 0,8 \times 7,11 = 5,69$$

$$\rho_0 = 1 ; \rho_1 = \gamma_1/\gamma_0 = 0,8 ; \rho_2 = 0,64 ; \rho_3 = 0,51 \text{ et } \rho_k = \phi_1\rho_{k-1} = \phi_1^k$$

d) Fonction d'autocorrélation partielle

$$\rho_{11} = \rho_1 = 0,8$$

$$\rho_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = 0$$

$$\rho_{33} = \frac{\rho_3^3 - \rho_1\rho_2(2 - \rho_2) + \rho_3(1 - \rho_1^2)}{1 - \rho_2^2 - 2\rho_1^2(1 - \rho_2)} = 0$$

$$\rho_{kk} = 0 \quad \forall k > 1$$

5) $x_t = 0,5x_{t-1} + a_t + 0,8a_{t-1} \quad a_t \rightarrow i.i.d.(0, \sigma_a^2 = 1)$

a) $(1 - 0,5B)x_t = (1 - 0,8B)a_t$

Un processus ARMA est inversible par sa partie MA et stationnaire par sa partie AR.

$(1 - 0,5B) = 0 \rightarrow B = 2 > 1 \rightarrow$ le processus est stationnaire.

$(1 - 0,8B) = 0 \rightarrow B = 1,25 > 1 \rightarrow$ le processus est inversible.

b) Le terme constant du processus ARMA est nul, donc le processus est centré.

$$E[x_t] = 0,5E[x_{t-1}] + E[a_t] + 0,8E[a_{t-1}] \rightarrow m_x = 0,5m_x + 0 + 0 \rightarrow m_x = 0$$

c) La fonction d'autocovariance s'écrit :

$$\gamma_k = E[x_t x_{t-k}] = 0,5E[x_{t-1} x_{t-k}] + E[a_t x_{t-k}] + 0,8E[a_{t-1} x_{t-k}]$$

$$\gamma_0 = \sigma_x^2 = E[(0,5x_{t-1} + a_t + 0,8a_{t-1})(0,5x_{t-1} + a_t + 0,8a_{t-1})]$$

$$\gamma_0 = 0,5^2 E(x_{t-1}^2) + 0,5 \times 0,8 E(x_{t-1} a_{t-1}) + E(a_t^2) + 0,8 \times 0,5 E(a_{t-1} x_{t-1}) + 0,8^2 E(a_{t-1} a_{t-1})$$

$$\gamma_0 = 0,5^2 \gamma_0 + 0,5 \times 0,8 \sigma_a^2 + \sigma_a^2 + 0,8 \times 0,5 \sigma_a^2 + 0,8^2 \sigma_a^2$$

$$\gamma_0 = \frac{1 + 0,8^2 + 2 \times 0,8 \times 0,5}{1 - 0,5^2} \sigma_a^2 \rightarrow \gamma_0 = 3,25 \quad (\sigma_a^2 = 1)$$

$$\gamma_1 = E[x_t x_{t-1}] = 0,5E[x_{t-1} x_{t-1}] + E[a_t x_{t-1}] + 0,8E[a_{t-1} x_{t-1}]$$

$$\gamma_1 = 0,5\gamma_0 + 0,8\sigma_a^2 = 2,425$$

$$\gamma_2 = E[x_t x_{t-2}] = 0,5E[x_{t-1} x_{t-2}] + E[a_t x_{t-2}] + 0,8E[a_{t-1} x_{t-2}] = 0,5\gamma_1 = 1,21$$

$$\gamma_3 = E[x_t x_{t-3}] = 0,5E[x_{t-1} x_{t-3}] + E[a_t x_{t-3}] + 0,8E[a_{t-1} x_{t-3}] = 0,5\gamma_2 = 0,61$$

$$\gamma_0 = 1 ; \rho_1 = \gamma_1/\gamma_0 = 0,75 ; \rho_2 = 0,375 ; \rho_3 = 0,19 \text{ et } \rho_k = \phi_1 \rho_{k-1}$$

d) Fonction d'autocorrélation partielle

$$\rho_{11} = \rho_1 = 0,75$$

$$\rho_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = -0,428$$

$$\rho_{33} = \frac{\rho_3^3 - \rho_1 \rho_2 (2 - \rho_2) + \rho_3 (1 - \rho_1^2)}{1 - \rho_2^2 - 2\rho_1^2 (1 - \rho_2)} = 0,34$$

V. La pratique de l'identification des processus

Il existe actuellement de nombreuses méthodes qui sont des aides à l'identification des processus, même si l'utilisation de la FAC et de la FAP préconisée par Box et Jenkins reste d'actualité. Le recours à d'autres techniques est surtout justifié lorsque se combinent des structures saisonnières et non saisonnières dans le processus ou lorsque le nombre d'observations est insuffisant.

Nous avons retenu dans ce paragraphe deux méthodes qui utilisent d'autres fonctions d'autocorrélations : la fonction d'autocorrélation inverse (et la fonction d'autocorrélation partielle inverse) d'une part, et la fonction d'autocorrélation étendue d'autre part.

A. La fonction d'autocorrélation inverse et la fonction d'autocorrélation partielle inverse

Cleveland (1972) définit la FACI (fonction d'autocorrélation inverse) comme étant l'autocorrélation associée à la fonction inverse de la densité spectrale d'un processus ; Chatfield (1979) en précise les caractéristiques dans le domaine du temps.

Soit une série temporelle réalisation d'un processus stationnaire x_t . La fonction génératrice d'autocovariance du processus est définie par :

$$\gamma(B) = (\dots + \gamma(-1)B^{-1} + \gamma(0)B^0 + \gamma(1)B^1 + \dots) = \sum_{k=-\infty}^{k=\infty} \gamma(k)B^k$$

avec $B =$ opérateur retard.

La fonction génératrice d'autocorrélation est :

$$\rho(B) = \sum_{k=-\infty}^{+\infty} \rho(k)B^k = \frac{\gamma(B)}{\gamma(0)}$$

$\rho(k)$ est la fonction d'autocorrélation de x_t

La fonction génératrice d'autocovariance inverse¹ notée $\gamma_I(B)$ est définie par :

$$\gamma_I(B)\gamma(B) = 1$$

et est telle que :

$$\gamma_I(B) = \sum_{k=-\infty}^{+\infty} \gamma_i(k)B^k$$

où $\gamma_i(k)$ est le coefficient d'autocovariance inverse de retard k . Le coefficient d'autocorrélation inverse est alors :

$$\rho_I(k) = \frac{\gamma_i(k)}{\gamma_i(0)}$$

avec $\gamma_i(0) = \frac{1}{\gamma(0)}$. La fonction génératrice d'autocorrélation inverse est donc :

$$\rho_I(B) = \frac{\rho_I(B)}{\rho_i(0)}$$

Comme les fonctions d'autocovariance et d'autocorrélation constituent des transformées de Fourier des fonctions spectrales, la même définition existe dans le domaine des fréquences.

Si $f(\lambda)$ est le spectre du processus de fréquence λ on a :

$$f(\lambda) = \sum_{k=-\infty}^{k=\infty} \gamma(k)e^{-i2\pi\lambda k},$$

la fonction inverse de $f(\lambda)$ est $f_i(\lambda) = 1/f(\lambda)$ et l'autocovariance inverse

$$\gamma_i(k) = \int_0^1 e^{i2\pi\lambda k} f_i(\lambda) d\lambda.$$

D'où la transformée inverse :

$$f(\lambda) = \sum_{k=-\infty}^{k=\infty} \gamma_i(k)e^{-i2\pi\lambda k};$$

1. Cf. Chatfield, 1979, p. 364

en posant $B = e^{-i2\pi\lambda k}$, $\gamma(B)$ et $\gamma_l(B)$ deviennent respectivement $f(\lambda)$ et $f_l(\lambda)$ ce qui montre l'équivalence des définitions dans le domaine du temps et dans le domaine des fréquences.

Pour un processus ARMA(p, q), la fonction d'autocovariance inverse est égale à :

$$\gamma_l(B) = \frac{\phi_p(B)\phi_p(B^{-1})}{\theta_q(B)\theta_q(B^{-1})} \times \frac{1}{\sigma_a^2}.$$

On peut en déduire pour le retard k de l'autocorrélation inverse du processus :

$$\gamma_i(k) = \begin{cases} \frac{-\phi_k + \sum_{j=1}^{p-k} \phi_j \phi_{i+k}}{1 + \sum_{j=1}^p \phi_j^2} & k \leq p \\ 0 & k > p \end{cases}$$

Pour un AR(p), les $\gamma_i(k)$ s'annulent après le retard p . C'est la même propriété que pour la FAP d'un AR(p). On pourrait démontrer que pour un processus MA(q), la FAPI a un comportement identique à celui de la FAC d'un AR(p). Bhansali R. J. (1980) démontre que pour un bruit blanc donné, l'échantillon des autocorrélations inverses obéit asymptotiquement à une loi normale de moyenne nulle et de variance $\frac{1}{n}$. De sorte que la signification des valeurs de la

FACI peut être contrôlée au seuil de 5 % en les comparant à la quantité $\frac{1,96}{\sqrt{n}}$.

La fonction d'autocorrélation partielle inverse (FAPI) a été introduite par Bhansali R. J. (1980 et 1983). Soit un processus MA(q) : $x_t = \sum_{j=0}^q \beta_j a_{t-j}$ avec

$\beta_0 = 1$ et $a_t \rightarrow i.i.d. (0, \sigma_a^2)$ et le polynôme $\beta_q(B) = \sum_{j=0}^q \beta_j B^j$ avec $|B| \leq 1$.

Cleveland et Parzen ont alors démontré que les β_j satisfont aux équations suivantes :

$$\sum_{j=0}^q \beta_j \rho_i(u - j) = 0$$

pour $u = 1, 2, 3, \dots, q$ où ρ_i est le coefficient d'autocorrélation inverse. Ce sont des équations similaires à celles de Yule-Walker et donc par analogie Bhansali

appelle β_q le coefficient d'autocorrélation inverse partiel, β_q étant la solution de ces équations. En faisant varier q , on obtient ainsi les éléments constitutifs de la FAPI. Elle est donc le complément à l'identification du processus à partir de la FACI comme la FAP est celui de la FAC.

La FAPI d'un processus MA(q) a la même forme que la FAC d'un processus MA(q) et inversement. L'association FAC-FAPI permet l'identification des processus moyennes mobiles et l'association FAP-FACI celle des processus autorégressifs.

Exercice n° 6.5

Représentations graphiques des FAC, FAP, FACI, FAPI

On demande de tracer les FAC, FAP, FACI, FAPI des deux processus simulés suivants :

a) AR(1) : $x_t = 0,9x_{t-1} - 1 + a_t$

b) MA(1) : $x_t = a_t - 0,9a_{t-1}$

avec $a_t \rightarrow i.i.d. (0 ; 1)$

Solution

Les fonctions FACI et FAPI peuvent être calculées (cf. figure 6.2 et figure 6.3) avec le logiciel RATS.

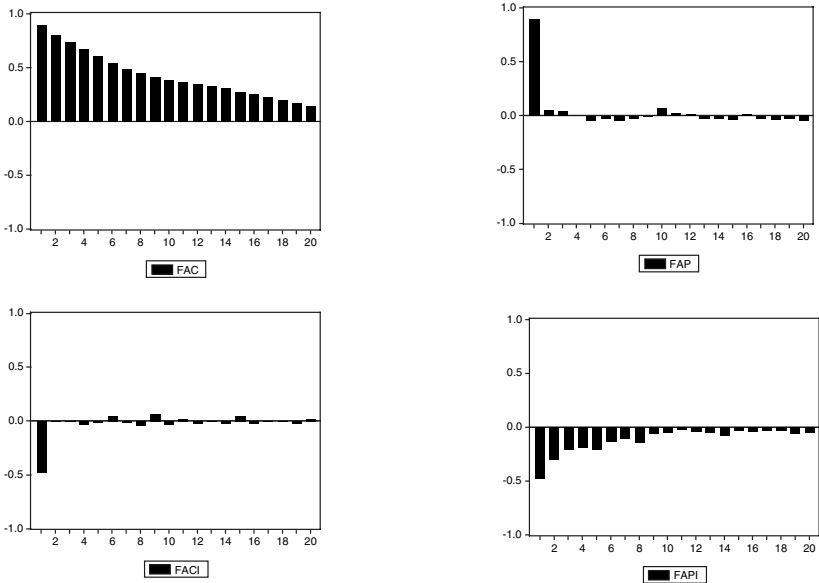


Figure 6.2 – Représentation graphique des FAC, FAP, FACI, FAPI d'un processus AR(1)

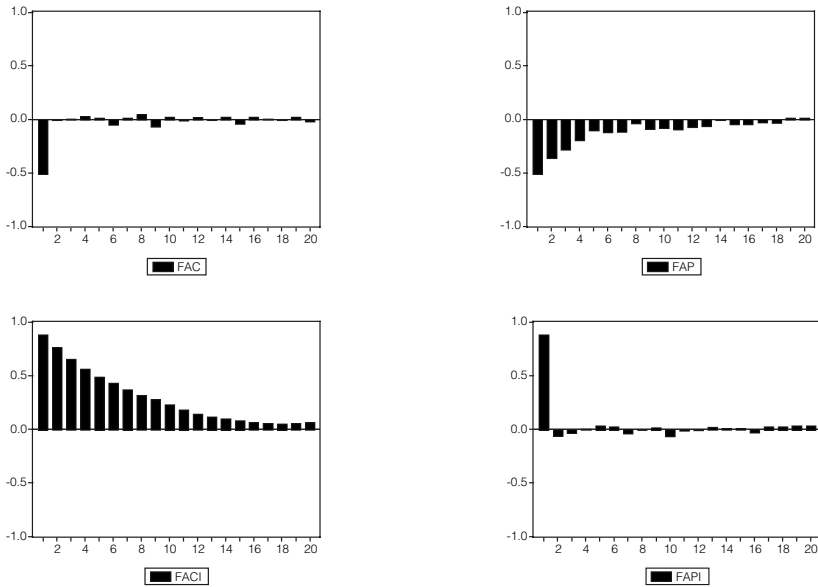
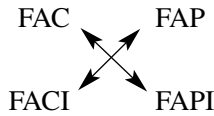


Figure 6.3 – Représentation graphiques des FAC, FAP, FACI, FAPI d'un processus MA(1)

Nous remarquons la correspondance suivante :



B. La fonction d'autocorrélation étendue

La fonction d'autocorrélation étendue (FACE) a été introduite par Tsay R. S. et Tiao G. C. (1984). Son principe est le suivant :

- on effectue des régressions ordinaires de manière itérative pour calculer des estimateurs consistants des paramètres autorégressifs d'un modèle ARMA stationnaire ou non ;
- à partir des estimateurs consistants, la série observée peut être transformée en un processus obéissant asymptotiquement à un MA pur ;
- la FACE est alors calculée sur la série transformée et on l'utilise pour identifier le processus générateur.

Soit une chronique x_t à n observations, réalisation d'un processus ARMA (p, q) :

$$\Phi(B)x_t = \theta(B)a_t \text{ où } a_t \rightarrow i.i.d. (0; \sigma_a^2)$$

$$\Phi(B) = U(B)\phi(B)$$

où $\phi(B)$ est un polynôme d'opérateurs pour lequel ses racines sont à l'extérieur du cercle unité et $U(B)$ un polynôme dont les racines sont sur le cercle unité. Par exemple :

$$\Phi(B) = (1 - B)^d \phi(B) \quad (\text{écriture de Box Jenkins}).$$

$\theta(B)$ est un polynôme d'opérateurs dont les racines sont à l'extérieur du cercle unité. Sur la partie AR du processus, on applique une régression ordinaire qui permet d'obtenir des estimateurs des moindres carrés ordinaires des paramètres AR. Les estimateurs, ainsi obtenus, ne sont pas toujours consistants, c'est-à-dire ne tendent pas en probabilité vers la vraie valeur du paramètre AR. Dans ce cas, les résidus de la régression ne sont pas un bruit blanc. Les auteurs définissent alors la première régression itérative qui s'écrit :

$$x_t = \sum_{l=0}^p \Phi_{l(p)}^{(1)} x_{t-l} + \beta_{1(p)}^{(1)} \widehat{e}_{p,t-1}^{(0)} + e_{p,t}^{(1)}.$$

Ce modèle est estimé par les MCO et les auteurs démontrent que $\widehat{\Phi}_{l(p)}^{(1)}$ est consistant si $q \leq 0$ ou si $\phi(B) = 1$.

On peut alors définir de manière équivalente la $j^{\text{ème}}$ régression itérative d'un AR(k) quelconque :

$$x_t = \sum_{l=1}^k \Phi_{ll(k)}^{(j)} x_{t-l} + \sum_{i=1}^j \beta_{i(k)}^{(j)} \widehat{e}_{k,t-i}^{(j-i)} + e_{k,t}^{(j)}$$

avec $t = k + j + 1, \dots, n$; $j = 0, \dots$ et $k = 1, 2, \dots$

où les $e_{k,t}^{(j)}$ sont les erreurs et les $\widehat{e}_{k,t}^{(j)}$ les résidus de la $i^{\text{ème}}$ régression :

$$\widehat{e}_{k,t}^{(i)} = x_t - \sum_{l=1}^k \widehat{\Phi}_{l(p)}^{(1)} x_{t-l} - \sum_{k=1}^i \widehat{\beta}_{k(k)}^{(i)} \widehat{e}_{k,t-k}^{(i-k)}.$$

Comme l'ordre de la partie autorégressive est inconnu, on utilise $k = 1, \dots, p$ et on effectue q régressions itératives ; p et q sont choisis arbitrairement mais suffisamment grand.

Les estimateurs sont alors calculés à partir de la formule récursive :

$$\widehat{\Phi}_{l(k)}^{(i)} = \frac{\widehat{\Phi}_{l(k+1)}^{(i-1)} - \widehat{\Phi}_{(l-1)(k)}^{(i-1)} \widehat{\Phi}_{(k+1)(k+1)}^{(i-1)}}{\widehat{\Phi}_{k(k)}^{(i-1)}}$$

$$\widehat{\Phi}_{0(k)}^{(i-1)} = -1$$

$$l = 1, \dots, k \quad k \geq 1 \quad i \geq 1$$

Ces estimateurs consistants possèdent des propriétés résumées dans les deux théorèmes suivants.

Théorème 1

Soit x_t un ARMA (p, q) stationnaire ou non.

$$\widehat{\Phi}_{l(k)}^{(i)} = \Phi_l + O_p(n^{-1/2}) \quad l = 1, \dots, k \quad \text{pour} \quad \begin{cases} 1) k \geq p & \text{et } i = q \\ 2) k = p & \text{et } i > k \end{cases}$$

Théorème 2

Soit x_t un ARMA (p, q) non stationnaire.

$$\widehat{\Phi}_{l(d)} = U_l + O_p(n^{-1}) \quad \text{pour } l = 1, \dots, d.$$

Ce deuxième théorème est le plus important pour des séries non stationnaires puisqu'il indique que les estimateurs des moindres carrés ordinaires $\widehat{\Phi}_{l(d)}^{(i)}$ d'une $i^{\text{ème}}$ régression itérée d'un AR (d) sont asymptotiquement consistants pour les coefficients AR non stationnaires noté U_l .

En utilisant le théorème 1 Tsay et Tiao (1984) définissent alors la FACE :

$$r_i(k) = r_i \left(w_{k,t}^{(i)} \right)$$

où $w_{k,t}^{(i)} = x_t - \sum_{l=1}^k \widehat{\Phi}_{l(k)}^{(i)} x_{t-l}$; r_i pour $i = 1, 2, \dots$ représente la FACE du processus x_t . On vérifie que $r_i(0)$ est la FAC habituelle de la série étudiée. En général, la FACE d'un modèle ARMA stationnaire ou non, se comporte asymptotiquement de la façon suivante :

$$r_i(k) \begin{cases} = 0 & \text{pour } i > q \text{ et } k = p \\ \neq 0 & \text{pour } i = q \text{ et } k = p \end{cases}$$

On peut remarquer que la p -ième FACE d'un processus ARMA (p, q) a la même propriété d'identification que la FAC d'un MA (q) puisque les étapes du calcul de la FACE peuvent être résumées par la figure 6.4 :

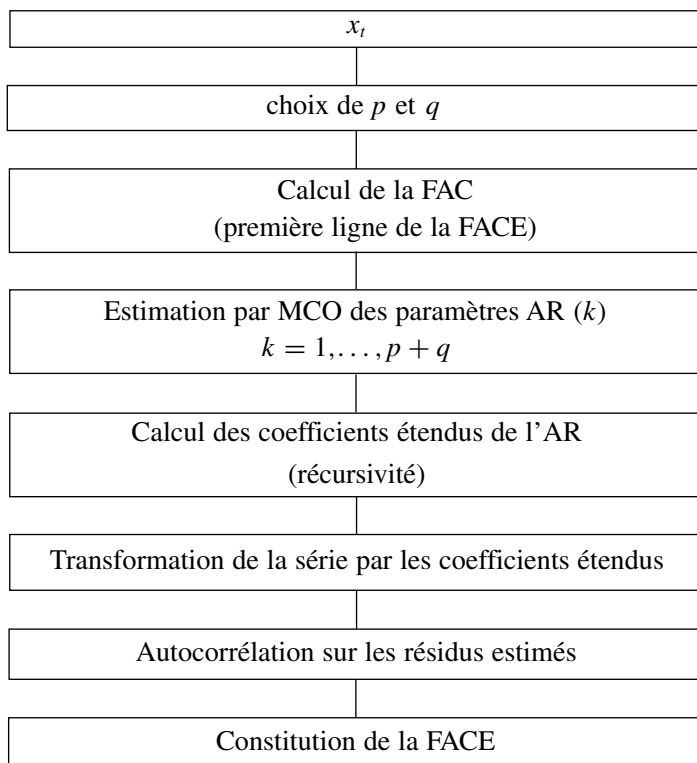


Figure 6.4 – Étapes du calcul de la FACE

Pour identifier les modèles $ARMA(p, q)$, on range les $r_i(k)$ (valeurs de la FACE) dans un tableau à double entrée dans lequel la première ligne du tableau est la FAC de la série, la deuxième ligne est la première FACE, etc. Les lignes sont numérotées de $0, 1, \dots, p$ pour indiquer l'ordre de la partie AR et les colonnes sont numérotées identiquement de $0, 1, \dots, q$ pour indiquer l'ordre de la partie MA. On constitue un tableau (cf. tableau 6.2) dit table FACE :

Tableau 6.2 – Table de FACE

		MA					
		0	1	2	3	...	
AR	0	$r_1(0)$	$r_2(0)$	$r_3(0)$	$r_4(0)$...	← FAC
	1	$r_1(1)$	$r_2(1)$	$r_3(1)$	$r_4(1)$...	← FACE1
	2	$r_1(2)$	$r_2(2)$	$r_3(2)$	$r_4(2)$...	← FACE2
	3
	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Dans le cas d'une simulation, la table FACE présente des valeurs nulles formant un triangle. Les coordonnées du sommet de ce triangle correspondent aux valeurs recherchées p et q du processus. Dans la pratique, les coefficients r ne sont pas exactement nuls, on effectue alors un test de signification de ces coefficients et on remplace les valeurs de la table FACE par une table FACE simplifiée où on repère les valeurs supérieures en valeur absolue à $\frac{1,96}{\sqrt{n}}$ par une croix \times et pour les valeurs à l'intérieur de l'intervalle par 0. Pour illustrer l'utilisation de cette méthode, supposons que le vrai modèle de départ soit un ARMA (1, 2).

Considérons la première ligne de la table FACE, $k = p = 0$ donc le modèle testé est un MA(2) qui a la propriété suivante : $r_i(0) = 0$ pour $i \geq 2$ et comme le vrai modèle n'est pas un MA(2), cette propriété n'est pas vérifiée. La deuxième ligne ($k = 1$ première FACE) a pour propriété $r_i(1) = 0$ pour $i \geq 3$, la troisième ligne ($k = 2$) $r_i(2) = 0$ pour $i \geq 4$. La table FACE simplifiée se présente donc comme sur le tableau 6.3.

Tableau 6.3 – Table FACE simplifiée

		MA					
		0	1	2	3	4	5
AR	0	\times	\times	\times	\times	\times	\times
	1	\times	\times	0	0	0	0
	2	\times	\times	\times	0	0	0
	3	\times	\times	\times	\times	0	0
	\vdots	\times	\times	\times	\times	\times	0

La méthode d'identification consiste à rechercher les coordonnées en ligne et en colonne de la pente du plus grand triangle ne contenant que des 0 que l'on peut dessiner sur la table FACE simplifiée. Cette méthode peut être aussi employée pour la recherche de processus ARMA saisonniers et elle peut s'automatiser sur micro-ordinateur puisqu'il suffit de déterminer l'intersection des deux droites d'équation k et $j - k$. Cependant, la présence d'un 0 ou d'un signe \times est liée au résultat du test de signification des $r_i(h)$. Il est donc recommandé de faire varier dans de légères proportions les intervalles du test pour vérifier la robustesse du triangle de l'identification. L'avantage de cette méthode par rapport aux autres fonctions d'autocorrélation, est son application à des chroniques non stationnaires, non saisonnières et saisonnières. Lorsque cette méthode est apparue, est né l'espoir de pouvoir rendre automatique l'algorithme de Box et Jenkins (ce qui n'est pas le cas avec les FACI et les FAPI) ; mais à l'expérience, cette méthode très séduisante *a priori* possède autant d'incertitude dans ses

résultats que les précédentes. Elle constitue donc un simple complément à l'identification sans plus¹.

Exemple d'utilisation de la FACE

La série étudiée est simulée par un ARMA(2, 1) sur 150 observations² :

$$x_t = 1,32x_{t-1} - 0,68x_{t-2} + a_t - 0,8a_{t-1}$$

avec $a_t \rightarrow i.i.d. (0 ; 1)$

Les coefficients étendus de l'AR (2) sont les suivants :

j	$\hat{\Phi}_{1(2)}^{(j)}$	$\hat{\Phi}_{2(2)}^{(j)}$
0	0,88	- 0,44
1	1,33	- 0,71
2	1,47	- 0,74
3	1,44	- 0,89
4	1,47	- 0,93

On retrouve en $j = 1$ pratiquement les valeurs des paramètres autorégressifs du processus ARMA simulé (1,33 pour 1,32 et - 0,71 pour - 0,68) ; une pré-estimation du paramètre MA est obtenue par le calcul des fonctions d'autocorrélation inverse.

La table de la FACE est présentée au tableau 6.4 et la face simplifiée au tableau 6.5.

Tableau 6.4 – Table de FACE

		MA			
		0	1	2	3
AR	0	0,60	0,08	- 0,30	- 0,50
	1	0,54	0,13	- 0,25	- 0,32
	2	- 0,44	- 0,08	0,12	0,02
	3	- 0,51	- 0,42	0,12	0,03
	4	- 0,32	0,21	- 0,31	0,01
	5	0,30	- 0,07	- 0,30	0,07

1. La méthode FACE est disponible in « PC SCA Statistical System » produit par la Scientific Computing Associated Corp, Chicago, USA.
 2. Cf. Simon, Terraza, 1984.

Tableau 6.5 – Table FACE simplifiée

		MA			
		0	1	2	3
AR	0	×	0	×	×
	1	×	0	×	×
	2	×	0	0	0
	3	×	×	0	0
	4	×	×	×	0
	5	×	0	×	0

On remarque la présence d'un triangle de 0 dont le sommet est en $AR = 2$ et $MA = 1$.

C. Les autres méthodes d'identification

D'autres techniques d'identification d'utilisation courante existent aussi¹. On peut citer, en particulier, les méthodes de type FACE, c'est-à-dire celles qui recherchent dans une table le processus générateur comme : la méthode « R et S » mise au point par Gray, Kelley et Mc Intire (1978) complétée par Woodward, Gray (1981), la méthode du coin de Beguin, Gourieroux et Monfort (1980), et celles utilisant le domaine des fréquences de Terraza (1981) puisque nous avons vu (*cf.* Chapitre 3) qu'il est possible de spécifier la forme analytique du spectre d'un processus ARMA(p, q).

1. En effet, des recherches dans des directions différentes ont été menées par Glasbey, 1982 ; Berliné, 1985 ; Berliné et Van Cutsem, 1986 (Epsilon Algorithme) ou encore par Fuchs (1985) (perturbations de matrices).

7. L'estimation, les tests de validation et la prévision des processus ARMA

Après avoir identifié l'ordre du processus ARMA, il convient d'estimer (section I) les paramètres du modèle, puis de vérifier à partir d'un certain nombre de tests statistiques que l'estimation du modèle est valide (section II). Les conditions étant réunies, nous pouvons alors utiliser le modèle à des fins de prévision (section III). L'algorithme de Box et Jenkins systématise cette démarche qui fait l'objet de ce chapitre.

I. Le problème de l'estimation

L'objectif est de trouver les estimateurs des paramètres de la partie AR et MA du processus de la chronique. L'échantillon utilisé pour l'estimation peut être le résultat d'une transformation par les filtres, par régression, ou encore par la méthode de Box-Cox (*cf.* chapitre 5.IV).

Dans ce dernier cas, deux possibilités se présentent :

- la transformation est effectuée en choisissant la valeur de λ de façon empirique ($\lambda = 0,5$ par exemple). Le problème de l'estimation reste inchangé ;
- la valeur de λ n'a pas été sélectionnée de façon définitive ; on la considère comme un paramètre supplémentaire à estimer simultanément avec les

paramètres ϕ et θ (paramètres du modèle ARMA)¹. Pour réaliser cette estimation, nous pouvons considérer tout d'abord les a_t du processus comme étant les résidus prévisionnels (cf. étape de la prévision) et appliquer le critère des moindres carrés sur ces résidus. Il faut alors minimiser la somme des carrés de ces résidus en utilisant des algorithmes d'optimisation non linéaire. Il est possible aussi de déterminer la vraisemblance du processus ARMA et après quelques transformations, de maximiser cette vraisemblance à l'aide de ces mêmes algorithmes.

A. Détermination et estimation de la vraisemblance des processus ARMA

Soit un processus ARMA(p, q) centré² : $\phi_p(B)x_t = \theta_q(B)a_t$

$$x_t = \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

Pour déterminer la vraisemblance des processus ARMA, il faut faire l'hypothèse que les a_t obéissent à un bruit blanc normal. Si a_t est multinormal, on peut en déduire la distribution d'un vecteur aléatoire issu d'un processus ARMA. Comme le modèle ARMA est linéaire, la distribution du vecteur est alors aussi multivariée. Elle s'écrit :

$$\frac{1}{(2\pi\sigma_a^2)^{n/2}} \frac{1}{(\det \Omega)^{1/2}} \exp \left[-\frac{1}{2\sigma_a^2} x' \Omega^{-1} x \right]$$

où $\sigma_a^2 \Omega$ est une matrice de dimension ($n \times n$) qui contient les variances-covariances du vecteur aléatoire $x(x_1, x_2, \dots, x_n)$. Elle dépend des ϕ_i et des θ_j du processus ARMA.

Cette vraisemblance est difficile à calculer à cause du déterminant de Ω et de son inverse Ω^{-1} quand n est très grand. On contourne cette difficulté en utilisant d'autres expressions de la vraisemblance et en particulier celle fournie par Newbold.

1. Méléard, 1990.

2. Si le processus est non centré : $\phi_p(B)x_t = \theta_q(B)a_t + \delta$ alors $E(x_t) = \frac{\delta}{\left(1 - \sum_{j=1}^p \phi_j\right)}$.

Les autres moments restent inchangés.

$$\frac{1}{(2\pi\sigma_a^2)^{n/2}} \frac{1}{(\det z'z)^{1/2}} \exp\left[-\frac{s}{2\sigma_a^2}\right]$$

où z est une matrice $(p + q + n ; p + q)$ qui dépend des ϕ_i et des θ_j du processus ARMA.

$$s = \sum_{t=1}^n [a_t^2] \text{ avec : } [a_t] = E[a_t/x, \phi, \theta, \sigma_a^2] \text{ l'espérance conditionnelle de } a_t$$

selon x , les paramètres de l'ARMA et la variance de a_t .

La vraisemblance pour la chronique échantillon en logarithme est donnée par :

$$L = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma_a^2 - \frac{1}{2} \log [\det(z'z)] + \log J - \frac{s}{2\sigma_a^2}$$

où $J = \prod_{t=1}^n x_t^{\lambda-1}$ est le jacobien de la transformation Box-Cox si celle-ci est employée.

Il faut donc maximiser L connaissant $\lambda, \phi_i, \theta, \sigma_a^2$: $\text{Max } L/\lambda, \phi_i, \theta_j, \sigma_a^2$.

Pour l'estimation de σ_a^2 , on calcule :

$$\frac{\partial L}{\partial \sigma_a^2} = -\frac{n}{2\sigma_a^2} + \frac{s}{2\sigma_a^2} = 0 \implies \hat{\sigma}_a^2 = \frac{s}{n}$$

On remplace dans l'écriture de L , σ_a^2 par $\hat{\sigma}_a^2$ et on constate alors que maximiser L revient à minimiser :

$$\left[\log \frac{s}{J^{2/n}} + \frac{1}{n} \log[\det(z'z)] \right].$$

B. Les méthodes d'estimation

L'expression précédente est relativement complexe. On fait alors appel à différentes procédures pour la minimiser et en particulier aux moindres carrés conditionnés ou à la méthode de la prévision arrière de Box et Jenkins.

1) Les moindres carrés conditionnés

Les moindres carrés conditionnés consistent à minimiser l'expression : $\frac{\sum_{t=1}^n a_t^2}{J^{2/n}}$ après avoir fixé des valeurs initiales à x_t et à a_t . Si la transformation de Box-Cox n'est pas optimisée, cette méthode revient à minimiser la somme des carrés des résidus prévisionnels à l'aide d'algorithmes d'optimisation non linéaire.

2) La méthode de la prévision arrière de Box et Jenkins

Cette méthode (« *Back Forecasting* ») est fondée sur le calcul itératif des espérances conditionnelles de a_t , de la somme des carrés s . Pour cela, un calcul des valeurs manquantes du processus de départ, inhérent à l'écriture des ARMA, est nécessaire. Au lieu de se fixer des valeurs *a priori*, on utilise la technique de la prévision arrière. Elle consiste à parcourir la chronique « en arrière » à l'aide du processus ARMA, estimé à partir de valeurs terminales fixées *a priori*. L'estimation des paramètres est alors effectuée comme dans le cas précédent en recourant à des algorithmes d'optimisation non linéaire. Des auteurs (cf. Mélard, 1990) ont montré que pour n grand, les solutions trouvées par les différentes méthodes sont identiques.

3) Les algorithmes d'optimisation non linéaire

On utilise couramment deux classes de méthodes ¹ d'optimisation non linéaire :

- les méthodes qui recourent aux notions de géométrie analytique et qui emploient les dérivées partielles de la fonction critère (S dans notre cas). Elles ont pour origine la méthode du gradient. La plus simple d'entre elles appelée « *Steepest Descent* » a été employée par Makridakis et Wheelwright (1974) pour estimer les paramètres des processus ARMA. Elle a conduit à l'élaboration par ces auteurs d'une méthode de prévision autonome (appelée filtre adaptatif généralisé) convenant pour des chroniques comportant un faible nombre d'observations ². L'algorithme de Newton-Raphson fait aussi partie de cette classe de méthodes ; il est utilisé dans la méthode de Box et Jenkins pour pré-estimer les paramètres des processus MA (cf. chapitre 4). Les algorithmes de Fletcher-Powell et de Marquardt sont parmi les plus performants ce qui justifie l'utilisation de celui de Marquardt par Box et Jenkins, pour estimer les paramètres des processus ARMA et actuellement la version améliorée de l'algorithme de Levenberg-Marquardt ;
- les méthodes directes qui n'utilisent pas le calcul des dérivées partielles de la fonction critère. La technique de Rosenbrock (1960) de rotation des axes, approfondissement de celle de Hooke et Jeeves, est une des plus perfectionnée. Elle est considérée comme une généralisation de l'algorithme de dichotomie, utilisé pour calculer les paramètres des lissages exponentiels (cf. chapitre 2). Mais les applications récentes font plutôt appel à la méthode du polyèdre flexible (algorithme du simplexe) de Nelder et Mead (1965) qui converge rapidement vers le minimum de la fonction critère.

Notons que ces algorithmes d'optimisation non linéaire convergent d'autant mieux qu'ils sont appliqués avec des coefficients pré-estimés. Pour les méthodes de type gradient, il s'agit même d'une nécessité, ce qui explique l'importance de la pré-estimation dans la méthode de Box et Jenkins qui recourt à l'algorithme

1. Références sur les algorithmes d'optimisation non linéaire : Wilde, 1976 ; Minoux, 1983 ; Himmelblau, 1972 ; Rao, 1979.

2. Terraza, 1981b.

de Marquardt. Enfin, on peut noter qu'avec des valeurs initiales différentes (pré-estimation, valeurs bornées...), ces algorithmes peuvent conduire à des résultats différents.

II. Les tests de validation

À l'étape de l'identification, les incertitudes liées aux méthodes employées font que plusieurs modèles, en général, sont estimés et c'est l'ensemble de ces modèles qui subissent alors l'épreuve des tests. Il en existe de très nombreux permettant d'une part de valider le modèle retenu, d'autre part de comparer les performances entre modèles.

A. L'analyse des racines

Après avoir estimé un modèle ARMA, on vérifie que la partie AR est stationnaire et la partie MA est inversible puisque les polynômes d'opérateur de la partie AR et de la partie MA ne possèdent pas de racine commune¹ au moyen, par exemple, des algorithmes de Newton-Raphson. Lorsque c'est le cas, on dit qu'il y a redondance et les coefficients estimés du modèle sont instables et peuvent conduire à des prévisions erronées. Il faut alors éliminer dans le modèle ARMA la ou les variables responsables de cette redondance.

B. Le test de Student des paramètres

Les méthodes d'estimation, en particulier celles qui font appel à la vraisemblance du modèle ARMA et donc à l'hypothèse de normalité des a_i , permettent de calculer la matrice des variances covariances des estimateurs. Pour n grand, les covariances se calculent à partir de la matrice d'information de Fisher appro-

chée par $\left(\frac{\delta^2 L}{\delta p_i \delta p_j} \right)_{p=\hat{p}}$ où p_i est un paramètre de la partie AR ou MA du processus estimé.

On effectue alors un test classique de Student sur chacun des paramètres du processus ARMA en divisant le paramètre par son écart type. Il peut arriver qu'un ou plusieurs paramètres ne soient pas significativement différents de 0 : le modèle est alors rejeté et on retourne à l'étape d'estimation en éliminant la variable dont le coefficient n'est pas significatif. En procédant ainsi, on peut construire des processus ARMA troués.

1. Pour des exemples de calcul de racines de polynôme d'opérateur cf. chapitre 3 section III.

C. Le coefficient de détermination

Les coefficients de détermination (R^2 classique ou \bar{R}^2 corrigé) des modèles estimés sont :

$$R^2 = 1 - \frac{\sum_i a_i^2}{\sum_i (x_i - \bar{x})^2}$$
$$\bar{R}^2 = 1 - \frac{n-1}{n-p-q} \frac{\sum_i a_i^2}{\sum_i (x_i - \bar{x})^2} \quad (a_i = \text{résidu d'estimation})$$

On utilise \bar{R}^2 dans de nombreux cas, puisqu'il permet de prendre en compte le nombre de variables explicatives, c'est-à-dire ici les retards p de l'AR et les retards q du MA.

La significativité de ce coefficient de détermination est testée à l'aide d'une statistique de Fisher classique (ce coefficient est proche de 1 si $\sum a_i^2$ tend vers 0). Ce test présente un intérêt limité car il permet seulement de savoir si un ordre de la partie AR ou MA est significatif.

D. Les tests sur les résidus de bruit blanc normal

Le processus estimé est bien évidemment de bonne qualité si la chronique calculée suit les évolutions de la chronique empirique. Les résidus entre les valeurs observées et les valeurs calculées par le modèle, doivent donc se comporter comme l'échantillon d'un bruit blanc normal. Les résidus (ou erreurs de prévision) sont notés par la suite a_t .

1) Test de nullité de la moyenne des résidus

Un bruit blanc est d'espérance mathématique nulle. On effectue donc sur les résidus a_t prévisionnels du modèle un test de nullité de leur moyenne. Pour n suffisamment grand ($n > 30$) $\bar{a} \rightarrow \mathcal{N}\left(m; \frac{\sigma_a}{\sqrt{n}}\right)$ et $\frac{\bar{a} - m}{\sigma_a} \sqrt{n} \rightarrow \mathcal{N}(0; 1)$.

Sous l'hypothèse H_0 de nullité de la moyenne et un seuil d'acceptation à 95 %, l'intervalle de confiance de \bar{a} est :

$$-\frac{1,96\sigma_a}{\sqrt{n}} \leq \bar{a} \leq \frac{1,96\sigma_a}{\sqrt{n}} \quad \text{avec } (t^{\alpha/2} = 1,96).$$

En général cette hypothèse de nullité des résidus est vérifiée car elle découle de la méthode d'estimation des paramètres des processus ARMA.

Remarque : n est le nombre d'observations de la chronique des résidus. Ce n'est pas obligatoirement le nombre d'observations de départ. En effet :

- si le processus est un ARMA (p, q) on perd p observations ;
- si la chronique est filtrée par un filtre aux différences, on perd $(d + sD)$ observations dans le cas général d'un processus saisonnier et non saisonnier à la fois.

2) Tests d'existence d'une tendance

Les résidus – par définition un bruit blanc – ne peuvent contenir des mouvements de long terme de type aléatoire ou déterministe. On peut s'en assurer rapidement à l'aide d'un test de recherche de racine unitaire de type Dickey et Fuller (cf. chapitre 5) ou par les tests présentés au chapitre 2 section A.

3) Tests de recherche d'autocorrélation

Si les résidus obéissent à un Bruit Blanc, il ne doit pas exister d'autocorrélation dans la série. Les tests suivants peuvent être utilisés :

- le test de Durbin-Watson, bien qu'il ne permette de détecter que des autocorrélations d'ordre 1 ;
- la FAC et la FAP mais aussi la FACI et la FAPI ou encore la table FACE, c'est-à-dire toutes les techniques d'identifications présentées au chapitre 6 et qui doivent indiquer une série sans structure ;
- l'analyse du spectre de la série qui doit présenter des fluctuations relativement peu importantes les unes par rapport aux autres. On doit pouvoir tracer dans l'intervalle de confiance du spectre, une ligne parallèle à l'axe des abscisses (cf. chapitres 4 et 6).
- les tests du « porte-manteau » ou test d'adéquation globale du modèle ;

Ces derniers sont bâtis sur le fait que la FAC d'un bruit blanc ne doit pas révéler d'autocorrélations significativement différentes de 0. On recourt alors aux tests statistiques de Box-Pierce et de Ljung-Box pour détecter une éventuelle autocorrélation (cf. chapitre 6) en l'absence d'hétéroscédasticité.

a) Le test de Box et Pierce

Il est établi à partir de la statistique $Q = n \sum_{k=1}^K r_k^2(a_t)$ qui est fonction de la somme des carrés des autocorrélations $r_k^2(a_t)$ de la FAC et du nombre d'observations n . Il permet de vérifier l'hypothèse :

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_K = 0$$

H_1 : il existe au moins un ρ_i significativement différent de 0.

Cette statistique Q en l'absence d'autocorrélation obéit à un χ^2 à $v = K - (p + q)$ degrés de liberté où p est l'ordre de la partie l'AR (saisonnier ou non), q est l'ordre de la partie MA (saisonnaire ou non), et K est le nombre de retards choisis pour calculer les autocorrélations.

Pour effectuer ce test, il est conseillé de choisir K proche du tiers du nombre d'observations. L'hypothèse H_0 est rejetée au seuil de 5 % si Q est supérieur au quantile 0,95 de la loi du χ^2 .

b) Le test de Ljung et Box

La statistique de Ljung-Box est donnée par :

$$Q' = n(n + 2) \sum_{k=1}^K \frac{r_k^2(a_t)}{n - k}.$$

Le test se déroule de manière identique à celui de Box-Pierce.

Dans le cas d'un résidu hétéroscédastique, il convient d'utiliser la statistique de Box-Pierce corrigée¹. On remplace dans la statistique initiale la variance d'autocorrélation de ρ_i par la variance asymptotique $V(\rho_i)$ calculée dans le cas de l'hétéroscédasticité par Lo et Mc Kinlay² selon la formule :

$$V(\rho_i) = \frac{\sum_{t=i+1}^n (a_t^2 a_{t-i}^2)}{\sum_{t=i+1}^n a_t^2}. \text{ Cette statistique suit une loi du } \chi^2 \text{ à } h \text{ degrés de liberté}$$

(h = nombre de retards retenus).

c) Le Runs Test

Ce test examine la fréquence de caractères répétitifs dans une chronique. Le nombre R de « runs » est le nombre de passage du signe « + » au signe « - » et inversement. Si les observations sont indépendantes (donc non corrélées), il est improbable d'avoir un nombre faible de « runs ». Dans le cas d'une indépendance, le nombre attendu de runs est égal à : $m = \frac{1}{n} \left(n(n+1) - \sum_{i=1}^3 n_i^2 \right)$ où n_i est le nombre de runs respectivement positifs, négatifs ou nuls ($i = 1, 2, 3$).

La variance de m est :

$$\sigma_m^2 = \frac{1}{n^2(n-1)} \left(\left(\sum_{i=1}^3 n_i^2 \right) \left(\sum_{i=1}^3 n_i^2 + n(n+1) \right) - 2n \sum_{i=1}^3 n_i^3 - n^3 \right).$$

1. Cf. Mignon, 1998.

2. Lo et Mc Kinlay, 1989.

Pour n grand la distribution est considérée comme normale et on démontre que :

$$z = \frac{R + 0,5 - m}{\sigma_m} \rightarrow \mathcal{N}(0 ; 1).$$

L'hypothèse d'indépendance (H_0) des observations est refusée si $z > 1,96$.

d) Le test de Von Neumann's

On range par ordre croissant les valeurs des résidus. Soit R_i la nouvelle chronique obtenue. Le coefficient du *Von Neumann's Ratio Test* est :

$$RVN = \frac{\sum_{i=1}^{n-1} (R_i - R_{i+1})^2}{\sum_{i=1}^{n-1} (R_i - \mu_R)^2}$$

où μ_R est la moyenne des R_i . L'hypothèse H_0 d'indépendance des valeurs des résidus est rejetée si $RVN > \tau$ où τ est la valeur critique de Bartels (pour un seuil à 95 %, $\tau = 1,67$).

4) Les tests de stabilité des résidus

Ces tests ont pour but de vérifier que les résidus sont stables sur la totalité de la période d'estimation. Dans le cas d'une non stabilité, nous suspectons un changement structurel dans le modèle. On procède alors à une estimation par sous-périodes.

a) Les tests à partir de l'estimation réursive (estimation roulante)

On appelle estimation ou régression réursive, la procédure qui consiste à estimer une succession de régressions en augmentant progressivement le nombre d'observations.

On estime d'abord le modèle avec les $k + 2$ premières observations ($k = p + q$, modèle à un degré de liberté) puis on incorpore l'observation suivante et on re-estime le modèle, et ainsi de suite jusqu'à n (le nombre d'observations total disponible).

Un simple examen graphique de l'évolution des coefficients du modèle - assortis de leurs intervalles de confiance à \pm deux écarts types - permet éventuellement de déterminer des changements structurels.

b) Les tests du « CUSUM » fondés sur la dynamique de l'erreur de prévision

Ces tests¹ permettent de détecter les instabilités structurelles des équations de régression au cours du temps. Contrairement au test de Chow, ils ne nécessitent pas de connaître *a priori* la date de modification structurelle. L'idée générale de ces tests est d'étudier l'évolution au cours du temps de l'erreur de prévision normalisée, on appelle *résidu récuratif* cette succession d'erreur de prévision calculée en $t - 1$ pour t .

1. Brown, Durbin et Evans, 1975, p. 149-163.

Le résidu récursif w_t au temps t est donc défini comme le rapport :

- de l'écart entre la prévision calculée en $t - 1$ pour t et la réalisation en t ;
- et son écart-type calculé pour cette même période t ,

$$\text{soit : } w_t = \frac{e_t = y_t - \hat{y}_t}{S_e} = \frac{y_t - \hat{y}_t}{\sqrt{(1 + x_t'(X'_{t-1}X_{t-1})^{-1}x_t)}}$$

avec $t = K + 2, K + 3, K + 4, \dots, n$ ($K = p + q$ est le nombre total de paramètres estimés du modèle). Ce résidu récursif suit donc une loi normale $\mathcal{N}(0, \sigma_e^2)$.

On utilise alors deux tests :

- le *CUSUM* (CUmulative SUM) fondé sur la somme cumulée des résidus récursifs ;
- le *CUSUM SQ* fondé sur la somme cumulée du carré des résidus récursifs.

• La statistique du *CUSUM*

À partir du résidu récursif w_t , on calcule la statistique *CUSUM* :

$$w_t = \frac{n - K}{SCR} \sum_{j=K+2}^t w_j \text{ avec } t = K + 2, \dots, n \text{ (} K = k + 1 \text{ est le nombre total de}$$

paramètres estimés du modèle) et *SCR* la somme des carrés des résidus calculée à partir du modèle à n observations.

Si les coefficients sont stables au cours du temps, alors les résidus récursifs doivent rester dans l'intervalle défini par les deux droites : $[K, \pm\alpha\sqrt{n - K}]$ et $[n, \pm3\alpha\sqrt{n - K}]$ avec $\alpha = 1,143 - 0,948$ et $0,850$ respectivement pour des seuils de confiance de 1 %, 5 % et 10 %. Dans le cas contraire, le modèle est réputé instable.

• La statistique du *CUSUM SQ*

Elle est donnée par le carré du résidu récursif : $S_t = \frac{\sum_{j=K+2}^t w_j^2}{\sum_{j=K+2}^n w_j^2}$ avec

$t = K + 2, \dots, n$ et $0 \leq S_t \leq 1$; ce test permet de détecter des modifications aléatoires (ponctuelles) dans le comportement du modèle. Si les coefficients ¹ sont stables au cours du temps, alors les résidus récursifs carrés doivent rester dans l'intervalle défini par :

$$\left[\pm C \frac{(t - K)}{n - K} \right] \text{ où } C \text{ est la statistique de Kolmogorov-Smirnov } ^2.$$

1. Démonstration cf. Brown, Durbin et Evans, 1975.
 2. Pour la table cf. Johnston et Dinardo, 1999, p. 379-380.

5) Tests d'homoscédasticité

Un Bruit Blanc est, par définition, homoscédastique. Tous les tests d'hétéroscédasticité peuvent être employés pour vérifier cette hypothèse.

a) Le test de Goldfeld-Quandt a pour but de comparer la somme des carrés des résidus d'estimation après avoir scindé les résidus en deux sous-échantillons. Nous excluons arbitrairement C observations situées au centre de l'échantillon. La valeur de C doit être approximativement égale au quart du nombre d'observations totales. Nous effectuons alors deux régressions sur deux sous-échantillons, et à l'aide d'un test de Fisher nous comparons les deux sommes des carrés des résidus. L'hypothèse d'hétéroscédasticité est acceptée s'il existe une différence significative entre les deux sommes des carrés des résidus.

Soit $SCR1$ la somme des carrés des résidus de la première sous période et $ddl1$ le degré de liberté correspondant et $SCR2$ la somme des carrés des résidus de la deuxième sous période et $ddl2$ le degré de liberté correspondant.

Soit l'hypothèse H_0 : $SCR1 = SCR2$, sous l'hypothèse H_0 d'homoscédasticité, le rapport :

$$F^* = \frac{SCR2/ddl2}{SCR1/ddl1} \text{ suit une loi de Fisher à } ddl2 \text{ et } ddl1 \text{ degrés de liberté. Si}$$

$F^* > F_{ddl2,ddl1}^{0,05}$, l'hypothèse d'homoscédasticité H_0 est rejetée.

Attention : Au numérateur doit figurer le carré moyen le plus élevé afin que le Fisher empirique soit toujours supérieur à 1.

b) les modèles de type ARCH¹ (*AutoRegressive Conditionnal Heteroscedasticity*) permettent de modéliser des chroniques (la plupart du temps financières²) qui ont une volatilité (ou variance ou variabilité) instantanée qui dépend du passé. Il est ainsi possible d'élaborer une prévision dynamique de la chronique en termes de moyenne et de variance. Le test est fondé soit sur un test de Fisher classique, soit sur le test du multiplicateur de Lagrange (*LM*). Le test ARCH ou du multiplicateur de Lagrange suppose que les résidus prévisionnels sont non corrélés et qu'ils obéissent à un modèle ARCH (dans la plupart des cas un modèle ARCH simple d'ordre p). On construit alors une régression entre les résidus au carré et les résidus au carré décalés jusqu'à l'ordre p . Il s'agit en fait d'un modèle AR sur les carrés des résidus. On calcule le coefficient de détermination de ce modèle et on démontre que la quantité $n \times R^2$ obéit asymptotiquement sous l'hypothèse de la nullité de tous les coefficients de régression du modèle à un χ^2 à p degrés de liberté.

Si $n \times R^2 > \chi_{1-\alpha=0,95}^2$ (p degrés de liberté), on rejette H_0 , c'est-à-dire que l'on accepte un effet ARCH (p), soit encore l'hétéroscédasticité des résidus. Ce test et les modèles ARCH sont développés au chapitre 8.

1. Engle, 1982.

2. En effet, les séries financières sont particulièrement concernées par les modèles ARCH, car on constate des périodes de forte spéculation (variabilité élevée) suivie de périodes d'accalmie (variabilité faible).

c) Un autre test consiste à étudier la distribution des carrés des résidus. L'analyse des termes du corrélogramme des résidus au carré permet de tester l'existence d'une hétéroscédacité. Si certaines valeurs de la FAC (tests de Box-Pierce ou Ljung-Box) sont significativement différentes de 0, nous pouvons conclure à la présence d'une hétéroscédacité.

6) Tests de normalité

Nous verrons par la suite que pour calculer des intervalles de confiance prévisionnels et aussi pour effectuer les tests de Student sur les paramètres, il faut que le bruit blanc a_t suive une distribution normale. Le test le plus classique de Jarque et Bera est fondé sur la notion de Skewness (asymétrie) et du Kurtosis (aplatissement).

a) Les tests du Skewness et de Kurtosis

Soit $\mu_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k$ le moment centré d'ordre k , le coefficient de Skewness ($\beta_1^{1/2}$) est égal à : $\beta_1^{1/2} = \frac{\mu_3}{\mu_2^{3/2}}$ et le coefficient de Kurtosis $\beta_2 = \frac{\mu_4}{\mu_2^2}$.

Si la distribution est normale et le nombre d'observations grand :

$$\beta_1^{1/2} \longrightarrow \mathcal{N}\left(0; \sqrt{\frac{6}{n}}\right)$$

et

$$\beta_2 \longrightarrow \mathcal{N}\left(3; \sqrt{\frac{24}{n}}\right)$$

On construit alors les statistiques :

$$v_1 = \frac{\beta_1^{1/2} - 0}{\sqrt{\frac{6}{n}}} \text{ et } v_2 = \frac{\beta_2 - 3}{\sqrt{\frac{24}{n}}} \text{ que l'on compare à } 1,96 \text{ au seuil de } 5\%.$$

Si les hypothèses $H_0 : v_1 = 0$ (symétrie) et $v_2 = 0$ (aplatissement normal) sont vérifiées alors $|v_1| < 1,96$ et $|v_2| < 1,96$, dans le cas contraire, l'hypothèse de normalité est rejetée.

b) Le test de Jarque et Bera

Il s'agit d'un test qui regroupe les résultats précédents, si $\beta_1^{1/2}$ et β_2 obéissent à des lois normales alors la quantité s :

$$s = \frac{n}{6}\beta_1 + \frac{n}{24}(\beta_2 - 3)^2 \text{ suit un } \chi_{1-\alpha}^2(2) \text{ à } 2 \text{ degrés de liberté.}$$

Donc si $s \geq \chi_{1-\alpha}^2(2)$, on rejette l'hypothèse H_0 de normalité des résidus au seuil α .

Ces tests de normalité servent également dans le cas où il y a hétéroscédasticité. En effet, l'hétéroscédasticité se manifeste sur le graphe de la distribution par des queues plus épaisses (distribution leptokurtique) que les queues de la loi normale.

Exercice n° 7.1

Tests de bruit blanc et de normalité sur l'indice CAC40

fichier C7EX1

On demande de vérifier l'hypothèse de normalité de la distribution de l'indice CAC40 des valeurs boursières sur 2 265 observations journalières. Les données sont rendues stationnaires par le passage aux différences premières de la série des logarithmes népériens (cf. chapitre 5).

Solution

Le corrélogramme (cf. Tableau 7.1) ne présente aucun terme significativement différent de 0, ce qui est confirmé par la Q -statistique dont les probabilités critiques sont toutes supérieures à 0,05 (acceptation de H_0 tous les termes sont nuls).

Tableau 7.1 – FAC et FAP de la série $LCAL = \text{Log}(x_t/x_{t-1})$

Sample: 1 2265
Included observations: 2264

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.022	0.022	1.1317	0.287
		2	0.028	0.027	2.8655	0.239
		3	-0.024	-0.025	4.1288	0.248
		4	0.014	0.014	4.5699	0.334
		5	0.001	0.001	4.5708	0.470
		6	-0.000	-0.002	4.5709	0.600
		7	0.033	0.034	7.0304	0.426
		8	-0.010	-0.012	7.2608	0.509
		9	0.045	0.044	11.832	0.223
		10	0.019	0.019	12.627	0.245
		11	0.014	0.009	13.070	0.289
		12	0.042	0.043	17.135	0.145
		13	-0.039	-0.042	20.536	0.083
		14	-0.029	-0.031	22.435	0.070
		15	-0.011	-0.005	22.716	0.090
		16	0.022	0.018	23.825	0.093
		17	-0.007	-0.008	23.932	0.121
		18	0.007	0.005	24.049	0.153

La figure 7.1 présente l'histogramme de la distribution des observations de LCAC ainsi que les calculs des Skewness, Kurtosis et statistique de Jarque-Bera.

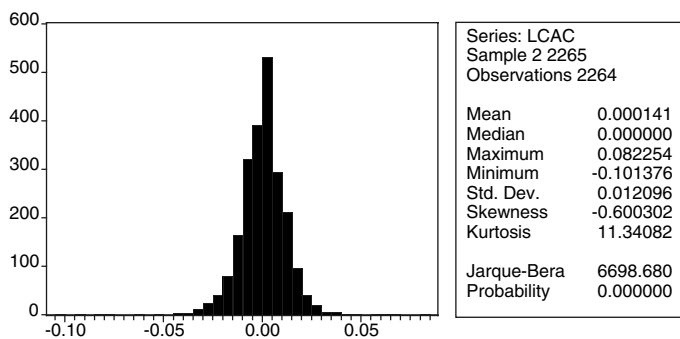


Figure 7.1 – Distribution de l'évolution du CAC40 filtré

La moyenne de la série n'est pas significativement différente de 0.

Nous avons :

$$v_1 = \frac{\beta_1^{1/2} - 0}{\sqrt{\frac{6}{n}}} = \frac{-0,6}{\sqrt{\frac{6}{2264}}} = -11,65 \text{ et } v_2 = \frac{\beta_2 - 3}{\sqrt{\frac{24}{n}}} = \frac{11,34 - 3}{\sqrt{\frac{24}{2264}}} = 81,$$

$|v_1|$ et $|v_2|$ sont supérieurs à 1,96 ; la statistique de Jarque-Bera = 6698 est très largement supérieure au $\chi_{1-\alpha}^2(2)$ lu dans la table (la probabilité critique est nulle), on accepte donc l'hypothèse H_0 de non normalité de la distribution.

La statistique $n \times R^2$ d'un ARCH(1) est 38,0 supérieur au $\chi_{0,95}^2 = 3,84$, on rejette H_0 , c'est-à-dire que l'on accepte un effet ARCH(1).

La série LCAC est une marche aléatoire hétéroscédastique non normale.

E. Les critères de comparaison de modèles

Il arrive fréquemment qu'à l'issue de tous les tests précédents plusieurs modèles se montrent résistants. Pour choisir le meilleur d'entre eux, on peut utiliser des critères de comparaison des modèles. Ces critères sont forts nombreux et jouent, parfois, un rôle important en économétrie.

Ces différents critères, que l'on cherche à minimiser sont fondés sur a_t , l'erreur de prévision, nous pouvons citer :

1) Les critères standard

- L'erreur absolue moyenne (MAE, *Mean Absolute Errors*) :

$$\text{MAE} = \frac{1}{n} \sum_t |a_t|$$

- L'erreur quadratique moyenne (MSE, *Mean Squared Errors*) :

$$\text{MSE} = \frac{1}{n} \sum_t a_t^2$$

$$\text{Soit } m = \frac{1}{n} \sum_{t=1}^n a_t \text{ et } \sigma^2 = \frac{1}{n-1} \sum_{t=1}^n (a_t - m)^2$$

$$\text{MSE} = m^2 + \frac{n-1}{n} \sigma^2$$

- La racine carrée de l'erreur quadratique moyenne (RMSE, *Root Mean Squared Errors*) : $\text{RMSE} = \sqrt{\text{MSE}}$.

- L'écart absolu moyen (MAD, *Mean Absolute Deviation*) :

$$\text{MAD} = \frac{1}{n} \sum_t |a_t - m|$$

- L'écart absolu moyen en pourcentage (MAPE, *Mean Absolute Percentage Error*) :

$$\text{MAPE} = \frac{1}{n} \sum_t \frac{|a_t|}{x_t}$$

- La statistique « U » de Theil

$$U = \left[\frac{\sum_t \left(\frac{\hat{x}_t - x_t}{x_{t-1}} \right)^2}{\sum_t \left(\frac{x_t - x_{t-1}}{x_{t-1}} \right)^2} \right]^{1/2}$$

Ce critère compare l'erreur quadratique moyenne (MSE) avec celle résultant d'une prévision naïve (c'est-à-dire $x_t^{\text{prévu}} = x_{t-1}^{\text{observé}}$). Cette statistique montre que si $U = 0$, la prévision est parfaite ; si $U = 1$, la prévision naïve est aussi bonne que la méthode considérée. Si $U \in]0 ; 1[$, la méthode considérée est supérieure à la naïve. Si $U > 1$, la méthode naïve est supérieure à la méthode considérée.

- L'erreur de prévision finale « *Final Prediction Error* » (FPE) d'Akaike. Soit un $\text{AR}(p)$ stationnaire, si $p = \text{Min} [\text{FPE}(j)/j = 0, \dots, L]$ où L est la limite supérieure de p avec $L < n$ alors :

$$\text{FPE}(p) = \hat{\sigma}_a^2 \frac{n+p}{n-p}$$

Bhansali et Downham (1977) proposent une généralisation plus intéressante de ce critère :

$$\text{FPE}(p) = \hat{\sigma}_a^2 \left[1 + \frac{\alpha p}{n} \right] \text{ avec } p = 0, \dots, L.$$

α est une constante positive utilisée pour réduire la probabilité de sur – ajustement de l'ordre vrai du modèle.

- La fonction transfert autorégressive (CAT, *Criterion for Autoregressing Transfer Functions*) :

$$\text{CAT}(p) = \left(\frac{1}{n} \sum_{j=0}^n \widehat{\sigma}_a^{-2}(j) \right) - \widehat{\sigma}_a^{-2}(p)$$
 avec $p = 0, \dots, L$ et $\widehat{\sigma}_a^2(j)$ est la variance correspondant à l'approximation par un $\text{AR}(p)$ d'un processus $\text{AR}(\infty)$ choisi pour représenter une chronique et pour lequel $V(a_t) = \sigma_\infty^2$.

2) Les critères spécialement construits pour les processus ARMA

- Le critère d'information de Akaike (AIC, *Akaike Informations Criterion*)

Présenté en 1973 pour un $\text{ARMA}(p, q)$, Akaike a démontré que le meilleur des modèles ARMA non filtré est celui qui minimise la statistique :

$$\text{AIC}(p, q) = n \log \widehat{\sigma}_a^2 + 2(p + q)$$

- Le critère d'information bayésien (BIC, *Bayesian Informations Criterion*, Akaike, 1979).

Pour un $\text{ARMA}(p, q)$, il s'écrit :

$$\begin{aligned} \text{BIC}(p, q) = n \log \widehat{\sigma}_a^2 - (n - p - q) \log \left[1 - \frac{(p + q)}{n} \right] \\ + (p + q) \log n + \log \left[(p + q)^{-1} \left[\frac{\sigma_x^2}{\widehat{\sigma}_a^2 - 1} \right] \right] \end{aligned}$$

De manière générale le critère BIC a des caractéristiques plus intéressantes que celles du critère AIC. Il est convergent et pénalise plus fortement les paramètres en surnombre que le critère AIC.

- Le critère d'estimation bayésien (BEC, *Bayesian Estimation Criterion*, Geweke et Meese, 1981). Il s'écrit :

$$\text{BEC}(p, q) = \widehat{\sigma}_a^2 + (p + q) \widehat{\sigma}_L^2 \left[\frac{\log n}{(n - L)} \right]$$

où $\widehat{\sigma}_L^2$ est l'estimateur du maximum de vraisemblance de la variance résiduelle de l'ordre du modèle le plus grand.

- Le critère de Schwarz (1978) :

$$\text{SC}(p, q) = n \log \widehat{\sigma}_a^2 + (p + q) \log n.$$

- Le critère de Hannan-Quin (1979)

$$HQ(p, q) = \log \hat{\sigma}_a^2 + (p + q)c \log \left[\frac{(\log n)}{(n)} \right]$$

où c est une constante à spécifier.

Parmi tous ces critères, les plus employés pour comparer les modèles entre eux sont les critères AIC, BIC, SC et HC. Les autres critères sont essentiellement utilisés pour vérifier des performances prévisionnelles.

III. La prévision

A. Les transformations de départ

À l'étape de la transformation, plusieurs techniques ont pu être employées afin de stationnariser le processus générateur pour qu'il soit identifiable dans la classe des processus ARMA.

Il est nécessaire lors de l'étape de prévision, de prendre en compte la transformation retenue (« recoloration de la prévision ») :

- si on a utilisé la régression afin de retirer une ou plusieurs composantes déterministes estimées par les MCO, ces dernières sont extrapolées jusqu'à l'horizon prévisionnel choisi puis combinées aux valeurs prévisionnelles de l'ARMA ;
- si la transformation résulte de l'utilisation d'un filtre linéaire, alors le processus sous-jacent complet est un processus de type ARIMA que l'on peut considérer – par extension – pour la prévision comme un processus ARMA ;
- si la transformation utilisée est celle de Box-Cox, il faudra recolorer la prévision issue du processus ARMA (prendre l'exponentielle des données si la transformation logarithmique a été appliquée sur la série brute).

La théorie des processus ARMA suppose que le processus de départ est un processus centré. Le fait de centrer les données peut introduire dans le modèle ARMA un terme constant. Deux cas peuvent se présenter :

- la chronique n'a pas subi de transformation, elle est centrée ($x_t - \bar{x}$) et le processus générateur obéit donc à un ARMA(p, q) : $\phi_p(B)(x_t - E(x_t)) = \theta_q(B)a_t$

Soit :

$$\begin{aligned}\phi_p(B)x_t &= \phi_p(B)E(x_t) + \theta_q(B)a_t \\ &= (1 - \phi_1 B - \dots - \phi_p B^p)E(x_t) + \theta_q(B)a_t \\ &= \left(1 - \sum_{j=1}^p \phi_j B^j\right)E(x_t) + \theta_q(B)a_t\end{aligned}$$

D'où $\phi_p(B)x_t = \mu + \theta_q(B)a_t$; $\mu = \left(1 - \sum_{j=1}^p \phi_j\right)E(x_t)$, du fait de la neutralité de l'opérateur B sur une constante. D'où : $E(x_t) = \frac{\mu}{\left(1 - \sum_{j=1}^p \phi_j\right)}$

– la chronique a été transformée par un filtre. Considérons, par exemple, une chronique centrée $x_t - \bar{x} = y_t$ du processus générateur $y_t = x_t - E(x_t)$ filtré : $z_t = (1 - B)y_t$ et supposons que z_t obéisse à un ARMA(p, q) (ou que y_t obéisse à un ARIMA($p, 1, q$)). On a :

$$\begin{aligned}\phi_p(B)(1 - B)y_t &= \theta_q(B)a_t \\ \phi_p(B)(1 - B)(x_t - E(x_t)) &= \theta_q(B)a_t \\ \phi_p(B)(1 - B)x_t &= \phi_p(B)(1 - B)E(x_t) + \theta_q(B)a_t \\ &= \phi_p(B)[E(x_t) - E(x_t)] + \theta_q(B)a_t \\ &= \theta_q(B)a_t\end{aligned}$$

On peut vérifier un résultat similaire à partir d'une chronique centrée après transformation par le filtre : $y_t = (1 - B)x_t = x_t - x_{t-1}$. D'où $\bar{y} \approx 0$ si n suffisamment grand. Si y_t centrée obéit à un ARMA(p, q) alors x_t suit le processus ARIMA($p, 1, q$) précédent.

B. Calcul du prédicteur

Supposons que l'on veuille prévoir un processus ARMA(p, q), (où p, q sont les ordres maximums saisonniers, non saisonniers des parties AR et MA). À un instant t quelconque, de l'axe des temps, le processus possède une valeur x_t , que l'on désire prévoir (calculer) à l'instant $t + l$ au moyen du processus ARMA (p, q) dont on connaît à présent les caractéristiques. On note cette prévision $\hat{x}_t(l)$.

l est dit date d'induction, et $\hat{x}_t(l)$ représente la prévision faite en t pour la date $t + l$. Comme le processus ARMA possède différentes écritures théoriques, la prévision théorique peut s'écrire :

$$\begin{aligned}
 x_{t+l} &= \phi_1 x_{t+l-1} + \dots + \phi_p x_{t+l-p} + a_{t+l} - \theta_1 a_{t+l-1} - \dots - \theta_q a_{t+l-q} \\
 x_{t+l} &= \sum_{j=0}^{\infty} \psi_j a_{t+l-j}; \psi_0 = 1 \\
 x_{t+l} &= \sum_{j=1}^{\infty} \pi_j x_{t+l-j} + a_{t+l} \quad (a_t \text{ est un bruit blanc gaussien})
 \end{aligned}$$

Considérons, par exemple, la seconde forme et le cas simple où $l = 1$. La vraie valeur de la prévision est donc :

$$x_{t+1} = a_{t+1} + \psi_1 a_t + \psi_2 a_{t-1} + \psi_3 a_{t-2} + \dots$$

Dans cette écriture, à l'instant t , a_{t+1} est inconnue et donc la valeur que l'on peut prévoir à l'aide de l'ARMA estimé est :

$$\hat{x}_t(1) = \psi_1 a_t + \psi_2 a_{t-1} + \dots$$

On commet donc une erreur à l'instant t pour la date d'induction 1 :

$$e_t(1) = x_{t+1} - \hat{x}_t(1) = a_{t+1}$$

Ainsi, on constate que les a_t utilisés dans les modèles ARMA comme étant les valeurs d'un bruit blanc sont aussi les erreurs de prévision ayant pour origine le pas de temps immédiatement antérieur. C'est la raison pour laquelle on appelle aussi les a_t les erreurs prévisionnelles ou les innovations.

Considérons à présent une date d'induction l quelconque et supposons que la meilleure prévision soit donnée par l'estimation :

$$\hat{x}_t(l) = \tilde{\psi}_l a_t + \tilde{\psi}_{l+1} a_{t-1} + \tilde{\psi}_{l+2} a_{t-2} + \dots$$

Dans cette expression, les coefficients ψ_l ont été estimés. L'erreur de prévision est alors :

$$e_t(l) = x_{t+l} - \hat{x}_t(l).$$

Pour trouver les $\tilde{\psi}_l$ on va utiliser le critère des MCO, c'est-à-dire calculer les $\tilde{\psi}_l$ qui minimisent l'erreur quadratique moyenne : $\text{Min} E[e_t(l)]^2$

avec

$$\begin{aligned}
 e_t(l) &= x_{t+l} - \hat{x}_t(l) \\
 &= a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1} + (\psi_l - \hat{\psi}_l) a_t + (\psi_{l+1} - \tilde{\psi}_{l+1}) a_{t-1} + \dots
 \end{aligned}$$

Comme a_t est un bruit blanc gaussien, $e_t(l)$ obéit donc à une loi $\mathcal{N}(0; V[e_t(l)])$

$$V[e_t(l)] = \sigma_a^2 \left(1 + \psi_1^2 + \dots + \psi_{l-1}^2 \right) + \sum_{j=1}^{\infty} \left(\psi_j - \tilde{\psi}_j \right)^2 \sigma_a^2$$

comme $E[e_t(l)] = 0$ alors $V[e_t(l)] = E[e_t(l)]^2$ et cette quantité est minimale si $\psi_j = \tilde{\psi}_j$.

On peut de ce fait écrire que :

$$\widehat{x}_t(l) = \psi_l a_t + \psi_{l+1} a_{t-1} + \dots$$

et l'erreur $e_t(l)$ s'écrit :

$$e_t(l) = a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1}$$

On réalise ainsi une prévision non biaisée et la variance de l'erreur de prévision est :

$$V[e_t(l)] = \left(1 + \psi_1^2 + \dots + \psi_{l-1}^2\right) \sigma_a^2$$

On peut noter que la fonction de prédiction $\widehat{x}_t(l)$ est de nature adaptative. En effet :

$$\begin{aligned} \widehat{x}_{t+1}(l) &= \psi_l a_{t+1} + \psi_{l+1} a_t + \psi_{l+2} a_{t-1} + \dots \\ \widehat{x}_{t+1}(l) &= \psi_l a_{t+1} + \widehat{x}_t(l+1) \\ \widehat{x}_{t+1}(l) &= \widehat{x}_t(l+1) + \psi_l e_t(l) \end{aligned}$$

Les prévisions issues des processus ARMA permettent donc de calculer le nouveau prédicateur $\widehat{x}_{t+1}(l)$ de x_{t+l+1} connaissant l'ancien prédicateur $\widehat{x}_t(l)$ d'origine t dès que la valeur du processus en $t+1$: x_{t+1} est connue. On peut ainsi adapter les prévisions faites sur la valeur x_{t+1} à l'époque t lorsque des transformations se manifestent entre t et $t+1$. Cette adaptativité montre la force des modèles ARMA par rapport au modèle de type déterministe. Cela explique aussi pourquoi les méthodes de lissage exponentiel (processus ARMA particulier) ont été abondamment utilisées pour réaliser des prévisions de court terme.

La prévision $\widehat{x}_t(l)$ de variance minimale est aussi l'espérance mathématique de x_{t+l} conditionnelle en x_t, x_{t-1}, \dots . En effet :

$$E[x_{t+l} / x_t, x_{t-1}, \dots] = \psi_l a_t + \psi_{l+1} a_{t-1} + \dots = \widehat{x}_t(l)$$

On constate qu'il est possible de calculer la prévision de variance minimale en utilisant la première écriture des processus ARMA. C'est ainsi que procèdent les logiciels informatiques. Compte tenu de la remarque précédente, on peut écrire :

$$\widehat{x}_t(l) = \sum_{i=1}^p \phi_i E[x_{t+l-i} / x_t, \dots] - \sum_{j=1}^q \theta_j E[a_{t+l-j} / x_t, \dots] + E[a_{t+l-j} / x_t, \dots]$$

or on sait que :

$$\begin{aligned} E[x_{t-j} / x_t] &= x_{t-j} & j &= 0, 1, 2, \dots \\ E[x_{t+j} / x_t] &= \widehat{x}_t(j) & j &= 1, 2, \dots \\ E[a_{t-j} / x_t] &= a_{t-j} = x_{t-j} - \widehat{x}_{t-j-1}(1) & j &= 0, 1, 2, \dots \\ E[a_{t+j} / x_t] &= 0 & j &= 1, 2, \dots \end{aligned}$$

Considérons, par exemple, un processus ARMA(1, 1) :

$$x_t = \phi_1 x_{t-1} + a_t - \theta_1 a_{t-1}$$

Calculons les prévisions en t pour des dates d'induction égales à $1, 2, \dots, l$:

$$\begin{aligned} \widehat{x}_t(1) &= \phi_1 x_t - \theta_1 a_t \\ \widehat{x}_t(2) &= \phi_1 \widehat{x}_t(1) \\ \widehat{x}_t(3) &= \phi_1 \widehat{x}_t(2) \\ &\vdots \\ \widehat{x}_t(l) &= \phi_1 \widehat{x}_t(l-1) \end{aligned}$$

Cet exemple montre que l'on peut prévoir la partie AR de l'ARMA(1, 1) à une date d'induction quelconque l puisque les valeurs observées pourront être remplacées par les valeurs prévues. En revanche, la partie MA ne peut être prévue que pour une date d'induction correspondant au retard du MA, égal ici à 1.

Un processus AR peut donc être prévu jusqu'à un horizon quelconque, tandis qu'un processus MA(q) ne pourra être prévu que jusqu'à l'horizon $t + q$. Pour réaliser les calculs de la prévision lorsqu'il y a une partie MA¹, il faut tenir compte soit des q premières valeurs de a_t inconnues soit du résultat de la prévision arrière de Box et Jenkins.

On peut enfin construire des intervalles de confiance autour des prévisions réalisées. En effet, par hypothèse les a_t sont un bruit blanc gaussien ; la distribution conditionnelle de la valeur à prévoir est donc une variable aléatoire d'espérance $\widehat{x}_t(l)$ et d'écart type $\sigma[e_t(l)]$.

La valeur observée :

$$\frac{x_{t+l} - \widehat{x}_t(l)}{\left(1 + \sum_{j=1}^{l-1} \psi_j^2\right)^{1/2} \sigma_a} \text{ suit donc une loi normale } \mathcal{N}(0; 1).$$

L'intervalle de confiance de la prévision est donné par :

$$x_{t+l} = \widehat{x}_t(l) \pm t^{\alpha/2} \left(1 + \sum_{j=1}^{l-1} \psi_j^2\right)^{1/2} s_{a_t}$$

où s_{a_t} est l'estimation de l'écart type du bruit blanc et $t^{\alpha/2}$ la valeur de la loi normale centrée réduite ($t^{\alpha/2} = 1,96$ pour un intervalle à 95 %).

1. C'est ainsi que l'on procède pour un LES (ARIMA(0, 1, 1)) ou un LED (ARIMA(0, 2, 2)).

Comme nous avons les relations : $x_t = \frac{\theta_q(B)}{\phi_p(B)}a_t$ et $x_t = \psi(B)a_t$, nous en déduisons que $\phi_p(B)\psi(B) = \theta_q(B)$. Les coefficients du polynôme d'opérateur $\psi(B)$ peuvent être calculés connaissant ceux de $\phi_p(B)$ et $\theta_q(B)$ par identification :

$$\begin{aligned}\psi_0 &= 1 \\ \psi_1 &= \phi_1 - \theta_1 \\ \psi_2 &= \phi_1\psi_1 + \phi_2 - \theta_2 \\ &\vdots \\ \psi_j &= \phi_1\psi_{j-1} + \dots + \phi_p\psi_{j-p} - \theta_j\end{aligned}$$

IV. Synthèse de la procédure et exercices d'application

La figure 7.2 illustre le traitement d'une chronique selon la méthodologie de Box-Jenkins.

Exercice n° 7.2

 fichier C7EX2

Analyse de la saisonnalité des immatriculations des voitures particulières en France

Soit la chronique mensuelle des immatriculations des voitures particulières en France connue sur 10 ans (soit $n = 120$), on demande de :

- a) construire le tableau de Buys-Ballot ;
- b) effectuer les tests par analyse de la variance de l'influence du facteur mois et du facteur année ;
- c) représenter sur un même graphique la série brute et la série lissée par une moyenne mobile d'ordre 12 ;
- d) calculer la fonction d'autocorrélation simple et partielle ;
- e) donner la représentation spectrale de la série ;
- f) déterminer le type de schéma de décomposition par le test de Buys-Ballot.

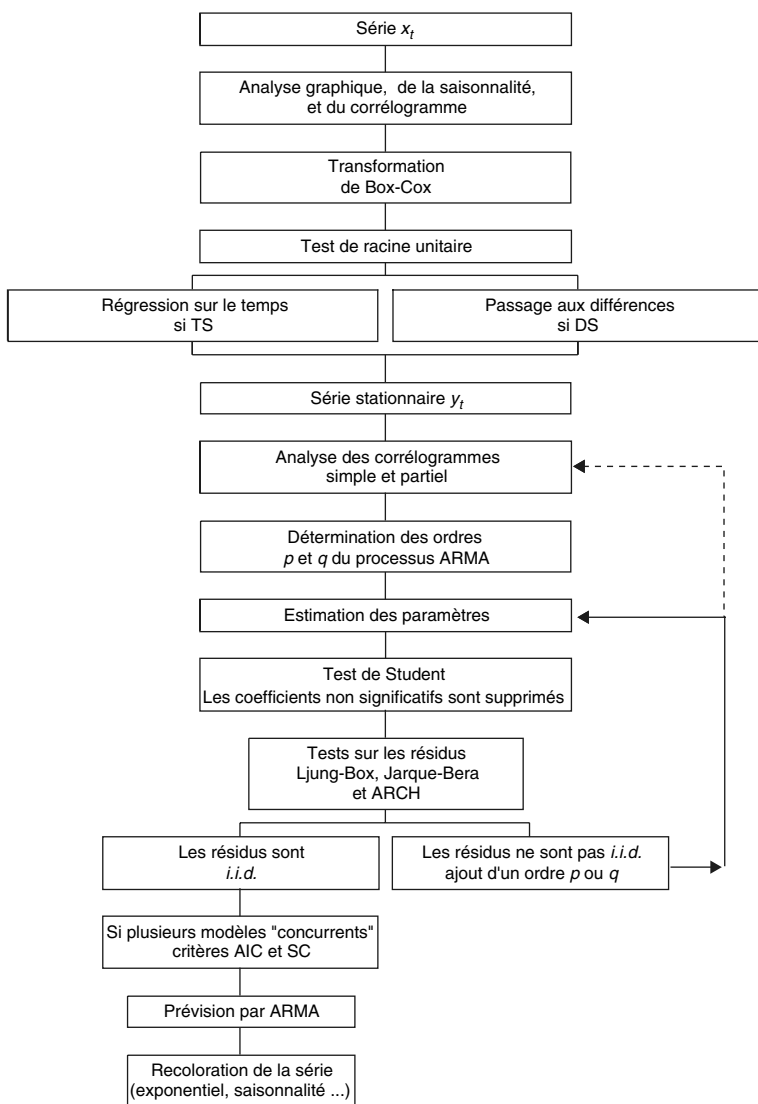


Figure 7.2 – Algorithme de traitement d'une chronique selon la méthodologie de Box-Jenkins

Solutions

La figure 7.3 présente la chronique à prévoir.

- Le tableau 7.2 illustre la construction du tableau de Buys-Ballot des immatriculations.
- Les tests par analyse de la variance de l'influence du facteur mois et du facteur année

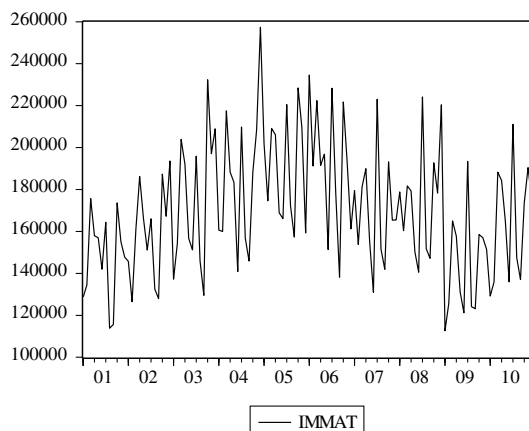


Figure 7.3 – Immatriculations des voitures particulières en France

• Test de l'effet mois (saisonnalité) : $F_c = 17,93$ que l'on compare à $F_{11;99}^{0,05} = 1,89$. La série est donc saisonnière.

• Test de l'effet année (tendance) : $F_c = 12,04$ que l'on compare à $F_{9;99}^{0,05} = 1,97$, l'hypothèse H_0 est rejetée ; la chronique est affectée d'une « tendance » significative.

c) Représentation graphique de la moyenne mobile d'ordre 12

La figure 7.4 présente les immatriculations brutes (IMMAT) et lissées (IMMATL).

Tableau 7.2 – Tableau de Buys-Ballot des immatriculations des voitures particulières en France

Année	Janvier	Février	Mars	Avril	Mai	Juin	Juillet
1	128760	134451	175688	157989	156978	142022	164392
2	145717	126423	161070	186197	166257	151148	166015
3	137236	154265	203865	192080	156703	151181	195919
4	160644	159998	217473	188304	183444	140816	209788
5	201914	174615	209100	206117	168910	166146	220553
6	234571	191225	222418	191541	196828	151378	228317
7	179464	153801	181281	189924	155031	131019	223071
8	178836	160337	181760	179264	150350	140561	224147
9	112659	125634	165040	157739	131237	121171	193383
10	129170	135855	188301	184343	164836	135987	211104
Moyenne	160897,1	151660,4	190599,6	183349,8	163057,4	143142,9	203668,9
Écart type	35963,1	20270,0	20422,4	14332,8	17086,5	11916,2	22152,8

Tableau 7.2 – Tableau de Buys-Ballot des immatriculations des voitures particulières en France (suite)

Année	Août	Septembre	Octobre	Novembre	Décembre	Moyennes	Écart type
1	113888	115637	173619	155112	147792	147194,0	19822,9
2	132501	127960	187352	167294	193587	159293,4	22133,5
3	146058	129552	232302	197077	208942	175431,7	31812,7
4	156838	145859	188181	208455	257349	184762,4	32878,4
5	172228	157306	228476	209686	159256	189525,6	24402,3
6	178136	138220	221772	193530	161194	192427,5	29792,0
7	151677	141960	193187	165398	165461	169272,8	24377,9
8	151789	147204	192756	178299	220397	175475,0	26138,1
9	124170	123147	158521	157063	151458	143435,2	22976,7
10	147451	137127	173721	190458	174566	164409,9	25680,8
Moyenne	147473,6	136397,2	194988,7	182237,2	184000,2		
Écart type	18905,5	11844,4	23627,9	19359,8	33695,6		

Moyenne générale	170122,8
Écart type général	30753,1

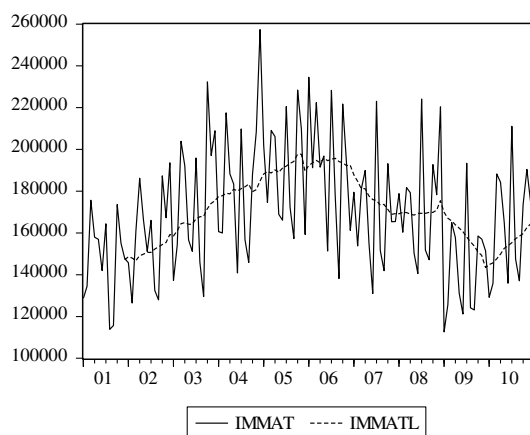


Figure 7.4 – Immatriculations brutes (IMMAT) et lissées (IMMATL)

d) Calcul de la fonction d'autocorrélation simple et partiel.

Sample: 0001M01 0011M08
Included observations: 120

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.140	0.140	2.4177	0.120
		2	0.006	-0.014	2.4223	0.298
		3	0.330	0.338	16.056	0.001
		4	0.294	0.229	26.970	0.000
		5	0.192	0.187	31.675	0.000
		6	0.108	0.003	33.160	0.000
		7	0.212	0.082	39.004	0.000
		8	0.231	0.064	46.000	0.000
		9	0.285	0.226	56.735	0.000
		10	-0.048	-0.228	57.046	0.000
		11	0.042	-0.083	57.287	0.000
		12	0.586	0.451	103.88	0.000
		13	0.029	-0.168	103.99	0.000
		14	-0.056	-0.040	104.43	0.000
		15	0.143	-0.159	107.29	0.000

e) Représentation graphique du spectre.

La figure 7.5 présente le spectre des immatriculations (en logarithme).

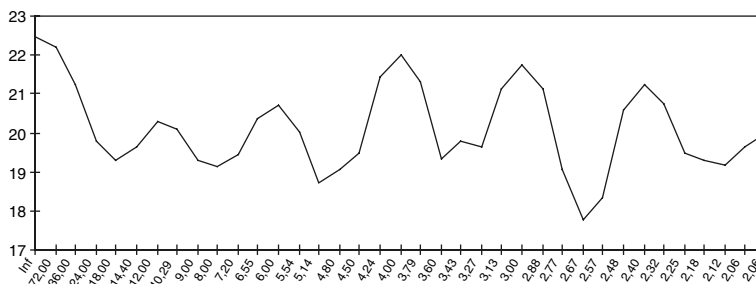


Figure 7.5 – Spectre des immatriculations (en logarithme)

La chronique des immatriculations est saisonnière mais cette cyclicité n'est pas régulière : le mois le plus fort est juillet (nouveau millésime) ou octobre (conséquence des prises de commandes lors du mondial de l'automobile qui a lieu une année sur deux). Cette irrégularité se manifeste par des valeurs des autocorrélations faibles et par l'apparition d'harmoniques de la période 12 sur le spectre. La série temporelle filtrée par la moyenne mobile de longueur 12 ne présente pas de tendance croissante ou décroissante mais plutôt un aspect cyclique.

f) Détermination du type de schéma de décomposition par le test de Buys-Ballot

Pour effectuer ce test nous estimons, par les MCO, les paramètres a_1 et a_0 de l'équation $\sigma_i = a_1 \bar{x}_i + a_0 + \varepsilon_i$ dans laquelle σ_i et \bar{x}_i représentent respectivement les moyennes et les écarts types des dix années disponibles (cf. Tableau 7.2). Les résultats de l'estimation sont :

$$\sigma_i = 0,178 \bar{x}_i - 4432,68 + e_i$$

(2,81)

$$R^2 = 0,50$$

$$n = 10$$

(.) = t de Student.

Le coefficient a_1 est significativement différent de 0 ($t^{cal} = 2,81 > t_8^{0,05} = 2,306$), il existe une relation légèrement significative entre moyenne et écart type, une nouvelle estimation sur les huit dernières années donne le résultat suivant :

$$\sigma_i = 0,13 \bar{x}_i - 4588,36 + e_i$$

(1,61)

$$n = 8$$

(.) = t de Student.

Le coefficient a_1 n'est plus significativement différent de 0 ($t^{cal} = 1,61 < t_6^{0,05} = 2,44$).
Le schéma retenu est donc plutôt additif.

Exercice n° 7.3

Comparaison de méthodes classiques de prévision du nombre des immatriculations en France



fichier C7EX2

On demande de calculer, pour les huit premiers mois de l'année 11, une prévision selon les méthodes suivantes :

- a) lissage exponentiel double ($\alpha = 0,3$) avec saisonnalité additive ;
- b) régression sur variables dichotomiques et tendance linéaire ;
- c) régression sur fonctions trigonométriques et tendance linéaire ;
- d) modèle de Holt-Winters ($\alpha = 0,3 - \beta = \gamma = 0,2$) avec saisonnalité additive.

Solutions (Fichier C7EX3.PRG)

Nous constatons (cf. tableau 7.3) que la méthode de régression sur fonctions trigonométriques présente la somme des carrés des erreurs de prévision la plus faible (série des résidus de IMPFT) et la plus mauvaise technique est le lissage exponentiel double (série des résidus de IMPLIED). En général, pour l'ensemble de ces méthodes, les performances prévisionnelles sont relativement mauvaises.

Tableau 7.3 – Résultats du calcul des prévisions du nombre des immatriculations en France et comparaisons avec les valeurs réalisées

Date	IMMAT	IMPLED	RESIDUS	IMPDIC	RESIDUS	IMPFT	RESIDUS	IMPHW	RESIDUS
11:01	155915	159510	- 3595	160917	- 5002	158584	- 2669	156813	- 898
11:02	151644	148388	3256	151680	- 36	153944	- 2300	149108	2536
11:03	185916	186962	- 1046	190619	- 4703	187732	- 1816	187764	- 1848
11:04	161355	180919	- 19564	183369	- 22014	186087	- 24732	180903	- 19548
11:05	157665	158465	- 800	163077	- 5412	160631	- 2966	159328	- 1663
11:06	198012	137857	60155	143162	54850	145542	52470	139393	58619
11:07	192012	199365	- 7353	203688	- 11676	200727	- 8715	206293	- 14281
11:08	133231	144135	- 10904	147493	- 14262	150866	- 17635	145033	- 11802
Moyennes			2519		- 1032		- 1045		1390
Carrés moyens			524 945 255		488 659 891		472 028 581		521 864 640

IMPLED : Prévission par lissage exponentiel double

IMPDIC : Prévission par variables dichotomiques

IMPFT : Prévission par fonctions trigonométriques

IMPHW : Prévission par modèle de Holt-Winters

IMMAT : valeurs réalisées

Exercice n° 7.4

Prévision par la méthodologie de Box et Jenkins du nombre des immatriculations en France

fichier C7EX2

On demande de prévoir à partir de la méthodologie de Box et Jenkins le nombre des immatriculations des voitures particulières en France (cf. tableau 7.2) pour les huit premiers mois de l'année en utilisant :

- le filtre obtenu par le test de Franses (cf. chapitre 5),
- le filtre aux différences saisonnières d'ordre 3 et de période 12.

Solutions

• Tests de Franses

La tableau 7.4 montre que le modèle sans tendance, et avec constante et variables dichotomiques, noté (c, d, n_t) présente les meilleurs indicateurs statistiques (en particulier la statistique $Q(25)$ de Ljung-Box et le ARCH-Test). Au seuil de 5 % les statistiques t de π_1 et F de $\pi_9 \cap \pi_{10}$ indiquent des valeurs significativement différentes de 0 pour les coefficients de la régression auxiliaire. On en déduit le filtre de la chronique :

$$D3IM = (1 - B)(1 + B + B^2)IMMAT$$

$$D3IM = (1 - B^3)IMMAT = IMMAT - IMMAT(-3)$$

On constate que ce filtre contient une racine unitaire de fréquence 0 qui se justifie par l'aspect cyclique de la composante de long terme.

- Prévision par la méthode de Box-Jenkins pour la série des immatriculations filtrées par les différences saisonnières d'ordre 3.

Le programme Eviews¹ ( C7EX4.PRG) est en téléchargement

Tableau 7.4 – Tests de Franses de racines unitaires saisonnières des immatriculations

Statistique t	nc, nd, nt	c, nd, nt	c, d, nt	c, nd, t	c, d, t
π_1	0,57	-1,59	-1,82	-1,87	-2,04
π_2	-2,11	-2,11	-2,77	-2,09	-2,73
π_3	-0,53	-0,53	-0,24	-0,53	-0,17
π_4	-1,12	-1,13	-3,71	-1,03	-3,65
π_5	-1,53	-1,52	-3,03	-1,42	-2,95
π_6	-2,00	-1,99	-3,74	-1,90	-3,66
π_7	1,23	1,44	2,12	1,66	2,24
π_8	-2,66	-2,83	-3,41	-2,92	-3,44
π_9	-1,08	-1,07	-1,86	-1,02	-1,79
π_{10}	-1,53	-1,54	-3,29	-1,53	-3,20
π_{11}	0,61	0,63	1,050	0,63	1,09
π_{12}	-2,50	-2,48	-3,51	-2,29	-3,38

1. Ce programme calcule les prévisions pour la série des immatriculations filtrée par les différences d'ordre 3 et d'ordre 12.

Tableau 7.4 (suite) – Tests de Franses de racines unitaires saisonnières des immatriculations

Statistique F					
$\pi_3 \cap \pi_4$	0,77	0,78	6,95	0,68	6,68
$\pi_5 \cap \pi_6$	2,10	2,08	7,11	1,94	6,84
$\pi_7 \cap \pi_8$	6,84	7,03	8,27	6,42	7,93
$\pi_9 \cap \pi_{10}$	1,25	1,25	5,48	1,23	5,17
$\pi_{11} \cap \pi_{12}$	3,45	3,37	6,62	2,84	6,06
$\pi_3 \cap \dots \cap \pi_{12}$	3,24	3,25	8,74	2,86	7,92
$Q(25)$ Stat.	21,10	22,10	13,30	23,10	13,50
\bar{R}^2	0,20	0,21	0,43	0,24	0,44
AIC	-1,11	-1,12	-1,35	-1,14	-1,37
SBIC	-3,65	-3,63	-3,60	-3,63	-3,58
ARCH	8,05	11,31	8,10	11,95	9,58
F	3,46	3,43	4,50	3,56	4,48

L'estimation du modèle final, pour la série filtrée par les différences saisonnières, est la suivante :

Dependent Variable: D3IM				
Sample (adjusted): 0002M01 0010M12				
Included observations: 108 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(3)	-0.834026	0.074707	-11.16391	0.0000
AR(6)	-0.836887	0.076203	-10.98233	0.0000
AR(9)	-0.662314	0.075617	-8.758744	0.0000
MA(1)	0.151001	0.094998	1.589527	0.1150
MA(4)	0.260171	0.095365	2.728167	0.0075

Presque tous les coefficients du modèle sont significativement différents de 0, à l'exception du coefficient de MA (1) et le résidu est bien un bruit blanc : la probabilité critique de la statistique Q est toujours supérieure à 0,05 et tous les termes du corrélogramme sont dans l'intervalle de confiance.

Sample: 0002M01 0010M12
Included observations: 108
Q-statistic probabilities adjusted for 5 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.026	0.026	0.0756	
		2 0.093	0.092	1.0360	
		3 0.058	0.054	1.4187	
		4 0.020	0.010	1.4658	
		5 0.097	0.087	2.5471	
		6 0.060	0.052	2.9629	0.085
		7 0.172	0.157	6.4595	0.040
		8 0.081	0.062	7.2348	0.065
		9 0.063	0.031	7.7135	0.103
		10 -0.042	-0.080	7.9297	0.160
		11 0.011	-0.015	7.9443	0.242
		12 -0.121	-0.156	9.7420	0.204

Jarque-Bera = 14,9 (probabilité critique = 0,00)

Les résidus ne sont pas gaussiens : la statistique de Jarque-Bera a une probabilité critique nulle (on accepte l'hypothèse H_0 de non normalité des résidus).

La prévision de la série des différences d'ordre 3 (D3IM), à l'horizon de huit périodes, est la suivante :

$$(1 - B^3) \text{IMMAT} = \text{D3IM}$$

0011M01	- 25 548,35
0011M02	- 37 383,67
0011M03	10 920,56
0011M04	38 511,62
0011M05	6 701,26
0011M06	- 41 195,29
0011M07	14 020,66
0011M08	- 2 787,23

























• Prévision par la méthode de Box-Jenkins pour la série des immatriculations filtrées par les différences saisonnières d'ordre 12.

L'estimation du modèle final, pour la série filtrée par les différences saisonnières, est la suivante :

Dependent Variable: D12IM				
Sample (adjusted): 0002M04 0010M12				
Included observations: 105 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.231505	0.096385	2.401890	0.0181
AR(2)	0.231888	0.096683	2.398435	0.0183
AR(3)	0.201706	0.095175	2.119313	0.0365
MA(12)	- 0.923147	0.026867	- 34.35952	0.0000

Tous les coefficients du modèle sont significativement différents de 0 et le résidu est bien un bruit blanc : la probabilité critique de la statistique Q est toujours supérieure à 0,05 et tous les termes du corrélogramme sont dans l'intervalle de confiance.

Sample: 0002M04 0010M12
Included observations: 105
Q-statistic probabilities adjusted for 4 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.022	-0.022	0.0503	
		2 -0.045	-0.045	0.2680	
		3 -0.077	-0.080	0.9273	
		4 0.033	0.027	1.0484	
		5 -0.008	-0.014	1.0558	0.304
		6 0.101	0.098	2.2090	0.331
		7 0.020	0.029	2.2544	0.521
		8 -0.035	-0.027	2.3927	0.664
		9 0.152	0.172	5.1093	0.403
		10 0.032	0.035	5.2335	0.514
		11 -0.000	0.013	5.2335	0.631
		12 0.082	0.110	6.0432	0.642

Jarque-Bera = 35,91 (probabilité critique = 0,00)

Les résidus ne sont pas gaussiens : la statistique de Jarque-Bera a une probabilité critique nulle (on accepte l'hypothèse H_0 de non normalité des résidus).

La prévision de la série des différences saisonnières (D12IM), à l'horizon de huit périodes, est la suivante :

$$(1 - B^{12}) \text{IMMAT} = \text{D12IM}$$

0011M01	26 460,85
0011M02	15 747,08
0011M03	2 830,59
0011M04	395,81
0011M05	- 550,25
0011M06	6 089,38
0011M07	- 4 450,16
0011M08	718,52

Les prévisions des immatriculations (cf. Tableau 7.5, IMPBJ3 = prévisions pour la série filtrée par les différences d'ordre 3 et IMPBJ12 = prévisions pour la série filtrée par les différences d'ordre 12) sont alors égales à :

Tableau 7.5 – Prévisions par la méthode de Box-Jenkins

Date	IMMAT	IMPBJ12	Résidu	IMPBJ3	Résidu
0011M01	155 915	155 630,85	284,15	148 172,64	7 742,36
0011M02	151 644	151 602,08	41,92	153 074,33	- 1 430,33
0011M03	185 916	191 131,59	- 5 215,59	185 486,56	429,44
0011M04	161 355	184 738,81	- 23 383,81	186 684,26	- 2 5329,26
0011M05	157 665	164 285,75	- 6 620,75	159 775,60	- 2 110,60
0011M06	198 012	142 076,38	55 935,62	144 291,27	53 720,73
0011M07	192 012	206 653,84	- 14 641,84	200 704,93	- 8 692,93
0011M08	133 231	148 169,52	- 14 938,52	156 988,37	- 23 757,37
		Carrés Moyens	523 032 218		529 262 181

Nous constatons que la qualité de la prévision est à peu près similaire à celle des techniques précédentes, c'est-à-dire relativement médiocre. La chronique des immatriculations est une série difficile à prévoir du fait des actions spéciales (primes à la casse, etc.) qui perturbent le marché et ne peuvent pas être prises en compte dans une modélisation univariée.

Exercice n° 7.5

Analyse par la méthodologie de Box et Jenkins et prévision de trois processus aléatoires



fichier C7EX5

Soit les trois processus x_1 , x_2 et x_3 connus sur 220 périodes, on demande de :

- analyser leurs propriétés stochastiques ;
- déterminer, le cas échéant, l'ordre du processus ARMA(p , q) ;
- d'effectuer une prévision à un horizon de 5 périodes.

Solution

Les résultats sont fournis sans les calculs intermédiaires.

1) Analyse du processus x_1

La figure 7.6 présente l'évolution du processus x_1 et son corrélogramme, à leur lecture le processus semble non stationnaire.

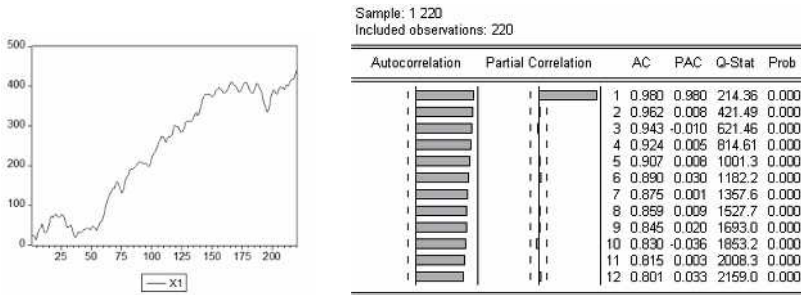


Figure 7.6 – Évolution et fonction d'autocorrélation simple du processus x_1

a) Tests de racines unitaires : tests de Phillips-Perron (fenêtre de Newey-West = 4)

Modèle [1] $t^* = 2,22$ $\alpha c = 0,99$ $\rightarrow H_0$, il existe une racine unitaire.

Modèle [2] $t^* = -0,62$ $\alpha c = 0,86$ $\rightarrow H_0$, il existe une racine unitaire.

Modèle [3] $t^* = -1,65$ $\alpha c = 0,77$ $\rightarrow H_0$, il existe une racine unitaire.

b) Stratégie de tests et tests d'hypothèses jointes

$F_3 = 0,61 < \phi_3 = 6,34 \rightarrow$ acceptation de H_0^3

$F_2 = 6,98 > \phi_2 = 4,75 \rightarrow$ rejet de H_0^2

Estimation du modèle : $\Delta x_1 = c + bt + a_t$

Dependent Variable: D(X1)				
Included observations: 219 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.088617	0.861662	2.423941	0.0162
@TREND	-0.001661	0.006792	-0.244533	0.8070

Acceptation de l'hypothèse $b = 0$

Il s'agit donc d'un processus DS avec dérive : $x_{1t} = x_{1t-1} + c + a_t$

Nous le stationarisons par le passage aux différences premières.

c) Analyse des fonctions d'autocorrélation simple et partielle sur la série stationnaire

Sample: 1 220
Included observations: 219

	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1			0.575	0.575	73.421	0.000
2			0.184	-0.219	80.993	0.000
3			-0.020	-0.036	81.081	0.000
4			-0.113	-0.068	83.969	0.000
5			-0.158	-0.074	89.628	0.000
6			-0.107	0.038	92.218	0.000
7			-0.046	-0.015	92.704	0.000
8			-0.010	-0.008	92.729	0.000
9			0.028	0.031	92.915	0.000
10			0.052	0.009	93.535	0.000
11			0.097	0.084	95.710	0.000
12			0.107	0.017	98.373	0.000

Il ne s'agit pas d'une marche au hasard (les probabilités critiques de la Q-Stat sont toutes très largement inférieures à 0,05), le processus est à mémoire, il existe donc une représentation dans la classe des processus ARMA.

d) Recherche des ordres p et q de la représentation ARMA

Compte tenu de la forme des corrélogrammes simple et partiel nous sélectionnons un modèle ARMA(2, 0). Nous constatons que la constante est significativement différente de 0, et des essais d'autres représentations concurrentes (ARMA(1, 1)) indiquent des valeurs des critères d'information (AIC ou SC) supérieures au modèle retenu.

Dependent Variable: D(X1)
Included observations: 217 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.980952	0.665293	2.977562	0.0032
AR(1)	0.724525	0.066183	10.94725	0.0000
AR(2)	-0.237118	0.066464	-3.567593	0.0004
R-squared	0.380932	Akaike info criterion		6.079308
Durbin-Watson stat	2.017052	Schwarz criterion		6.126035
Inverted AR Roots	.36-.33i	.36+.33i		

e) Validation de la représentation

– Tests de Student sur les coefficients : les coefficients sont tous significativement différents de 0 (probabilités critiques inférieures à 0,05).

– Analyse des racines du polynôme opérateurs : les racines inverses sont égales à $\sqrt{0,36^2 + 0,33^2} = 0,488$, la racine du polynôme est donc $B = 1/0,488 = 2,5 > 1$, le processus est stationnaire. Il est nécessairement inversible car $q = 0$.

Analyse des résidus

Corrélogrammes du résidu et du résidu au carré

Sample: 4 221
Included observations: 217
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	1	-0.013	-0.013	0.0346	
2	1	0.017	0.017	0.0989	
3	1	0.009	0.009	0.1156	0.734
4	1	-0.012	-0.012	0.1438	0.938
5	1	-0.106	-0.106	2.6495	0.449
6	1	-0.016	-0.017	2.8972	0.610
7	1	-0.013	-0.009	2.7345	0.741
8	1	-0.023	-0.021	2.8556	0.627
9	1	0.018	0.015	2.9270	0.832
10	1	-0.001	-0.011	2.9272	0.939
11	1	0.051	0.047	3.5213	0.940
12	1	0.004	0.004	5.1626	0.800

Sample: 4 222
Included observations: 217
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	1	0.052	0.052	0.9533	
2	1	-0.034	-0.036	1.1121	
3	1	-0.022	-0.017	1.2142	0.271
4	1	0.056	0.066	2.1996	0.333
5	1	0.333	0.296	22.738	0.000
6	1	0.052	0.038	23.346	0.000
7	1	-0.056	-0.042	24.044	0.000
8	1	0.028	0.017	24.323	0.000
9	1	0.036	0.002	24.515	0.001
10	1	0.077	-0.022	25.886	0.001
11	1	-0.026	-0.050	26.016	0.002
12	1	0.019	0.055	26.036	0.004

Le corrélogramme du résidu indique qu'il s'agit d'un bruit blanc, le corrélogramme du résidu au carré indique un seul terme ($\rho_5 = 0,30$) de significativement différent de 0 ; il s'agit d'une anomalie statistique que ne remet pas en cause l'homoscédasticité des résidus. Les résidus sont donc un processus de bruit blanc.

Les résidus sont-ils gaussiens ? La statistique de Jarque-Bera ($JB = 0,32$) indique une probabilité critique de 0,85, nous acceptons l'hypothèse H_0 de normalité des résidus.

La représentation est validée x_{1t} est un processus ARIMA(2, 1, 0) avec constante.

f) Prédiction

La prédiction est donnée par :

Attention à la valeur de la constante $\mu = E(x_{1t}) \times (1 - \theta_1 - \theta_2)$

$$\Delta x_{1t} = \theta_1 \times \Delta x_{1t-1} + \theta_2 \times \Delta x_{1t-2} + E(x_{1t}) \times (1 - \theta_1 - \theta_2)$$

$$\Delta x_{1t} = 0,724 \times \Delta x_{1t-1} - 0,237 \times \Delta x_{1t-2} + 1,98 \times (1 - 0,724 + 0,237)$$

$$x_{1t} = x_{1t-1} + \Delta x_{1t}$$

$$\Delta x_{1,221} = 0,724 \times \Delta x_{1,220} - 0,237 \times \Delta x_{1,219} + 1,016$$

$$\Delta x_{1,221} = 0,724 \times 9,76 - 0,237 \times 10,43 + 1,016 = 5,62$$

$$x_{1,221} = x_{1,220} + \Delta x_{1,221} = 439,50 + 5,62 = 445,12$$

$$\Delta x_{1,222} = 0,724 \times \Delta x_{1,221} - 0,237 \times \Delta x_{1,220} + 1,016$$

$$\Delta x_{1,222} = 0,724 \times 5,62 - 0,237 \times 9,76 + 1,016 = 2,77$$

$$x_{1,222} = x_{1,221} + \Delta x_{1,222} = 445,12 + 2,77 = 447,88$$

$$\Delta x_{1,223} = 0,724 \times \Delta x_{1,222} - 0,237 \times \Delta x_{1,221} + 1,016$$

$$\Delta x_{1,223} = 0,724 \times 2,77 - 0,237 \times 5,62 + 1,016 = 1,69$$

$$x_{1,223} = x_{1,222} + \Delta x_{1,223} = 447,88 + 1,69 = 449,57$$

Les résultats sont présentés sur le tableau 7.6.

Tableau 7.6 – Résultats des prévisions calculées à un horizon de 5 périodes

t	x_{1t}	Δx_{1t}
217	416,25	
218	419,31	3,05
219	429,74	10,43
220	439,50	9,76
221	445,12	5,62
222	447,88	2,77
223	449,57	1,69
224	451,16	1,58
225	452,92	1,76

2) Analyse du processus x_2

La figure 7.7 présente l'évolution du processus x_2 et son corrélogramme, à leur lecture le processus semble non stationnaire.

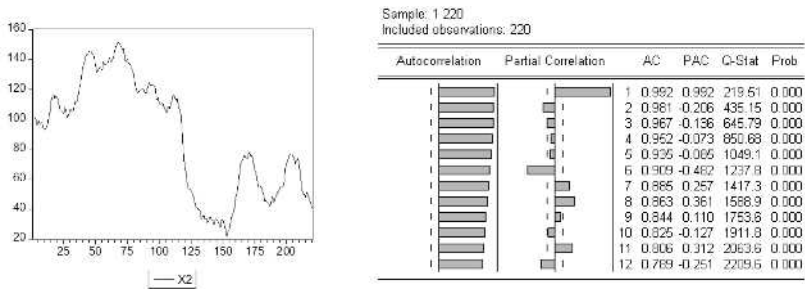


Figure 7.7 – Évolution et fonction d'autocorrélation simple du processus x_2

a) Tests de racines unitaires : tests de Phillips-Perron (fenêtre de Newey-West = 9)

Modèle [1] $t^* = -0,96$ $\alpha c = 0,30$ $\rightarrow H_0$, il existe une racine unitaire.

Modèle [2] $t^* = -0,94$ $\alpha c = 0,77$ $\rightarrow H_0$, il existe une racine unitaire.

Modèle [3] $t^* = -2,01$ $\alpha c = 0,58$ $\rightarrow H_0$, il existe une racine unitaire.

b) Stratégie de tests et tests d'hypothèses jointes

$F_3 = 1,73 < \phi_3 = 6,34 \rightarrow$ acceptation de H_0^3

$F_2 = 1,62 < \phi_2 = 4,75 \rightarrow$ acceptation de H_0^2

$F_1 = 0,71 < \phi_1 = 4,63 \rightarrow$ acceptation de H_0^1

Il s'agit donc d'un processus DS sans dérive : $x_{2t} = x_{2t-1} + a_t$

Nous le stationarisons par le passage aux différences premières.

c) Analyse des fonctions d'autocorrélation simple et partielle sur la série stationnaire

Sample: 1 220
Included observations: 219

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.449	0.449	44.850	0.000
		2 0.356	0.194	73.189	0.000
		3 0.308	0.120	94.482	0.000
		4 0.248	0.048	108.32	0.000
		5 0.190	0.007	116.51	0.000
		6 0.086	-0.084	118.20	0.000
		7 0.101	0.032	120.52	0.000
		8 0.001	-0.093	120.52	0.000
		9 -0.039	-0.054	120.87	0.000
		10 -0.001	0.048	120.87	0.000
		11 -0.111	-0.119	123.73	0.000
		12 -0.183	-0.132	131.54	0.000

Il ne s'agit pas d'une marche au hasard (les probabilités critiques de la Q-Stat sont toutes très largement inférieures à 0,05), le processus est à mémoire, il existe donc une représentation dans la classe des processus ARMA.

d) Recherche des ordres p et q de la représentation ARMA

Compte tenu de la forme des corrélogrammes simple et partiel nous sélectionnons un modèle ARMA(1, 1). Nous constatons que la constante n'est pas significativement différente de 0, et des essais d'autres représentations concurrentes (ARMA(2, 0) ou ARMA(0, 2)) indiquent des valeurs des critères d'information (AIC ou SC) supérieures au modèle retenu.

Dependent Variable: D(X2)				
Included observations: 218 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.806122	0.074302	10.84926	0.0000
MA(1)	-0.464814	0.110707	-4.198596	0.0000
R-squared	0.243089	Akaike info criterion		5.051100
Durbin-Watson stat	2.005677	Schwarz criterion		5.082150
Inverted AR Roots	.81			
Inverted MA Roots	.46			

e) Validation de la représentation

– Tests de Student sur les coefficients : les coefficients sont tous significativement différents de 0 (probabilités critiques inférieures à 0,05)

– Analyse des racines du polynôme opérateurs : les racines inverses sont inférieures à 1, les racines du polynôme sont donc pour la partie AR et MA supérieures à 1, le processus est stationnaire et inversible. Les deux racines sont différentes, il n'y a donc pas de racine commune.

Analyse des résidus

Corrélogrammes du résidu et du résidu au carré

Sample: 3 220
Included observations: 218
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	1	-0.005	-0.005	0.0062	
1	1	2	-0.014	-0.014	0.0510
1	1	3	0.032	0.032	0.2795
1	1	4	0.035	0.035	0.5521
1	1	5	0.031	0.032	0.7696
1	1	6	-0.069	-0.069	1.8347
1	1	7	0.063	0.061	2.7365
1	1	8	-0.045	-0.050	3.1960
1	1	9	0.055	-0.051	3.8806
1	1	10	0.083	0.083	5.4830
1	1	11	-0.066	-0.067	6.5023
1	1	12	-0.162	-0.165	12.643

Sample: 3 220
Included observations: 218
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
1	1	0.058	0.058	1.6965		
1	1	2	-0.057	-0.065	3.3966	
1	1	3	0.004	0.022	3.3980	0.056
1	1	4	-0.037	-0.049	3.7088	0.157
1	1	5	-0.050	-0.050	4.5217	0.210
1	1	6	-0.013	-0.011	4.5513	0.336
1	1	7	-0.103	-0.113	6.9874	0.223
1	1	8	0.006	0.025	6.9766	0.323
1	1	9	0.037	0.009	7.2937	0.389
1	1	10	-0.073	-0.079	6.5172	0.356
1	1	11	0.038	0.112	10.725	0.235
1	1	12	-0.042	-0.005	11.139	0.347

Le corrélogramme du résidu indique qu'il s'agit d'un bruit blanc, le corrélogramme du résidu au carré indique aucun terme de significativement différent de 0 ; les résidus sont donc homoscedastiques. Les résidus sont donc un bruit blanc.

Les résidus sont-ils gaussiens ? La statistique de Jarque-Bera (JB = 0,91) indique une probabilité critique de 0,63, nous acceptons l'hypothèse H_0 de normalité des résidus.

La représentation est validée x_{2t} est un processus ARIMA(1, 1, 1) sans constante.

f) Prévision

La prévision est donnée ($E(a_t) = 0$) par :

$$\Delta x_{2t} = \theta_1 \times \Delta x_{2t-1} + \phi_1 a_{t-1}$$

$$\Delta x_{2t} = 0,806 \times \Delta x_{2t-1} - 0,465 a_{t-1}$$

$$x_{2t} = x_{2t-1} + \Delta x_{2t}$$

$$\Delta x_{2,221} = 0,806 \times \Delta x_{2,220} - 0,465 \times a_{220}$$

$$\Delta x_{2,221} = 0,806 \times -3,37 - 0,465 \times -2,07 = -1,75$$

$$x_{2,221} = x_{2,220} + \Delta x_{2,221} = 40,81 - 1,75 = 39,06$$

$$\Delta x_{2,222} = 0,806 \times \Delta x_{2,221} - 0,465 \times a_{221}$$

$$\Delta x_{2,222} = 0,806 \times -1,75 - 0,465 \times 0 = -1,41$$

$$x_{2,222} = x_{2,221} + \Delta x_{2,222} = 39,06 + -1,41 = 37,64$$

$$\Delta x_{2,223} = 0,806 \times \Delta x_{2,222} - 0,465 \times a_{222}$$

$$\Delta x_{2,223} = 0,806 \times -1,41 - 0,465 \times 0 = -1,14$$

$$x_{2,223} = x_{2,222} + \Delta x_{2,223} = 37,64 - 1,14 = 36,50$$

Les résultats sont présentés sur le tableau 7.7.

Tableau 7.7 – Résultats des prévisions calculées à un horizon de 5 périodes

t	x_{2t}	Δx_{2t}	Résidu
217	49,75	- 1,51	- 1,95
218	46,08	- 3,67	- 3,36
219	44,18	- 1,90	- 0,50
220	40,81	- 3,37	- 2,07
221	39,06	- 1,75	0,00
222	37,64	- 1,41	0,00
223	36,50	- 1,14	0,00
224	35,59	- 0,92	0,00
225	34,84	- 0,74	0,00

3) Analyse du processus x_3

La figure 7.8 présente l'évolution du processus x_3 et son corrélogramme, à leur lecture le processus semble non stationnaire.

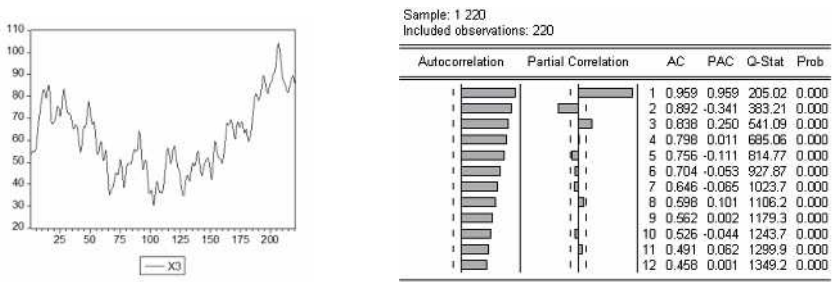


Figure 7.8 – Évolution et fonction d'autocorrélation simple du processus x_3

a) Tests de racines unitaires : tests de Phillips-Perron (fenêtre de Newey-West = 6)

Modèle [1] $t^* = 0,09$ $\alpha c = 0,71$ $\rightarrow H_0$, il existe une racine unitaire.

Modèle [2] $t^* = -1,66$ $\alpha c = 0,45$ $\rightarrow H_0$, il existe une racine unitaire.

Modèle [3] $t^* = -1,81$ $\alpha c = 0,70$ $\rightarrow H_0$, il existe une racine unitaire.

b) Stratégie de tests et tests d'hypothèses jointes

$F_3 = 1,21 < \phi_3 = 6,34 \rightarrow$ acceptation de H_0^3

$F_2 = 0,93 < \phi_2 = 4,75 \rightarrow$ acceptation de H_0^2

$F_1 = 1,00 < \phi_1 = 4,63 \rightarrow$ acceptation de H_0^1

Il s'agit donc d'un processus DS sans dérive : $x_{3t} = x_{3t-1} + a_t$

Nous le stationarisons par le passage aux différences premières.

c) Analyse des fonctions d'autocorrélation simple et partielle sur la série stationnaire

Sample: 1 220
Included observations: 219

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.525	0.525	61.249	0.000
		2 -0.179	-0.628	68.374	0.000
		3 -0.293	0.431	87.652	0.000
		4 0.015	-0.167	87.701	0.000
		5 0.155	-0.016	93.126	0.000
		6 -0.020	-0.086	93.214	0.000
		7 -0.180	0.008	100.61	0.000
		8 -0.105	-0.010	103.13	0.000
		9 0.031	-0.074	103.36	0.000
		10 0.054	0.080	104.03	0.000
		11 -0.044	-0.176	104.48	0.000
		12 -0.158	-0.067	110.34	0.000

Il ne s'agit pas d'une marche au hasard (les probabilités critiques de la Q-Stat sont toutes très largement inférieures à 0,05), le processus est à mémoire, il existe donc une représentation dans la classe des processus ARMA.

d) Recherche des ordres p et q de la représentation ARMA

Compte tenu de la forme des corrélogrammes simple et partiel nous sélectionnons un modèle ARMA(2, 1). Nous constatons que la constante n'est pas significativement différente de 0, et des essais d'autres représentations concurrentes (ARMA(2, 0), ARMA(0, 2), ARMA(1,1)) indiquent des valeurs des critères d'information (AIC ou SC) supérieures au modèle retenu.

Dependent Variable: D(X3)				
Included observations: 217 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.583539	0.070300	8.300700	0.0000
AR(2)	- 0.507840	0.066947	- 7.585706	0.0000
MA(1)	0.643773	0.064124	10.03949	0.0000
R-squared	0.667197	Akaike info criterion		4.294141
Durbin-Watson stat	1.900895	Schwarz criterion		4.340867
Inverted AR Roots	.29-.65i	.29+.65i		
Inverted MA Roots		-.64		

e) Validation de la représentation

– Tests de Student sur les coefficients : les coefficients sont tous significativement différents de 0 (probabilités critiques inférieures à 0,05)

– Analyse des racines du polynôme opérateurs : la racine inverse de la partie AR est égale à $\sqrt{0,29^2 + 0,65^2} = 0,71$, la racine du polynôme est donc $B = 1/0,71 = 1,41 > 1$, le processus est stationnaire. La racine inverse de la partie MA est inférieure à 1, la racine du polynôme est donc supérieure à 1, le processus est inversible. Les deux racines sont différentes, il n'y a donc pas de racine commune.

Analyse des résidus

Corrélogrammes du résidu et du résidu au carré

Sample: 4 220
Included observations: 217
Q-statistic probabilities adjusted for 3 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
1	1	1	0.042	0.042	0.3688	
1	1	2	0.023	0.027	0.5715	
1	1	3	-0.005	-0.008	0.5799	
1	1	4	0.093	0.090	2.3688	0.122
1	1	5	-0.003	0.010	2.3520	0.503
1	1	6	-0.045	-0.050	2.9465	0.416
1	1	7	0.169	0.195	9.3040	0.054
1	1	8	0.038	0.045	9.3020	0.087
1	1	9	-0.035	-0.030	9.3672	0.129
1	1	10	-0.049	-0.043	10.4530	0.165
1	1	11	0.024	0.031	10.568	0.228
1	1	12	-0.131	-0.150	14.551	0.104

Sample: 4 220
Included observations: 217
Q-statistic probabilities adjusted for 3 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
1	1	1	-0.049	-0.049	0.5232	
1	1	2	0.013	0.011	0.5626	
1	1	3	0.026	0.025	0.7036	
1	1	4	-0.069	-0.062	1.4811	0.224
1	1	5	-0.099	-0.105	3.6636	0.150
1	1	6	0.054	0.065	4.3091	0.290
1	1	7	0.009	0.001	4.2296	0.365
1	1	8	0.055	0.049	5.0437	0.411
1	1	9	0.015	0.026	5.0952	0.532
1	1	10	0.068	0.048	6.1476	0.522
1	1	11	0.090	0.092	8.0395	0.430
1	1	12	0.016	0.031	8.9447	0.678

Le corrélogramme du résidu indique qu'il s'agit d'une marche au hasard, le corrélogramme du résidu au carré indique aucun terme de significativement différent de 0 ; les résidus sont donc homoscedastiques. Les résidus sont donc un bruit blanc.

Les résidus sont-ils gaussiens ? La statistique de Jarque-Bera (JB = 0,878) indique une probabilité critique de 0,64, nous acceptons l'hypothèse H_0 de normalité des résidus.

La représentation est validée x_{3t} est un processus ARIMA(2, 1, 1) sans constante.

f) Prévision

La prévision est donnée ($E(a_t) = 0$) par :

$$\Delta x_{3t} = \theta_1 \times \Delta x_{3t-1} + \theta_2 \times \Delta x_{3t-2} + \phi_1 a_{t-1}$$

$$\Delta x_{3t} = 0,583 \times \Delta x_{3t-1} - 0,508 \times \Delta x_{3t-2} + 0,644 a_{t-1}$$

$$x_{3t} = x_{3t-1} + \Delta x_{3t}$$

$$\Delta x_{3,221} = 0,583 \times \Delta x_{3,220} - 0,508 \times \Delta x_{3,219} + 0,644 \times a_{220}$$

$$\Delta x_{3,221} = 0,583 \times -3,40 - 0,508 \times 0,89 + 0,644 \times -2,28 = -3,90$$

$$x_{3,221} = x_{3,220} + \Delta x_{3,221} = 85,99 - 3,90 = 82,09$$

$$\Delta x_{3,222} = 0,583 \times \Delta x_{3,221} - 0,508 \times \Delta x_{3,220} + 0,644 \times a_{221}$$

$$\Delta x_{3,222} = 0,583 \times -3,90 - 0,508 \times -3,40 + 0,644 \times 0 = -0,55$$

$$x_{3,222} = x_{3,221} + \Delta x_{3,222} = 82,09 - 0,55 = 81,54$$

$$\Delta x_{3,223} = 0,583 \times \Delta x_{3,222} - 0,508 \times \Delta x_{3,221} + 0,644 \times a_{222}$$

$$\Delta x_{3,223} = 0,583 \times -0,55 - 0,508 \times -3,90 + 0,644 \times 0 = 1,66$$

$$x_{3,223} = x_{3,222} + \Delta x_{3,223} = 81,54 + 1,66 = 83,19$$

Les résultats sont présentés sur le tableau 7.8.

Tableau 7.8 – Résultats des prévisions calculées à un horizon de 5 périodes

t	x_{3t}	Δx_{2t}	Résidu
217	86,59	2,00	- 1,16
218	88,49	1,90	2,87
219	89,38	0,89	- 1,05
220	85,99	- 3,40	- 2,28
221	82,09	- 3,90	0,00
222	81,54	- 0,55	0,00
223	83,19	1,66	0,00
224	84,44	1,25	0,00
225	84,33	- 0,11	0,00

8. *P*rocessus à mémoires longues et processus non linéaires

Dans les chapitres précédents l'hypothèse de linéarité est admise implicitement pour le processus générateur de la chronique qui obéit à un ARIMA. C'est pourquoi nous abordons à présent deux extensions importantes de l'analyse des processus aléatoires.

La première, la recherche de mémoire longue dans les processus concerne d'une part les modèles ARIMA fractionnaires, notés ARFIMA, pour lesquels d ou D , les degrés de différenciation des filtres ne sont pas des entiers et d'autre part les processus non linéaires chaotiques.

La seconde a trait à la modélisation non linéaire de type ARCH du phénomène d'hétéroscédasticité que l'on rencontre couramment dans les résidus de régression ou de processus ARMA qui modélisent les évolutions des prix de marchés financiers ou des produits de base par exemple.

I. *P*rocessus ARFIMA et processus chaotiques

Les processus ARMA étudiés jusqu'à présent (ou ARCH dans la section II.) sont des processus à mémoire courte dans le sens où l'effet d'un choc à un instant donné n'est pas durable et n'affecte donc pas l'évolution future de la chronique. Les processus à mémoire infinie comme les processus DS ont un comportement opposé : l'effet d'un choc est permanent et se répercute sur l'ensemble des valeurs futures de la série temporelle. Cette dichotomie s'avère insuffisante pour rendre compte de phénomènes à long terme comme l'ont montré les travaux de Hurst (1951) dans le domaine de l'hydrologie.

Les processus à mémoire longue, mais non infini, constituent un cas intermédiaire dans le sens où l'effet d'un choc à des conséquences durables pour des valeurs futures de la série mais celle-ci retrouvera son niveau d'équilibre « naturel » (Mignon V. 1998). Ce type de comportement a été formalisé par Mandelbrot et Wallis (1968) et par Mandelbrot et Van Ness (1968) par l'utilisation de mouvements browniens fractionnaires et des bruits gaussiens fractionnaires. À partir de ces travaux Granger et Joyeux (1980) et Hoskins (1981) définissent les processus ARIMA fractionnaires notés ARFIMA. Plus récemment ces processus ont été étendus au cas saisonnier (Ray 1993, Porter-Hudak 1990) et sont notés processus SARFIMA.

A. Les processus ARFIMA

1) Définitions

Rappelons (cf. chapitre 3) qu'un processus réel x_t de Wold : $x_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$

avec $\psi_0 = 1$, $\psi_j \in \mathbb{R}$ et a_t est *i.i.d.*(0, σ_a^2) est stationnaire sous la condition que

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty.$$

Le processus x_t stationnaire est à mémoire longue si $\sum_{j=0}^{\infty} |\psi_j^2| = \infty$.

Considérons un processus centré x_t , $t = 1, \dots, n$. On dit que x_t est un processus intégré stationnaire, noté ARFIMA(p, d, q) s'il s'écrit :

$$\phi_p(B)(1 - B)^d x_t = \theta_q(B)a_t$$

où :

- $\phi_p(B)$ et $\theta_q(B)$ sont respectivement les polynômes d'opérateurs en B des parties AR(p) et MA(q) du processus,

- a_t est *i.i.d.*(0, σ_a^2),
- $d \in \mathbb{R}$.

$(1 - B)^d$ porte le nom d'opérateur de différences fractionnaires et s'écrit à partir du développement en série :

$$\begin{aligned} (1 - B)^d &= \sum_{j=0}^{\infty} C_j^d (-B)^j \\ &= 1 - dB - \frac{d(1-d)}{2} B^2 - \dots - \frac{d(1-d)\dots(j-d-1)}{j!} B^j - \dots \\ &= \sum_{j=0}^{\infty} \pi_j B^j \end{aligned}$$

Avec $\pi_j = \frac{\Gamma(j-d)!}{\Gamma(j+1)\Gamma(-d)}$ $j = 0, 1, \dots$ et Γ est la fonction eulérienne de seconde espèce.

2) Processus FI(d)

Soit le processus ARFIMA(0, d , 0) : $(1 - B)^d x_t = a_t$ appelé aussi processus FI(d). C'est ce processus qui contient les composantes de long terme, la partie ARMA rassemblant les composants de court terme.

- Lorsque $d < 1/2$, le processus est stationnaire et possède la représentation moyenne mobile infinie : $x_t = (1 - B)^{-d} a_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} = \psi(B) a_t$ avec

$\psi_j = \frac{\Gamma(d+j)}{\Gamma(d)\Gamma(j+1)}$ où la fonction $\Gamma(h)$ est telle que :

$$\Gamma(h) = (h-1)! = \begin{cases} \int_0^{\infty} t^{h-1} e^{-t} dt & \text{si } h > 0 \\ \infty & \text{si } h = 0 \\ \frac{\Gamma(1+h)}{h} & \text{si } h < 0 \end{cases}$$

et $\Gamma(1/2) = \pi^{1/2}$

- Lorsque $d > -1/2$, le processus est inversible et possède la représentation autorégressive infinie :

$$(1 - B)^d x_t = \pi(B) x_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} = a_t \text{ avec } \pi_j = \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)}$$

Les valeurs asymptotiques des coefficients ψ_j et π_j :

$\lim_{j \rightarrow \infty} \psi_j \approx \frac{j^{d-1}}{\Gamma(d)}$ et $\lim_{j \rightarrow \infty} \pi_j \approx \frac{j^{-d-1}}{\pi(-d)}$ décroissent à une vitesse hyperbolique

qui est plus faible que la vitesse exponentielle des processus ARMA. La FAC possède ce même type de comportement ce qui permet de caractériser les processus FI(d).

- En définitive si (Hoskins 1981) :

$0 < d < 1/2$, le processus FI(d) est un processus à mémoire longue,

$d < 0$, le processus FI(d) est un processus anti-persistant,

$-1/2 < d < 0$, le processus FI(d) n'est pas à mémoire longue mais n'a pas le comportement des ARMA. Ce cas intermédiaire appelé anti-persistence par Mandelbrot correspond à des alternances de hausses et de baisses dans le processus. Ce comportement est aussi appelé « effet Joseph » par référence à la bible.

Pourquoi effet « Joseph » ?

Dans l'ancien testament, Joseph fils d'Israël a été vendu par ses frères puis emmené en Égypte. Le Pharaon fit venir Joseph afin de lui expliquer le songe des sept années de vaches grasses et des sept années de vaches maigres. Initié grâce à Joseph, le Pharaon reconnu l'esprit de Dieu et le nomma gouverneur de sa maison.

C'est ce « hasard lent » ou anti-persistence qui est appelé l'effet Joseph : 7 ans de sécheresse, 7 ans de fertilité (alternances de hausses et de baisses dans le processus).

L'effet Noé, appelé ainsi par Mandelbrot, correspond à une discontinuité brutale comme le déluge (un « *crash* » boursier par exemple).

Le processus $FI(d)$ est donc stationnaire et inversible pour $-1/2 < d < 1/2$.

3) Processus ARFIMA

Les processus ARFIMA(p, d, q) s'écrivent : $\phi_p(B)(1 - B)^d x_t = \theta_q(B)a_t$ où a_t est *i.i.d.*($0, \sigma_a^2$) ou encore $(1 - B)^d x_t = u_t$ avec u_t le processus ARMA(p, q) : $\phi_p(B)u_t = \theta_q(B)a_t$.

Il est stationnaire si :

- $d < 1/2$,
- les racines $\phi_p(B)$ sont à l'extérieur du cercle unité du plan complexe,

Il est inversible si :

- $d > -1/2$,
- les racines $\theta_q(B)$ sont à l'extérieur du cercle unité du plan complexe.

La fonction d'autocorrélation des ARFIMA possède le comportement de décroissance hyperbolique mentionné pour les processus $FI(d)$. On peut vérifier qu'au-delà d'un certain retard l'influence de la mémoire courte (due à la partie ARMA) s'estompe pour ne laisser subsister que l'influence de la partie intégrée fractionnaire.

Le spectre du processus ARFIMA(p, d, q) est : $f_x(\omega) = \left[2 \sin\left(\frac{\omega}{2}\right) \right]^{-2d} \times f_u(\omega)$ avec $\omega \in [0, \pi]$ et où $f_u(\omega)$ est la densité spectrale d'une composante u_t de mémoire courte comme un processus ARMA.

On peut démontrer si ω est petit et si u_t est un ARMA que :

$$f_x(\omega) \underset{\omega \rightarrow 0}{\approx} \frac{\sigma^2}{2\pi} \left[\frac{1 - \theta_1 - \dots - \theta_q}{1 - \phi_1 - \dots - \phi_p} \right]^2 \omega^{-2d} \approx C \omega^{-2d}, C \text{ est une constante positive}$$

qui dépend de la variance du bruit blanc et du processus ARMA. La densité spectrale d'un ARFIMA a la forme d'une hyperbole à l'image d'un processus non stationnaire, ce qui peut engendrer des confusions lors des interprétations.

Dans les travaux récents le principe d'intégration fractionnaire est étendu pour mêler dépendance de long terme et périodicité. Deux voies sont envisagées selon que la fréquence du cycle est connue ou non.

- Si elle est connue (série mensuelle, journalière...) Hassler (1993) propose de recourir aux processus ARFIMA périodiques, notées PARFIMA, qui sont des extensions des processus SARFIMA (ARFIMA saisonnier).

- Si elle est inconnue (lors de la recherche d'un cycle conjoncturel par exemple) Gray, Zang et Woodward (1989) suggèrent un recours aux processus ARMA généralisés, notés GARMA, développés par Chung (1996). Ces modèles permettent d'estimer conjointement la périodicité et le niveau de la structure de long terme.

4) Estimation

L'estimation d'un modèle ARFIMA peut être réalisée à partir de deux familles de méthodes : les méthodes dites « en deux étapes » et celles du maximum de vraisemblance. Nous ne présentons ici que les premières ¹.

Les méthodes en deux étapes consistent à estimer (étape 1) le coefficient d'intégration fractionnaire puis (étape 2) – sur la série différenciée fractionnaire – les coefficients de la composante de court terme.

La technique la plus utilisée est celle de l'estimation de Geweke et Porter-Hudak (1983). Après transformation du spectre du processus ARFIMA et à partir du logarithme de son périodogramme évalué à la fréquence $\omega_j = \frac{2\pi_j}{n}$ on obtient la relation ² :

$$\text{Ln} [I_x(\omega_j)] = \text{Ln} [f_u(0)] - d \text{Ln} \left[4 \sin^2 \left(\frac{\omega_j}{2} \right) \right] + \text{Ln} \left[\frac{I_x(\omega_j)}{f_x(\omega_j)} \right]$$

où $I_x(\omega_j) = \rho_{0,x} + 2 \sum_{\theta=1}^{n-1} \rho_{\theta,x} \cos(\omega \theta)$ est le périodogramme de x_t (cf. chapitre 4).

On ne conserve que les premières fréquences $\omega_j, j = 1, \dots, m$ où m est une fonction du nombre d'observations n telle que $\lim_{n \rightarrow \infty} m(n) = 0$; $\lim_{n \rightarrow \infty} \frac{m}{n} = 0$. En général $m = n^\alpha$ avec $0 < \alpha < 1$ (le plus souvent ($\alpha = 0,5$)).

On pose :

$$Y_j = \text{Ln} [I_x(\omega_j)] ; X_j = \text{Ln} \left[4 \sin^2 \left(\frac{\omega_j}{2} \right) \right] ; e_j = \text{Ln} \left[\frac{I_x(\omega_j)}{f_x(\omega_j)} \right] + \gamma ;$$

1. Pour les méthodes de maximum de vraisemblance cf., par exemple, Brockwell-Davis (1987), Sowell (1992), Chung (1996).
 2. Hassler – 1993 page 373.

$a = \text{Ln}[f_u(0)] - \gamma$ et $\gamma = 0,5772$ (constante d'Euler).

On obtient donc l'équation :

$$Y_j = a + dX_j + e_j$$

que l'on estime par les MCO. L'estimateur de Geweke et Porter-Hudak s'écrit :

$$d^{ghp} = \frac{\sum_{j=1}^m (X_j - \bar{X})a_j}{\sum_{j=1}^m (X_j - \bar{X})^2}$$

et la variance asymptotique de l'estimateur est :

$$\text{Var}(d^{ghp}) = \frac{\pi^2}{6} \left[\sum_{j=1}^m (X_j - \bar{X})^2 \right]^{-1}$$

Les auteurs démontrent alors que : $\frac{d^{ghp} - d}{[\text{Var}(d^{ghp})]^{1/2}} \rightarrow N(0,1)$ si

$$\lim_{n \rightarrow \infty} \frac{m}{n} \rightarrow 0 \text{ et } \lim_{n \rightarrow \infty} \frac{m}{\text{Ln}^2(n)} \rightarrow \infty.$$

Cet estimateur peut être calculé en utilisant des fenêtres spectrales (Chen, Abraham, Peiris (1994)) dans le périodogramme. Par ailleurs, Hurvich et Ray (1995) proposent une fenêtre spectrale particulière qui permet d'estimer d pour des cas non stationnaires ($1/2 < d < 1$) et/ou non inversibles ($d < -1/2$).

5) Tests de mémoire longue

a) La statistique « Rescaled Range »¹

La statistique R/S a été présentée en 1951 dans une étude sur les débits du Nil par l'hydrologue Harold Edwin Hurst. Son but est de rechercher l'intensité d'une composante cyclique apériodique dans une chronique qui est un des aspects de la dépendance à long terme (mémoire longue) développé par Mandelbrot.

Soit x_t une chronique réalisation d'un processus aléatoire stationnaire avec $t = 1, \dots, n$ et $x_t^* = \sum_{u=1}^t x_u$ la chronique cumulée. La statistique R/S notée Q_n est l'étendue R_n des sommes partielles des écarts de la série à sa moyenne divisée par son écart type S_n :

1. On peut consulter pour ce paragraphe Mignon (1998).

$$Q_n = \frac{R_n}{S_n} = \frac{\max_{1 \leq k \leq n} \sum_{j=1}^k (x_j - \bar{x}_n) - \min_{1 \leq k \leq n} \sum_{j=1}^k (x_j - \bar{x}_n)}{\left[\frac{1}{n} \sum_{j=1}^n (x_j - \bar{x}_n)^2 \right]^{1/2}}$$

Le premier terme du numérateur est le maximum sur k des sommes partielles des k écarts de x_j par rapport à sa moyenne. Ce terme (en max) est donc toujours positif ou nul. Par définition le second terme (en min) est toujours négatif ou nul. Par conséquent R_n est toujours positif ou nul.

La statistique Q_n est toujours non négative.

La statistique H de Hurst appliquée à une chronique x_t est fondée sur le découpage du temps en intervalles de longueur d , pour d donné on obtient $(T + 1)$ tronçons du temps. La statistique H est calculée sur chaque tronçon (Mandelbrot) en appliquant la méthode précédente de Hurst en tenant compte du décalage opéré sur l'échelle du temps. Dans ce cas :

$R_n(t, d) = \max_{0 \leq u \leq d} [\Delta(u)] - \min_{0 \leq u \leq d} [\Delta(u)]$ où $[\Delta(u)]$ est l'interpolation linéaire de

$x_t^* = \sum_{s=1}^t x_s$ entre t et $t + d$; soit $\Delta(u) = [x_{t+u}^* - x_t^*] - \frac{u}{d} [x_{t+d}^* - x_t^*]$ il s'agit

de l'expression ramenée au décalage d de $\sum_{j=1}^k (x_j - \bar{x}_n)$ utilisée pour le calcul de R_n .

L'écart type s'écrit alors :

$$S_n(t, d) = \left[\frac{1}{d} \left[\sum_{0 \leq u \leq d} x_{t+u}^2 \right] - \frac{1}{d^2} \left[\sum_{0 \leq u \leq d} x_{t+u} \right]^2 \right]^{1/2}$$

On peut donc calculer R_n/S_n pour chacun des $(T + 1)$ tronçons de longueur d ainsi que leur moyenne arithmétique. On peut aussi démontrer (Mandelbrot-Wallis) que l'intensité de la dépendance à long terme est donnée par le coefficient H compris entre zéro et un de la relation :

$$Q_n = \frac{R_n}{S_n} \underset{d \rightarrow \infty}{\approx} cd^H.$$

Soit $\log(R_n/S_n) = \log c + H \log d$,

H est l'estimateur par les MCO de cette relation linéaire.

Dans la pratique on construit M échantillons fictifs en choisissant M points de départ arbitraires de la série. Un point de départ est donné par :

$$t = \frac{n(j-1)}{M} + 1$$

et la longueur du $j^{\text{ème}}$ échantillon est :

$$l = n \frac{(M-j-1)}{M}.$$

L'ajustement linéaire du nuage ¹ obtenu conduit à l'estimation de l'exposant de Hurst.

Le r^2 du nuage dépend du décalage initial choisit. On retient (Hernad, Mouillard, Strauss-Khan, 1978, 1979) pour H celui qui fournit le r^2 maximum pour un décalage initial d .

L'interprétation des valeurs de H est la suivante :

Si $0 < H < 1/2 \implies$ processus anti-persistant,

Si $H = 1/2 \implies$ processus purement aléatoire ou processus ARMA. Il y a absence de dépendance de long terme (mémoire courte).

Si $1/2 < H < 1 \implies$ processus à mémoire longue, la dépendance est d'autant plus forte que H tend vers 1.

b) La statistique de Lo

La statistique de l'exposant de Hurst ne peut être testée car elle est trop sensible à la dépendance de court terme. Lo (1991) montre que l'analyse proposée par Mandelbrot peut conclure à la présence d'une mémoire longue alors que la série ne présente qu'une dépendance de court terme. En effet, dans ce cas, l'exposant de Hurst par l'analyse R/S est biaisé à la hausse. Lo propose donc une nouvelle statistique R/S modifiée notée :

$$\tilde{Q}_n = \frac{R_n}{\hat{\sigma}_n(q)}$$

$$\tilde{Q}_n = \frac{\max_{1 \leq k \leq n} \sum_{j=1}^k (x_j - \bar{x}_n) - \min_{1 \leq k \leq n} \sum_{j=1}^k (x_j - \bar{x}_n)}{\left[\frac{1}{n} \sum_{j=1}^n (x_j - \bar{x}_n)^2 + \frac{2}{n} \sum_{j=1}^q \omega_j(q) \left(\sum_{i=j+1}^n (x_j - \bar{x}_n) (x_{i-j} - \bar{x}_n) \right) \right]^{1/2}}$$

Cette statistique diffère de la précédente Q_n par son dénominateur qui prend en compte non seulement les variances des termes individuels mais aussi les autocovariances pondérées en fonction des décalages q avec :

$$\omega_j(q) = 1 - \frac{j}{q+1} \text{ où } q < n.$$

1. Pour lequel les premiers points sont supprimés (phase transitoire).

Andrews-Lo (1991) proposent la règle suivante pour le choix de q :

$$q = [k_n] = \text{partie entière de } k_n \text{ avec } k_n = \left[\left(\frac{3n}{2} \right)^{1/3} \left(\frac{2\hat{\rho}}{1-\hat{\rho}} \right)^{2/3} \right] \hat{\rho} \text{ est l'esti-}$$

mation du coefficient d'autocorrélation d'ordre 1 et dans ce cas $\omega_j = 1 - \left| \frac{j}{k_n} \right|$.

Lo démontre que sous l'hypothèse $H_0 : x_t \implies i.i.d.(0, \sigma_x^2)$ et pour n tendant vers l'infini, la distribution asymptotique de \tilde{Q}_n converge faiblement vers une variable $V : \frac{1}{\sqrt{n}} \tilde{Q}_n \longrightarrow V$ où V est le rang d'un pont Brownien, processus à accroissements indépendants gaussien contraint à l'unité et pour lequel $H = 1/2$.

La distribution de la variable aléatoire V est donnée par :

$$F_V(v) = 1 + 2 \sum_{k=1}^{\infty} (1 - 4k^2 v^2) e^{-2(kv)^2}$$

Les valeurs critiques de cette distribution symétrique les plus couramment utilisées sont :

$P(V < v)$	0,005	0,025	0,05	0,1	0,2	0,3	0,4	0,5	0,543	0,6	0,7	0,8	0,9	0,95	0,975	0,995
v	0,721	0,809	0,861	0,927	1,018	1,09	1,157	1,223	$\sqrt{\pi/2}$	1,294	1,374	1,473	1,620	1,747	1,862	2,098

Le calcul de H s'effectue comme précédemment et Lo analyse le comportement de \tilde{Q}_n sous des alternatives de dépendance à long terme. Il démontre alors que :

$$V = \frac{1}{\sqrt{n}} \tilde{Q}_n \xrightarrow{P} \begin{cases} \infty & \text{pour } H \in [0,5; 1] \\ 0 & \text{pour } H \in [0; 0,5] \end{cases}$$

Soit l'hypothèse H_0 : il existe une mémoire courte dans la série ($H = 0,5$). Pour un seuil d'acceptation à 5 % H_0 est acceptée si $v \in [0,809; 1,862]$.

On peut vérifier qu'il existe une relation entre les valeurs d des processus ARFIMA et la valeur H de l'exposant de Hurst ($d = H - 0,5$).

En 1996 Moody et Wu proposent une statistique modifiée notée R/S^* . Ils montrent, en effet, que la statistique de Lo peut être biaisée pour des petits échantillons. Ils proposent un estimateur non biaisé du dénominateur de la statistique notée S^* .

Exercice n° 8.1

 fichier C5EX3

Simulation et calcul des statistiques de Hurst, Lo et de GPH

On demande :

a) de simuler un bruit blanc de 1 000 observations comprises entre (-100) et $(+ 100)$ et de calculer les statistiques de Hurst et de Lo.

b) de simuler une marche au hasard $x_t = x_1 + \sum_{i=1}^{1000} a_i$ avec $a_i \implies i.i.d.(0, 1)$ et $x_1 = 1$ sur 1 000 observations et de calculer les statistiques de Hurst et de Lo,

c) de calculer les statistiques de Hurst et de Lo sur les données du CAC40 ($n = 1\,109$) étudié au chapitre 5,

d) enfin, sur cette même série du CAC40, d'utiliser la méthode GPH et le maximum de vraisemblance afin d'estimer l'ordre d du processus générateur $FI(d)$ de la chronique brute et en différences premières.

Solutions

a) Pour la statistique de Hurst nous avons retenu : un décalage initial de 20, 10 échantillons et un décalage seuil de 200, ce qui permet d'interpoler H sur les trente valeurs les plus significatives soit 36 % des observations.

Tableau 8.1 – Résultats

Décalage seuil	50	100	200	230
H	0,203	0,302	0,481	0,503
R^2	0,629	0,761	<u>0,863</u>	0,853

Pour la statistique de Lo on dispose de 3 échantillons et l'interpolation est réalisée sur 25 % des estimations.

Tableau 8.2 – Résultats

Décalage seuil	50	110	150	210
H	0,241	0,384	0,428	0,503
R^2	0,679	0,883	<u>0,897</u>	0,871
ν	0,893	1,03	1,11	1,07

Quelle que soit la méthode, les valeurs les plus fiables de H sont celles pour lesquelles R^2 tend vers un. C'est le cas pour le décalage compris entre 200 et 300 (Hurst) et 150 et 210 (Lo). On constate que H tend vers 0,5 conformément à la théorie.

Pour des valeurs de H tendant vers 0,5, la variable ν doit être comprise entre 0,861 et 1,620 pour accepter l'hypothèse nulle de mémoire courte : c'est le cas avec cet exercice. La méthode de Hurst évalue la mémoire à environ 200 fois le pas de temps alors que celle de Lo ne l'estime qu'à 150.

b) Pour la statistique de Hurst nous avons choisi un décalage initial de 25 et 10 échantillons. Pour l'interpolation, nous avons retenu les décalages supérieurs à 120 soit 72 estimations sur 220 (32 %).

Tableau 8.3 – Résultats

Décalage seuil	50	70	110	200
H	0,683	0,817	0,889	0,903
R^2	0,914	0,975	<u>0,986</u>	0,970

Le calcul de la statistique de Lo s'effectue dans les mêmes conditions mais avec un seul échantillon.

Tableau 8.4 – Résultats

Décalage seuil	50	100	150	210
H	0,789	0,885	0,903	0,892
R^2	0,978	<u>0,993</u>	0,989	0,986
ν	6,92	7,52	7,62	7,57

La valeur la plus fiable de H par la méthode de Hurst est 0,889 qui permet donc de conclure à la présence d'une mémoire longue. Celle donnée par Lo se trouve pour un décalage compris entre 100 et 150 soit environ $H = 0,9$. La valeur de ν supérieure à 1,620 confirme la structure de dépendance de long terme.

Ces calculs obtenus à partir d'une série non stationnaire montrent l'interprétation erronée que l'on peut effectuer avec ces tests.

c) Nous avons vu (cf. chapitre 5) que le processus générateur du CAC40 contient une racine unitaire. La statistique H de Hurst calculée sur la série brute est de l'ordre de 0,9 pour un décalage seuil compris entre 50 et 60 et possède une valeur équivalente pour un décalage de 40 avec la statistique de Lo ($\nu = 7,14$ supérieure à 1,62).

Nous pourrions en déduire l'existence d'une dépendance positive comprise entre 40 et 70 valeurs. En réalité lorsque le processus générateur est stationnarisé par le passage aux différences premières, les statistiques de Hurst et de Lo sont respectivement de 0,46 et 0,45 et la valeur du coefficient ν est de 1,294 inférieure à 1,620. Nous pouvons alors conclure qu'il n'existe pas de dépendance à long terme dans la série CAC40 en différences premières et que les résultats issus de la série brute ne sont pas conformes aux hypothèses d'application des tests.

d) Les calculs sont réalisés avec le logiciel Gauss et TSM sous Gauss. Les résultats sont les suivants.

Tableau 8.5 – Estimation GPH

	Série brute	Série différenciée	Écart type (série différenciée)
Pas de fenêtre	1,130	- 0,096	0,161
Rectangulaire	1,079	- 0,166	0,178
Bartlett	1,048	- 0,126	0,103
Daniell	1,043	- 0,126	0,126
Tukey	1,060	- 0,139	0,112
Parzen	1,0602	- 0,123	0,092
Bartlett– Priestley	1,067	- 0,151	0,138

Estimation par le maximum de vraisemblance

Série brute

Nombre d'observations = 1 109

Nombre de paramètres estimés : 2

Valeur de la fonction de vraisemblance = - 5 202,185

Paramètre	Estimation	Écart type	t-statistique	Probabilité critique
<i>d</i>	1,155	0,019	62,28	0,000
Sigma	26,413	0,0561	47,10	0,000

Série en différences premières

Nombre d'observations = 1 108

Nombre de paramètres estimés : 2

Valeur de la fonction de vraisemblance = - 4 950,552

Paramètre	Estimation	Écart type	t-statistique	Probabilité critique
<i>d</i>	0,034	0,029	1,199	0,231
Sigma	21,097	0,448	47,08	0,000

Ces résultats montrent que la série possède une racine unitaire puisque, avec ou sans fenêtre, l'estimateur GPH est proche de 1 ainsi que celui du maximum de vraisemblance, ce qui confirme les résultats obtenus au chapitre 5.

Lorsque la série est différenciée, afin de la stationnariser conformément à la théorie, l'hypothèse H_0 de nullité du coefficient d'intégration fractionnaire est acceptée dans les deux cas :

- estimation GPH $\left(\frac{d^{sph}}{Std - error} < 1,96 \right)$,
- maximum de vraisemblance (cf. les probabilités critiques).

Enfin, à partir de la relation $d = H - 0,5$ (H = statistique de Hurst), nous pouvons vérifier qu'elle conduit à un résultat contradictoire pour la série brute ($d \approx 0,4$ par la relation et $d \approx 1$ par le calcul). Pour la série différenciée, on obtient $d \approx -0,037$ à partir du H de Hurst et $d \approx 0,05$ par la statistique de Lo. Ces résultats sont conformes aux estimations.

B. Les processus chaotiques

Les processus et les tests présentés au paragraphe précédent ne remettent pas en cause la linéarité d'une possible structure dans une chronique réalisation d'un processus stationnaire (ou intégré stationnaire).

La recherche de modèles structurants non linéaires de type déterministe ou aléatoire dans les séries brutes (ou dans les résidus d'une modélisation) peut être réalisée à partir de la théorie du Chaos et plus précisément de la technique de la dimension de corrélation de Grassberger et Procaccia de 1983, du test de Brock, Dechert et Scheinkman (BDS) de 1987 et des exposants de Lyapunov.

La statistique de test BDS est fondée sur l'intégrale de corrélation. Elle permet de tester l'hypothèse nulle qu'une chronique est indépendante statistiquement et identiquement distribué (*i.i.d.*) contre diverses hypothèses alternatives non spécifiées du type structure linéaire ou non linéaire.

1) La dimension de corrélation

Soit x_t , une série d'observations représentant l'état d'un système régi par une équation du type $x_t = f(x_{t-1}, x_{t-2}, \dots, x_{t-n})$ où n est le nombre d'observations et f est la loi inconnue de marche des variables d'état inconnues.

L'ensemble des n coordonnées caractérisant un système de ce type à un instant donné est considéré comme les coordonnées d'un point dans un espace abstrait de dimension n appelé espace des phases.

Le plongement de la série x_t est la substitution de l'espace des observations par l'espace reconstruit à partir des dérivées successives du signal x_t . On étudie les propriétés du système dynamique dans l'espace des phases en faisant croître la dimension de plongement notée m . Ce plongement est possible grâce au théorème de plongement démontré par Takens (1981).

La dimension de corrélation a été introduite par Grassberger et Procaccia (deux physiciens) en 1983. Elle mesure la probabilité pour que deux points de la chronique dans l'espace des phases que l'on reconstruit avec la dimension m du plongement soient proches selon une distance inférieure à une quantité ε donnée. Si on appelle x_t la chronique ($t = 1, \dots, n$), la formule de l'intégrale de corrélation est alors :

$$c(\varepsilon; m; n) = \left(\frac{1}{n_m(n_m - 1)} \right) \sum_{i=1}^{n_m} \sum_{j=1}^{n_m} I(x_i^m; x_j^m; \varepsilon)$$

où : $n_m = n - m + 1$ $i \neq j$

avec : m la dimension de plongement et $m > 1$,

ε un paramètre choisi représentant la distance maximale entre deux paires de points,

$$I(x_i^m; x_j^m; \varepsilon) \text{ une fonction indicatrice } \begin{cases} = 1 & \text{si } \|x_i^m - x_j^m\| \leq \varepsilon \\ = 0 & \text{sinon} \end{cases}.$$

Pour des petites valeurs de ε Grassberger et Procaccia (1983) montrent que $c(\varepsilon; m; n)$ croît de manière exponentielle soit :

$$d_m = \lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \left[\frac{\text{Ln } c(\varepsilon, m, n)}{\text{Ln } (\varepsilon)} \right] \quad (\text{Ln} = \text{Logarithme népérien}).$$

Alors $\text{Ln } c(\varepsilon, m, n) \approx d_m \text{Ln } \varepsilon \rightarrow \text{Ln } c(\varepsilon, m, n) \approx \text{Ln } \varepsilon^{d_m}$ et $c(\varepsilon, m, n) \approx \varepsilon^{d_m}$

L'intégrale $c(\varepsilon; m; n)$ évolue donc comme ε^{d_m} . On distingue la dimension de corrélation \widehat{d}_c de la dimension de plongement m . Quand la dimension de plongement augmente on obtient pour chaque m la dimension d_m . Lors de ces calculs d_m converge vers une valeur \widehat{d}_c , appelée dimension de corrélation :

$$\widehat{d}_c = \lim_{m \rightarrow \infty} d_m \quad \text{ou} \quad \widehat{d}_c = \lim_{m \rightarrow \infty} \left[\frac{d \text{Ln } c(\varepsilon, m, n)}{d \text{Ln } (\varepsilon)} \right],$$

si x_t est un bruit blanc, alors $\widehat{d}_c = m$, car \widehat{d}_c ne converge pas.

En pratique, on procède de la façon suivante :

- on choisit des m_i croissants $i = 1, \dots, 20$.
- pour chaque m , on calcule la dimension de corrélation, c'est-à-dire $c(\varepsilon, m, n)$,
- on estime par les moindres carrés ordinaires le modèle :

$$\text{Ln } c(\varepsilon, m, n) = k + p_i \text{Ln } \varepsilon_j + v_i$$

où k est une constante et p_i représente l'estimation de la dimension de corrélation d_c ,

$v \rightarrow i.i.d.$ est fixé et j varie. On obtient ainsi plusieurs valeurs de d_c .

p_i est l'élasticité instantanée entre la dimension de corrélation et la distance entre deux points.

Si la courbe des estimateurs ainsi obtenus évolue linéairement avec m , on en conclut que x_t est un processus stochastique.

Si la courbe sature au-delà d'une valeur notée m_s , cela signifie qu'une structure sous-jacente chaotique déterministe existe dans la chronique x_t .

2) Test BDS et exposants de Lyapunov

Les trois auteurs (Brock, Dechert et Scheinkman) ont développé une théorie statistique pour l'intégrale de corrélation reprise en 1996 par ces mêmes auteurs associés à Lebaron. Ils en déduisent un test d'hypothèse nulle d'une série *i.i.d.* contre une hypothèse non spécifiée.

Si x_t est une chronique ($t = 1, \dots, n$) et les $x_t^m = (x_t, x_{t-1}, \dots, x_{t+m-1})$ les m historiques, la statistique BDS s'écrit : $B(\varepsilon, m, n) = c(\varepsilon, m, n) - [c(\varepsilon, m, n)]^m$

On démontre que si $B(\varepsilon, m, n) = 0$ et si $n \rightarrow \infty$ alors x_t est *i.i.d.*

On utilise alors la statistique : $W(\varepsilon, m, n) = n^{1/2} \frac{B(\varepsilon, m, n)}{\sigma_n(\varepsilon, n)}$ qui sous l'hypothèse que x_t est *i.i.d.* tend en distribution, si $n \rightarrow \infty$, vers une loi normale d'écart type $\sigma_n(\varepsilon, n)$.

Comme on peut montrer que sous l'hypothèse nulle d'une série *i.i.d.* $c(\varepsilon, m, n) \rightarrow [c(\varepsilon, 1, n)]^m$ alors W peut s'écrire sous la forme :

$$W(\varepsilon, m, n) = n^{1/2} \frac{(c(\varepsilon, m, n) - [c(\varepsilon, 1, n)]^m)}{\sigma_m(\varepsilon, n)}$$

qui obéit à une loi normale centrée réduite. ε et m sont étroitement liés. Pour m donné, ε ne doit pas être trop élevé (respectivement petit). Nous pouvons choisir $\varepsilon = 0,5 \times$ l'écart type estimé de x_t et pour m la valeur 5 afin que $\frac{n}{m} > 200$.

Si $|W| > 1,96$ (valeur de la loi normale pour un test bilatéral à 5 %), l'hypothèse nulle d'indépendance est rejetée.

Un rejet de l'hypothèse nulle d'une série *i.i.d.* peut provenir d'une série non stationnaire, d'une structure de dépendance linéaire ou encore d'une structure de dépendance non linéaire (déterministe ou chaotique). Par conséquent, le test BDS n'est pas un test direct de non linéarité, ni un test de chaos : on ne connaît pas la distribution de cette statistique sous d'autres hypothèses nulles que celle d'indépendance. On peut cependant l'employer comme tests de non linéarité en l'appliquant à une série filtrée de toute forme de dépendance linéaire.

On complète souvent l'utilisation de la dimension de corrélation, qui teste la convergence globale, par le calcul des exposants de Lyapunov qui permettent de déterminer les propriétés de stabilité des systèmes dynamiques non linéaires. Ils s'écrivent : $\lambda_i = \lim_{n \rightarrow 0} \left[\frac{1}{n} \log_2 (\lambda_i^{(n)}) \right]$ avec $\lambda_i^{(n)}$ les valeurs propres de la matrice Jacobienne évaluée en x_0 du système dynamique.

Ils mesurent le taux moyen de divergence ou de convergence de deux trajectoires d'un système initialement proches. Si le plus grand exposant de Lyapunov est négatif, cela signifie qu'il y a une convergence vers un état stable (l'attracteur). S'il est positif, cela signifie qu'il y a divergence des trajectoires dans l'espace des phases et donc présence de chaos dans la chronique.

On confronte alors les résultats de la dimension de corrélation et des exposants de Lyapunov dans un test de validation de Brock appelé aussi test résiduel.

Nous pouvons résumer l'utilisation de ces tests par l'organigramme de la figure 8.1.

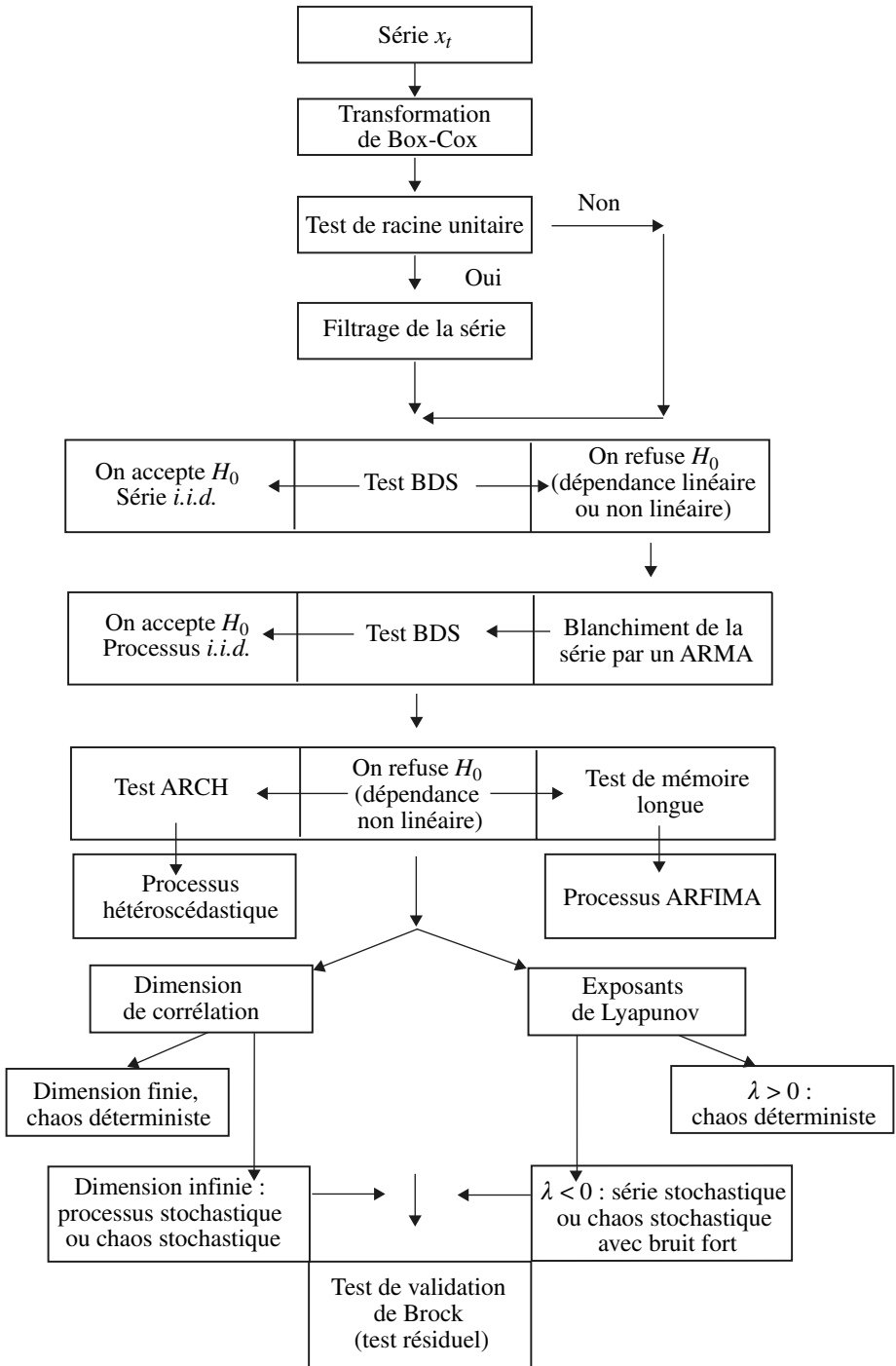


Figure 8.1 – Organigramme combinant les différents tests

Exercice n° 8.2

 fichier C5EX3

Applications des tests BDS, dimension de corrélation et exposant de Lyapunov

a) À partir des données du CAC40 ($n = 1\,109$) stationnarisées par les différences premières $y_t = x_t - x_{t-1}$, on demande d'appliquer le test BDS, de calculer la dimension de corrélation (d_c) et l'exposant de Lyapunov (λ_i) sur la série y_t .

b) On demande d'effectuer les mêmes tests et calculs sur une série de rentabilité boursière ($n = 919$). La série est transformée selon les différences premières des logarithmes naturels (après avoir vérifié la présence d'une racine unitaire sur le logarithme de la chronique) : $y_t = \text{Ln} \frac{x_t}{x_{t-1}}$

Solution

a) Chronique du CAC40

Tableau 8.6 – Test BDS sur y_t

ε/σ Dimension de plongement	0,5	1	1,5	2
$m = 2$	1,9268	2,7421*	4,2576*	5,5503*
$m = 3$	1,9979*	2,7075*	4,2351*	5,7090*
$m = 4$	3,1894*	3,7816*	5,1695*	6,5805*
$m = 5$	4,7690*	4,7868*	5,7564*	6,9341*

ε : distance maximale de l'intégrale de corrélation,

σ : écart type de la série,

* : valeur de $w(\varepsilon, m) > 1,96$

On constate qu'à l'exception d'une seule, toutes les valeurs de $w(\varepsilon, m)$ sont supérieures à 1,96. On rejette donc l'hypothèse H_0 : le processus est *i.i.d.* Il existe donc une dépendance linéaire ou non linéaire.

Tableau 8.7 – Dimension de corrélation (D_c) et exposant de Lyapunov (λ_i)

m	d_c	λ_i
2	2,040	0,883
3	3,056	0,645
4	3,887	0,382
5	4,507	0,231
6	5,175	0,142
7	5,464	0,116
8	5,940	0,099
9	6,408	0,097
10	6,723	0,090

La dimension de corrélation ne converge pas, elle tend vers l'infini avec l'accroissement de m : le processus est donc stochastique.

Les exposants de Lyapunov convergent lentement vers une valeur positive. Il semble que le processus est chaotique bien que les valeurs de l'exposant de Lyapunov puissent être sur estimées du fait du faible nombre d'observations et du bruit dans les données.

La série filtrée du CAC40 est un processus stochastique non linéaire (on pourrait vérifier le caractère hétéroscédastique de ce processus).

b) Chronique de la rentabilité boursière :

Tableau 8.8 – Test BDS sur y_t

ε/σ Dimension de plongement	0,25	0,5	0,75	1
$m = 2$	2,4872*	2,0436*	2,1326*	2,2928*
$m = 3$	2,3607*	2,3716*	0,2981	2,0443*
$m = 4$	-2,4726*	6,9219*	5,3687*	6,7409*
$m = 5$	-1,3995	-1,9611*	-2,5647*	7,6870*

ε : distance maximale de l'intégrale de corrélation,

σ : écart type de la série,

* : valeur de $w(\varepsilon, m) > 1,96$

Une majorité de valeurs de $w(\varepsilon, m)$ sont supérieures à 1,96, on rejette donc l'hypothèse H_0 : le processus est *i.i.d.* Il existe donc une possibilité de dépendance linéaire ou non linéaire.

Tableau 8.9 – Dimension de corrélation (D_c) et exposant de Lyapunov (λ_i)

m	d_c	λ_i
2	1,824	0,834
3	2,889	0,633
4	3,913	0,327
5	4,619	0,205
6	5,237	0,155
7	5,600	0,120
8	6,009	0,101
9	6,281	0,087
10	6,634	0,071

La dimension de corrélation ne converge pas, elle tend vers l'infini avec l'accroissement de m : le processus est donc stochastique.

Les exposants de Lyapunov convergent vers une valeur positive, le processus semble donc être chaotique.

On effectue le test résiduel de Brock en estimant un processus AR(1) sur la chronique y_t :

$$y_t = 0,1743y_{t-1} + e_t$$

et on utilise la série des résidus (e_t) pour effectuer le test BDS de non linéarité.

Tableau 8.10 – Test BDS sur e_t

ε/σ Dimension de plongement	0,5	0,75	1	1,5	2
$m = 2$	6,0481*	4,3600*	4,4630*	2,9715*	2,2030*
$m = 3$	- 0,3153	9,205*	7,173*	1,2079	2,1733*
$m = 4$	- 8,7150*	46,4*	15,959*	- 0,9226	3,9138*
$m = 5$	- 5,4407*	- 6,4092*	- 6,9434*	5,9458*	3,377*

La série des résidus n'obéit pas à un processus *i.i.d.* La série des rentabilités boursières présente une structure non linéaire, probablement stochastique.

Tableau 8.11 – Dimension de corrélation (D_c) et exposant de Lyapunov sur e_t

m	d_c	λ_i
2	1,976	0,852
3	2,912	0,598
4	3,970	0,343
5	4,574	0,219
6	5,325	0,154
7	5,591	0,110
8	6,113	0,087
9	6,339	0,074
10	6,722	0,063

La dimension de corrélation est infinie et les exposants de Lyapunov sont positifs. Les résultats sont donc identiques à ceux obtenus sur la série y_t .

La chronique y_t est donc un processus stochastique non linéaire.

II. Les modèles ARCH : présentation générale

Les modèles classiques de prévision fondés sur les modèles ARMA supposent des séries temporelles à variance constante (hypothèse d'homoscédasticité). Cette modélisation néglige donc, éventuellement, l'information contenue dans le facteur résiduel de la chronique.

Les modèles de type ARCH (*AutoRegressive Conditional Heteroscedasticity*) permettent de modéliser des chroniques (la plupart du temps financières ¹)

1. En effet, les séries financières sont particulièrement concernées par les modèles ARCH, car on constate des périodes de forte spéculation (variabilité élevée) suivie de périodes d'accalmie (variabilité faible).

qui ont une volatilité (ou variance ou variabilité) instantanée qui dépend du passé. Il est ainsi possible d'élaborer une prévision dynamique de la chronique en termes de moyenne et de variance.

Présentés initialement par Engle (1982), ces modèles ont connu des développements et des applications très importants par la suite. Dans ce chapitre nous allons aborder les différentes classes de modèles ARCH, GARCH (*Generalized AutoRegressive Conditionnal Heteroscedasticity*) et ARCH-M (*AutoRegressive Conditional Heteroscedasticity-Mean*), les tests statistiques permettant de les repérer, puis les méthodes d'estimation et de prévision.

L'étude des séries temporelles financières s'est donc trouvée confrontée à deux types de problèmes :

- la non stationnarité des séries ;
- le caractère leptokurtique ¹ de la distribution des données.

Soit, par exemple, le modèle de régression linéaire : $y = Xa + \varepsilon$

avec

y : vecteur de dimension n des observations de la variable à expliquer,

X : matrice de dimension $n \times (k + 1)$ des données des variables explicatives,

a : vecteur de dimension $k + 1$ des paramètres du modèle,

ε : erreur de spécifications.

Supposons que le modèle soit à erreurs autocorrélées de type AR(1) :

$\varepsilon_t = \rho\varepsilon_{t-1} + u_t = \sum_{n=0}^{\infty} \rho^n u_{t-n}$ avec $u_t \rightarrow i.i.d.(0; \sigma_u^2)$. Comme $E(\varepsilon_t) = 0$ par

hypothèse, on obtient alors $V(\varepsilon_t) = \frac{\sigma_u^2}{1 - \rho^2}$

L'espérance et la variance conditionnelles ² du modèle sont égales à :

$$E(\varepsilon_t / \varepsilon_{t-1}) = \rho\varepsilon_{t-1}$$

$$\text{Var}(\varepsilon_t / \varepsilon_{t-1}) = \sigma_u^2$$

1. Les queues de probabilité sont plus épaisses que celles d'une loi normale aux extrémités, les valeurs anormales sont donc plus fréquentes.

2. Rappel : $E(X/Y) = \sum_x xp(x/y)$ (cas discret) et $E(X/Y) = \int xf(x/y)dx$ (cas continu)

$$E(X/Y) = E(X)$$

Si X et Y sont indépendants :

$$\text{Var}(X/Y) = E(X^2/Y) - E^2(X/Y)$$

Par exemple concernant un cours de bourse : si on a $E(\text{cours}_{t+1} / \text{cours}_t) = \text{cours}_t$. On en déduit que le meilleur prédicateur du cours du jour est le cours de la veille. Le marché est parfaitement efficient.

Posons que : $\varepsilon_t = u_t \times z_{t-1}$ avec $u_t \rightarrow i.i.d.(0 ; \sigma_u^2)$ et z_{t-1} une variable aléatoire, à variations stochastiques indépendantes de u_t .

$E(\varepsilon_t) = 0$ et $V(\varepsilon_t) = \sigma_u^2 z_{t-1}^2$, le modèle est donc par construction hétéroscédastique.

Dans ce cas, $V(\varepsilon_t/\varepsilon_{t-1}) = \sigma_u^2 z_{t-1}^2$. On constate que les propriétés stochastiques des estimateurs conditionnels ou non sont les mêmes, ce qui limite leur utilisation dans un cadre de prévision, notamment dans le cas des séries financières, car l'effet spéculatif ne peut pas être pris en compte. On préfère dans ce cas utiliser les modèles ARCH.

A. Modèle de régression de type ARCH

1) Spécification du modèle

Soit un modèle $AR(p)$ à erreurs ARCH : $\phi_p(B)x_t = \varepsilon_t$ (ou un modèle de régression $y = Xa + \varepsilon$) avec $\varepsilon_t = u_t \times h_t$ où $u_t \rightarrow N(0, 1)$ (on remplace z_{t-1} par h_t

dans l'écriture précédente) et $h_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 = \alpha_0 + \alpha(B) \varepsilon_t^2$ tel que $\alpha_0 > 0, \alpha_i \geq 0 \quad \forall i$ et $\alpha(B) = \alpha_1 B + \alpha_2 B^2 + \dots + \alpha_p B^p$; h_t^2 est appelé processus ARCH d'ordre p et il est noté $ARCH(p)$.

Le modèle AR (ou de régression) est dit modèle AR à erreurs $ARCH(p)$.

Les espérances et les variances conditionnelles sont égales, alors que les variances sont différentes :

$$E(\varepsilon_t) = E(u_t)E(h_t) = 0$$

$$V(\varepsilon_t) = \sigma_\varepsilon^2$$

$$E(\varepsilon_t/I_{t-1}) = E(u_t/I_{t-1})E(h_t/I_{t-1}) = 0$$

I_t constitue l'ensemble de l'information par rapport à laquelle la condition est définie. En termes conditionnels, il vient : $\text{Var}(\varepsilon_t/I_{t-1}) = h_t^2$ (quantité variable dans le temps).

Le modèle reste correct sur sa trajectoire – en moyenne – car l'espérance conditionnelle est toujours nulle. On peut vérifier que la variance conditionnelle est finie si $\sum_{i=1}^p \alpha_i < 1$.

L'intérêt de cette formulation réside dans l'interdépendance d'une variable endogène au modèle. Pour Engle, cette notion est très importante en finance car le risque d'erreur n'est pas le même selon les périodes t : il y a alternance de périodes d'accalmie et de périodes d'euphorie.

2) Propriétés d'un modèle ARCH(1)

Soit le modèle AR(p) à erreurs ARCH(1) : $\phi_p(B)x_t = \varepsilon_t$ (ou le modèle $y = Xa + \varepsilon$) avec $\varepsilon_t = u_t \times h_t$ où $u_t \rightarrow N(0; 1)$ et $h_t = \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$
soit $\varepsilon_t = u_t \times \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$ où $\alpha_0 > 0$ et $0 \leq \alpha_1 < 1$.

La moyenne non conditionnelle est donnée par :

$$E(\varepsilon_t) = E \left[u_t \times (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^{1/2} \right] = E(u_t) \times E \left[(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^{1/2} \right] = 0$$

puisque par hypothèse nous avons $E(u_t) = 0$.

De même $E(\varepsilon_t \varepsilon_{t-i}) = 0$ puisque $E(u_t u_{t-i}) = 0$ pour $i \neq 0$.

La variance (non conditionnelle) de ε_t est donnée par :

$$E(\varepsilon_t^2) = E \left[u_t^2 \times (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2) \right] = E(u_t^2) \times E \left[(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2) \right]$$

Or $\sigma_u^2 = 1$ et la variance de ε_t est égale à celle de ε_{t-i} . La variance (non conditionnelle) de ε_t est donc égale à $E(\varepsilon_t^2) = \alpha_0 / (1 - \alpha_1)$.

La moyenne et la variance non conditionnelles ne sont donc pas fonctions du temps.

La moyenne conditionnelle est donnée par :

$$E(\varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) = E(u_t) \times E(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^{1/2} = 0$$

La variance conditionnelle est : $E(\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$ (puisque $\sigma_u^2 = 1$)

Ainsi dans une spécification de type ARCH seule la variance conditionnelle dépend des erreurs passées.

Kurtosis d'un processus ARCH

Le Kurtosis est le rapport du moment centré d'ordre 4 sur le carré du moment centré d'ordre 2. Pour un processus ARCH(1), il vient :

$$k = \frac{E(\varepsilon_t^4)}{E^2(\varepsilon_t^2)} = \frac{3(1 - \alpha_1^2)}{1 - 3\alpha_1^2}, k \text{ est toujours supérieur à 3, un processus ARCH a}$$

donc une distribution leptokurtique qui permet de modéliser les phénomènes rares.

Exercice n° 8.3

Simulation d'une erreur ARCH

On demande de simuler le processus ARCH suivant : $\varepsilon_t = u_t \times \sqrt{1 + 0,8\varepsilon_{t-1}^2}$

Solutions

Programme Eviews : (C8EX3.PRG) en téléchargement

```
CREATE U 200
' Initialisation des séries
GENR U = NRND
GENR H = 0
' Génération dynamique de H
smpl 2 200
GENR H = 1+0.8*U(-1)*U(-1)*H(-1)
GENR EPS = sqr(H)*U
smpl 1 200
GENR Y = 0
' Génération dynamique de Y
smpl 2 200
GENR Y = 0.8*Y(-1)+EPS
smpl 1 200
```

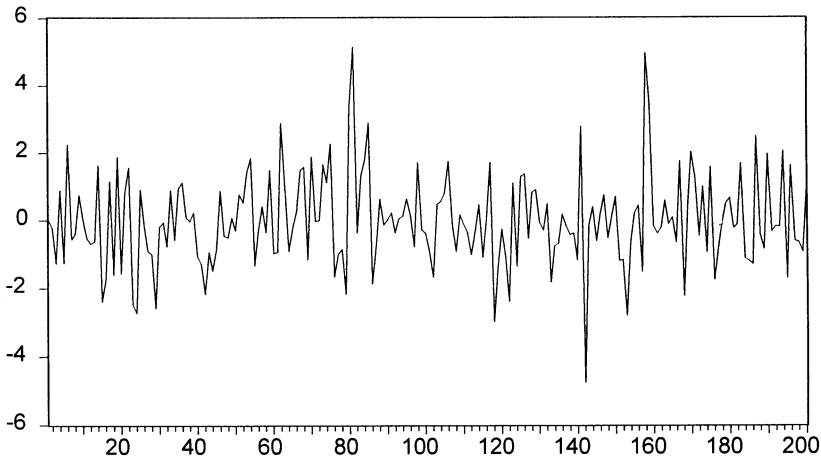


Figure 8.2 – Simulation de l'erreur $\varepsilon_t = u_t \times \sqrt{1 + 0,8\varepsilon_{t-1}^2}$

Nous observons sur la figure 8.2 et la figure 8.3 que la variabilité de l'erreur et du processus y_t n'est plus constante avec le temps. Elle connaît des phases de stabilité suivies de périodes de variance plus élevée (particulièrement vers les observations 80 et 140 à 160).

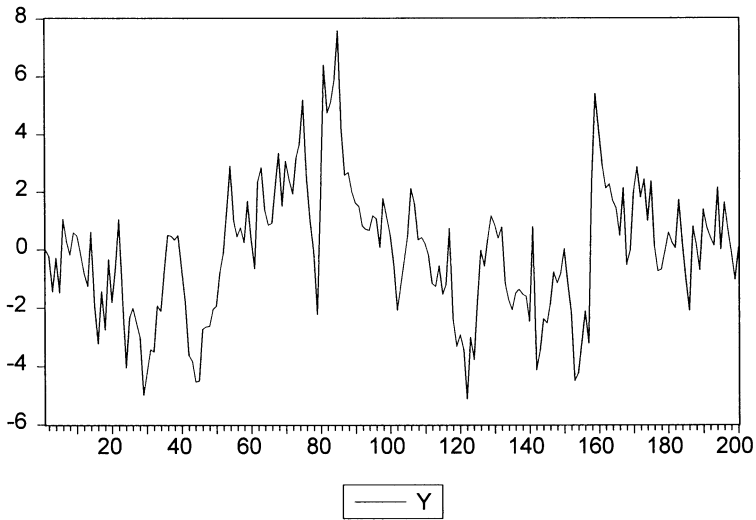


Figure 8.3 – Simulation du processus $y_t = 0,8y_{t-1} + \epsilon_t$

B. Test d'un modèle de type ARCH

Soit un modèle de régression $y = Xa + \epsilon$ (ou autorégressif) avec une spécification de type ARCH pour l'erreur ϵ_t telle que $\epsilon_t = u_t \times h_t$ avec $u_t \rightarrow N(0; 1)$

$$\text{et } h_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \epsilon_{t-i}^2 = \alpha_0 + \alpha(B) \epsilon_t^2$$

avec $\alpha(B) = \alpha_1 B + \alpha_2 B^2 + \dots + \alpha_p B^p$.

Soit l'hypothèse emboîtée $H_0, H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_p = 0$;
contre l'hypothèse alternative, $H_1 : \alpha_i$ non tous nuls.

Si l'hypothèse H_0 est acceptée, la variance de l'erreur est constante $\sigma_t^2 = \alpha_0$. Dans le cas contraire les termes de l'erreur suivent un ARCH dont l'ordre p est à déterminer.

Le test est fondé soit sur un test de Fisher classique, soit sur le test du multiplicateur de Lagrange (LM) :

De manière pratique, on procède de la manière suivante :

1^{re} étape : Calcul de e_t et le résidu du modèle de régression (ou d'un modèle ARMA).

2^e étape : Calcul des e_t^2

3^e étape : Régression autorégressive des résidus sur p retards (résidu décalé)

où seuls les retards significatifs sont conservés, $e_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i e_{t-i}^2$

4^e étape : Calcul de la statistique du multiplicateur de Lagrange, $LM = n \times R^2$ avec :

n = nombre d'observations servant au calcul de la régression de l'étape 3,
 R^2 = coefficient de détermination de l'étape 3.

Si $LM > \chi^2(p)$ à p degrés de liberté lu dans la table à un seuil α fixé (en général 0,05), on rejette H_0 ; on considère que le processus est justifiable d'un modèle ARCH(p).

C'est le test de significativité des coefficients α_i de la régression e_t^2 sur e_{t-p}^2 , qui permet de déterminer l'ordre p du processus ARCH¹ sachant qu'un processus ARCH d'ordre 3 semble un maximum, au-delà, le modèle sera justifiable d'un processus de type GARCH (*Generalized ARCH*) présenté ci-après.

Une autre approche consiste à calculer le corrélogramme des résidus aux carrés du modèle initial. Si des termes de ce corrélogramme sont significativement différents de 0, alors on peut conclure à une spécification de type ARCH ; on utilise pour cela la statistique Q de Ljung-Box vue précédemment.

C. P Procédure d'estimation et prévision

La technique du maximum de vraisemblance peut être utilisée pour estimer les coefficients du modèle ARCH. La log-vraisemblance conditionnelle à la date t est donnée par :

$$l_t = \text{Cte} - \frac{1}{2} \log h_t^2 - \frac{1}{2} \varepsilon_t^2 \cdot h_t^{-2}$$

et la log-vraisemblance totale par : $L = \text{Cte} - \frac{1}{2} \sum_{t=1}^{t=n} \log (h_t^2) - \frac{1}{2} \sum_{t=1}^{t=n} (\varepsilon_t^2 \cdot h_t^{-2})$

Le test précédent permet de déterminer l'ordre p du processus ARCH et donc d'estimer les paramètres de la régression (ou de l'ARMA) et α .

Nous pouvons recourir aussi à la méthode des MCG réalisables (moindres carrés pondérés), les différentes étapes sont les suivantes :

1^{re} étape : Estimation du modèle ARMA ou de régression $y = Xa + \varepsilon$

2^e étape : À partir du résidu e_t , on effectue une régression par les MCO ($e_t = h_t, e_t$ joue le rôle de h_t)

1. On peut très bien rencontrer un processus ARCH « troué ».

$$e_t^2 = \widehat{\alpha}_0 + \sum_{i=1}^p \widehat{\alpha}_i e_{t-i}^2 + u_t$$

3^e étape : On considère que l'on peut approcher la variance de l'erreur ε_t par h_t^2 .

On estime de nouveau les paramètres du modèle par les MCG :

$$\widehat{a} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} Y$$

Ce qui équivaut à faire une régression pondérée avec comme facteur de pondération $\omega = \frac{1}{h_t}$ et $\Omega = \text{diag}(h_t^2)$.

4^e étape : Estimation des coefficients α par MCG

$$h_t^2 = \widehat{\alpha}_0 + \sum_{i=1}^p \widehat{\alpha}_i e_{t-i}^2 + u_t \quad \text{avec } \widehat{\alpha} = (e' \Omega^{-1} e)^{-1} (e' \Omega^{-1} h)$$

La différence de spécification d'un modèle ARCH par rapport au modèle linéaire général réside dans la variance de l'erreur du modèle :

- dans le modèle linéaire, il est fonction de la variance résiduelle, $\widehat{\sigma}_\varepsilon^2$.
- dans le modèle ARCH(p), il va être fonction de h_t^2 .

En d'autres termes la variance de l'erreur n'est plus constante et de ce fait l'intervalle de confiance de la prévision est fonction de la volatilité de la série elle-même.

Exercice n° 8.4

 **fichier C8EX4**

Étude d'un processus ARCH

Soit un processus y_t , on demande d'en étudier les propriétés et d'estimer les paramètres du modèle par une méthode adéquate.

Solutions

Cet exercice est résolu à partir d'Eviews.

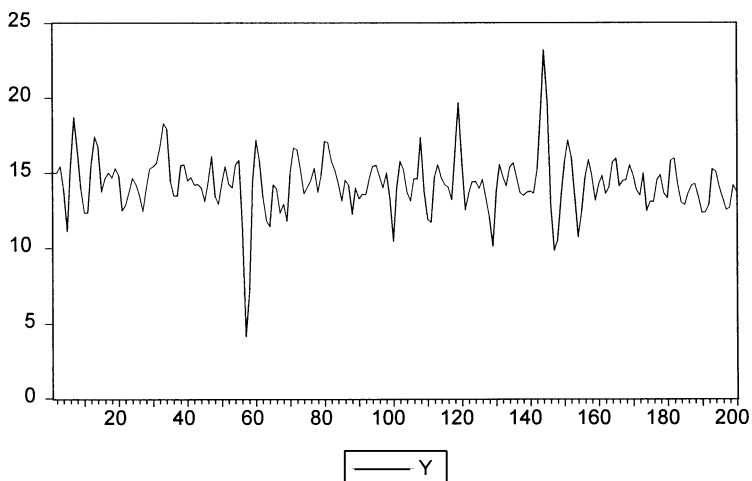


Figure 8.4 – Représentation du processus y_t

Étude du corrélogramme

Sample: 1 200
Included observations: 200

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.535	0.535	58.049	0.000
		2	-0.122	-0.571	61.069	0.000
		3	-0.351	0.117	86.332	0.000
		4	-0.174	0.001	92.551	0.000
		5	0.054	-0.020	93.159	0.000
		6	0.143	0.057	97.407	0.000
		7	0.086	-0.020	98.968	0.000
		8	-0.014	-0.005	99.010	0.000
		9	-0.055	0.020	99.655	0.000
		10	-0.033	-0.014	99.887	0.000

L'examen du corrélogramme laisse présager d'un modèle autorégressif d'ordre 2 de type AR(2). L'estimation des paramètres conduit aux résultats suivants :

Dependent Variable: Y				
Sample (adjusted): 3 200				
Included observations: 198 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	14.28529	0.131266	108.8268	0.0000
AR(1)	0.839452	0.058878	14.25747	0.0000
AR(2)	-0.569230	0.058859	-9.671123	0.0000

Le corrélogramme des résidus carrés est le suivant :

Sample: 3 200
 Included observations: 198
 Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.234	0.234	11.017	
		2	-0.050	-0.111	11.519	
		3	-0.048	-0.009	11.981	0.001
		4	-0.083	-0.080	13.403	0.001
		5	-0.083	-0.052	14.824	0.002
		6	-0.088	-0.073	16.418	0.003
		7	-0.046	-0.023	16.849	0.005
		8	-0.041	-0.050	17.207	0.009
		9	0.065	0.074	18.092	0.012
		10	0.094	0.040	19.960	0.010

Les premiers termes sont significativement différents de 0, on corrobore ce test par celui du multiplicateur de Lagrange. Ce test de spécification de type ARCH(1) conduit aux résultats :

ARCH Test :

F-statistic	11.30457	Probability	0.000930
Obs*R-squared	10.79472	Probability	0.001018

Test Equation :

LS // Dependent Variable is RESID^2
 Sample(adjusted) : 4 200
 Included observations : 197 after adjusting endpoints

Variable	CoefficienStd.	Error-Statistic	Prob.	
C	1.374943	0.334909	4.105421	0.0001
RESID^2(-1)	0.234064	0.069616	3.362226	0.0009
R-squared	0.054796	Mean dependent var	1.793670	
Log likelihood	-568.7646	F-statistic	11.30457	
Durbin-Watson stat	1.948383	Prob(F-statistic)	0.000930	

À la lecture des résultats, le F du Fisher empirique et la statistique LM du multiplicateur de Lagrange laissent supposer une spécification de type ARCH(1). Les deux probabilités critiques sont inférieures à 0,05, nous sommes donc amenés à rejeter l'hypothèse H_0 : F de nullité des coefficients α_i . Le test d'un ARCH(2) doit alors être effectué :

ARCH Test :

F-statistic	6.849449	Probability	0.001337
Obs*R-squared	12.98983	Probability	0.001511

Test Equation :

LS // Dependent Variable is RESID^2
Sample(adjusted) : 5 200
Included observations : 196 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.531062	0.349598	4.379489	0.0000
RESID^2(-1)	0.259801	0.071538	3.631640	0.0004
RESID^2(-1)	- 0.110554	0.071559	- 1.544921	0.1240

Le coefficient de l'ordre 2 de la variable RESID^2(-2) n'est pas significativement différent de 0, la spécification retenue est donc de type ARCH(1). L'estimation des paramètres peut alors être effectuée à l'aide de la méthode du maximum de vraisemblance (instruction EvIEWS) : ARCH(1,0) Y C AR(1) AR(2)

Dependent Variable: Y				
Method: ML - ARCH (Marquardt) - Normal distribution				
Sample (adjusted): 3 200				
Included observations: 198 after adjustments				
Variance backcast: ON				
GARCH = C(4) + C(5)*RESID(-1)^2				
	Coefficient	Std. Error	z-Statistic	Prob.
C	14.24263	0.088669	160.6274	0.0000
AR(1)	0.720230	0.058788	12.25135	0.0000
AR(2)	- 0.466271	0.061722	- 7.554369	0.0000
Variance Equation				
C	0.834466	0.113545	7.349206	0.0000
RESID(-1)^2	0.577660	0.143681	4.020423	0.0001
R-squared	0.505514	Durbin-Watson stat		1.642802
Inverted AR Roots	.36+.58i			.36-.58i

Le terme constant du processus est égal à :

$$\hat{a}_0 = \hat{b}_0 \times (1 - \hat{\theta}_1 - \hat{\theta}_2) = 14,24 \times (1 - 0,72 + 0,466) = 10,62$$

Le module de la racine du polynôme opérateur retard est égal à

$$\frac{1}{\sqrt{0,36^2 + 0,58^2}} = 1,46 > 1, \text{ le processus est stationnaire.}$$

Le modèle estimé s'écrit donc : $y_t = 10,6 + 0,72y_{t-1} - 0,47y_{t-2} + e_t$ dont les erreurs théoriques suivent un ARCH(1) : $\varepsilon_t = u_t \times h_t$ avec $u_t \rightarrow N(0; 1)$ et $h_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$ estimé par : $\hat{h}_t^2 = 0,83 + 0,58 \varepsilon_{t-1}^2$.

D. Processus de type GARCH

1) Spécification

Le modèle GARCH¹ est une généralisation (*Generalized*), due à Bollerslev (1986), des modèles de type ARCH. La spécification est la suivante :

$$y = Xa + \varepsilon$$

avec $\varepsilon_t = u_t \times h_t$, $u_t \rightarrow N(0; 1)$ et $h_t^2 = \alpha_0 + \alpha(D)\varepsilon_t^2 + \beta(D)h_t^2$

$$h_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2$$

qui est l'écriture d'un GARCH(p, q).

Remarques :

- Si $q = 0$ on a un GARCH(p, q) = GARCH($p, 0$) = ARCH(p) et si $q = 0$ et $p = 0$ alors $\varepsilon_t \rightarrow n.i.d.$
- Un processus de type GARCH(p, q) est équivalent à un processus de type ARCH(∞) ce que l'on peut démontrer par récurrence (en remplaçant h_t^2 par h_{t-1}^2 etc.). Cette équivalence permet de déterminer les conditions de stationnarité d'un processus de type GARCH : $\alpha(1) + \beta(1) < 1$.

Les processus GARCH sont similaires aux processus ARMA usuels dans le sens où le degré q apparaît comme le degré de la partie de la moyenne mobile et p comme celui de l'autorégressive ; cela permet d'introduire des effets d'innovation. La variance conditionnelle est déterminée par le carré des p erreurs précédentes et des q variances conditionnelles passées.

2) Test et estimation de modèle de type GARCH

a) Test d'un modèle de type GARCH

Dans le cas d'une hétéroscédasticité conditionnelle supposée, on ne peut tester une spécification de type ARCH que contre une spécification de type GARCH.

Le test porte sur l'hypothèse nulle H_0 d'une erreur ARCH(p) contre l'hypothèse H_1 d'une erreur GARCH(p, q). On va donc tester l'hypothèse H_0 que les β_j sont nuls.

$H_0 : \beta_j = 0, j = 1, \dots, q$; contre l'hypothèse H_1 : Il existe au moins un β_j non nul.

Le test le plus approprié est celui du multiplicateur de Lagrange : $nR^2 \rightarrow \chi^2(q)$ ($q =$ degré de liberté) où R^2 est le coefficient de détermination

obtenu dans la régression par les MCO dans l'équation $h_t^2 = \hat{\alpha}_0 + \sum_{i=1}^p \hat{\alpha}_i \varepsilon_{t-i}^2$

$$+ \sum_{j=1}^q \hat{\beta}_j h_{t-j}^2$$

1. Bollerslev T. (1988).

Si $nR^2 > \chi^2(q)$ lu dans la table à un seuil de confiance (en général 0,05) et q degrés de liberté alors on rejette l'hypothèse H_0 . Les erreurs obéissent à un processus GARCH(p, q).

b) Estimation d'un modèle de type GARCH

La procédure d'identification des ordres p et q du GARCH est la même que pour l'identification des ordres d'un processus ARMA. L'estimation des paramètres du modèle peut se faire à l'aide de l'algorithme de Berndt, Hall ou bien celui de Hall et Hausman (1974) fondé sur une méthode de maximum de vraisemblance.

Exercice n° 8.5

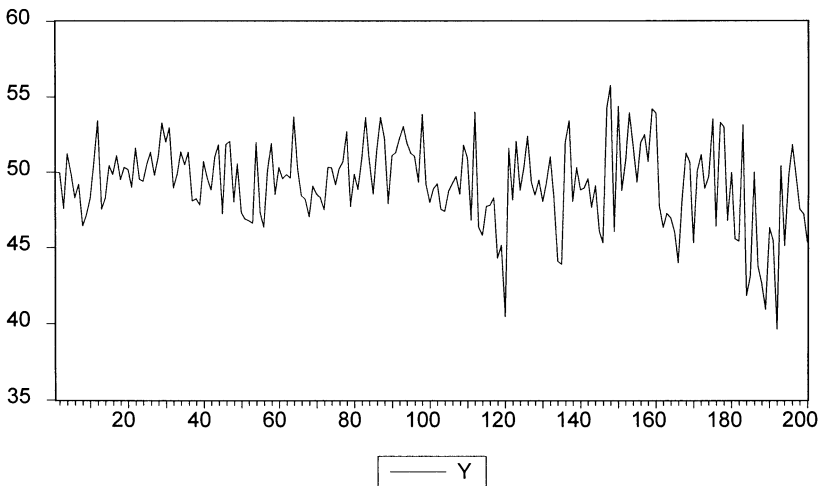
 fichier C8EX5

Étude d'un processus GARCH

Soit un processus y_t . On demande d'en étudier les propriétés et d'estimer les paramètres du modèle par une méthode adéquate.

Solutions

Cet exercice est résolu à partir d'Eviews.



L'étude complète (tests de spécifications...) de ce processus à partir de la méthodologie de Box et Jenkins laisse présager d'un ARMA(1, 1). L'estimation des paramètres conduit aux résultats suivants :





















Dependent Variable: Y				
Included observations: 199 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.639871	4.421883	1.727742	0.0856
Y(-1)	0.844845	0.089698	9.418815	0.0000
MA(1)	-0.647517	0.126607	-5.114389	0.0000

Le corrélogramme des résidus est le suivant :

Sample: 2 200

Included observations: 199

Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.010	-0.010	0.0199	
		2 -0.078	-0.078	1.2659	0.261
		3 0.157	0.156	6.2703	0.043
		4 -0.026	-0.031	6.4039	0.094
		5 0.070	0.098	7.4262	0.115
		6 0.007	-0.024	7.4367	0.190
		7 -0.086	-0.065	8.9717	0.175
		8 -0.054	-0.086	9.5845	0.213
		9 -0.090	-0.100	11.286	0.186
		10 0.027	0.034	11.438	0.247

















La statistique de Q Ljung-Box indique un corrélogramme dont les termes ne sont pas significativement différents de 0, les résidus sont donc non corrélés.

Le corrélogramme des résidus carrés est le suivant :

Sample: 2 200

Included observations: 199

Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.408	0.408	33.659	
		2 0.213	0.055	42.830	0.000
		3 0.208	0.124	51.648	0.000
		4 0.117	-0.018	54.441	0.000
		5 0.091	0.030	56.155	0.000
		6 0.042	-0.033	56.514	0.000
		7 0.033	0.014	56.740	0.000
		8 0.184	0.192	63.852	0.000

La statistique de Q Ljung-Box indique un corrélogramme dont les termes sont significativement différents de 0, une spécification de type ARCH est donc à retenir. Ceci est corroboré par la statistique du multiplicateur de Lagrange ($p = 1$) :

```

=====
ARCH Test :
=====
F-statistic          39.33688      Probability      0.000000
Obs*R-squared       33.09597      Probability      0.000000
=====

```

Après différents essais de spécification ARCH(1), ARCH(2) et GARCH(1, 1), le modèle dont les coefficients sont tous significatifs s'avère être un GARCH(1, 1). C'est donc cette dernière qui est retenue.

```

=====
ARCH // Dependent Variable is Y
Sample(adjusted) : 2 200
Included observations : 199 after adjusting endpoints
Convergence achieved after 45 iterations
=====
Variable      Coefficient Std. Error t-Statistic Prob.
=====
C              8.442109    5.493176    1.536836    0.1260
Y(-1)         0.830275    0.110566    7.509314    0.0000
MA(1)        -0.641980    0.147506   -4.352240    0.0000
=====
                    Variance Equation
=====
C              1.176666    0.554793    2.120910    0.0352
ARCH(1)       0.287708    0.121742    2.363256    0.0191
GARCH(1)      0.540634    0.141073    3.832314    0.0002
=====
R-squared      0.109181    Mean dependent var 49.28626
Log likelihood -457.0356    F-statistic      4.730923
Durbin-Watson stat 1.977453    Prob(F-statistic) 0.000423
=====

```

L'estimation des paramètres du modèle GARCH(1, 1) est donnée par l'instruction Eviews :

ARCH(1, 1) Y C Y(-1) MA(1)

Les coefficients des variables sont tous significatifs, le modèle estimé s'écrit donc :

$y_t = 8,4 + 0,83y_{t-1} + e_t - 0,64e_{t-1}$ dont les erreurs théoriques suivent un GARCH(1, 1) : $\varepsilon_t = u_t \times h_t$ avec $u_t \rightarrow N(0; 1)$ et $h_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}^2$ estimé par : $\hat{h}_t^2 = 1,18 + 0,28 \varepsilon_{t-1}^2 + 0,54 h_{t-1}^2$

La prévision de y_t est calculée de la manière suivante :

SMPL 201 210

GENR U = 0 'on considère le bruit blanc nul

GENR h = C(4) + C(5)*U(-1)*U(-1)*h(-1) + C(6)*h(-1)

GENR EPS = sqrt(H)*U

GENR YF = C(1) + C(2)*YF(-1) + EPS + C(3)*EPS(-1)

E. Autres processus : variantes des processus ARCH

Nous présentons ici, les diverses extensions des modèles ARCH et GARCH utilisées dans le domaine de la finance.

1) Processus de type ARCH-M et GARCH-M

Ces modèles introduits par Engle, Lilien et Robins (1987) sont une extension des modèles ARCH et GARCH.

La forme générale du modèle est : $y = Xa + f(h_t^2) + \varepsilon_t$
avec $h_t^2 = \alpha_0 + \alpha(D)\varepsilon_t^2$ pour un ARCH-M¹
et $y_t/I_{t-1} \rightarrow N(\mu_t; h_t^2)$ avec $\mu_t = Xa + f(h_t^2)$ et $h_t^2 = \alpha_0 + \alpha(D)\varepsilon_t^2$
 $f(\cdot)$ est une fonction quelconque.

Dans ce modèle l'espérance conditionnelle μ_t est fonction de la variance h_t^2 , ce qui signifie que le niveau atteint par la variable est fonction de sa volatilité (ce qui paraît assez réaliste pour des cours boursiers selon l'hypothèse de l'aversion pour le risque des agents et donc que l'espérance de gain est fonction de la variance).

La spécification GARCH-M est obtenu en écrivant la variance conditionnelle :

$$h_t^2 = \alpha_0 + \alpha(D)\varepsilon_t^2 + \beta(D)h_t^2$$

Cocco et Paruolo (1990) définissent un modèle dans lequel c'est l'accroissement de la volatilité (et non la volatilité elle-même) qui va influencer le niveau atteint par la variable à expliquer.

2) Processus de type GARCH-DM, GARCH-DLM et TARCH

On peut dans le même esprit, spécifier un processus ARCH-DM :

$$\text{GARCH-DM}^2 : y = x\beta + f(h_t^2 - h_{t-1}^2) + \varepsilon_t$$

Ce modèle peut être étendu en intégrant les effets passés de la volatilité à l'aide d'un polynôme de retards, il s'agit d'un modèle de type GARCH-DLM³ :

$$y = Xa + \Delta(D)(h_t^2 - h_{t-1}^2) + \varepsilon_t \text{ avec ARCH}(p,0) \quad h_t^2 = \alpha_0(D)\varepsilon_t^2$$

$$\text{GARCH}(p,q) \quad h_t^2 = \alpha_0 + \alpha(D)\varepsilon_t^2 + \beta(D)h_t^2$$

Dans une spécification de type TARCH⁴, on différencie l'effet d'une erreur positive et d'une erreur négative (erreur sur la variance).

-
1. ARCH-M comme Mean (moyenne).
 2. DM = Difference in Mean.
 3. Distributed Lag in Mean.
 4. Treshold ARCH.

La spécification TARCH est la suivante :

$$h_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^+ + \sum_{i=1}^p \beta_i \varepsilon_{t-i}^-$$

avec : $\varepsilon_t^+ = \text{Max}(\varepsilon_t, 0)$ et $\varepsilon_t^- = \text{Min}(\varepsilon_t, 0)$.

Un modèle de ce type permet de prendre en compte une dissymétrie de l'information : les agents suivent un comportement différent selon que la variable à expliquer est à la hausse ou à la baisse.

Les variantes des processus ARCH présentées dans ce chapitre ne sont pas exhaustives ¹, le lecteur peut consulter les ouvrages spécialisés dans ce domaine pour plus d'informations.

Exercice n° 8.6

fichier C5EX4

Modélisation du cours en dollar du baril de pétrole brut

Lors de l'exercice 4 du chapitre 5, nous avons appliqué la stratégie des tests DF, DFA et PP au prix du baril de pétrole brut connu jour par jour sur 2087 jours (2087 observations). Nous avons conclu que le cours est un processus DS sans dérive à mémoire. Nous allons donc maintenant le modéliser à l'aide de la méthode de Box-Jenkins.

Solutions

La Figure 8.6 présente le corrélogramme du cours du pétrole brut en différences premières. Nous constatons que ce n'est pas un bruit blanc. Nous allons donc chercher sa représentation dans la classe des processus ARMA.

Included observations: 2086

	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.163	0.163	55.372	0.000		
2	-0.072	-0.101	66.197	0.000		
3	-0.042	-0.013	69.890	0.000		
4	-0.036	-0.035	72.567	0.000		
5	-0.024	-0.018	73.777	0.000		
6	-0.029	-0.029	75.549	0.000		
7	-0.003	0.002	75.568	0.000		
8	0.062	0.058	83.713	0.000		
9	0.047	0.024	88.335	0.000		
10	-0.007	-0.012	88.452	0.000		
11	-0.042	-0.032	92.140	0.000		
12	-0.040	-0.025	95.461	0.000		
13	-0.010	-0.002	95.681	0.000		
14	0.001	-0.001	95.683	0.000		
15	0.033	0.032	97.951	0.000		

Figure 8.6 – Corrélogramme du cours du pétrole brut en différences premières

1. On peut par exemple proposer le processus FIGARCH qui permet de définir une double mémoire sur les moments des processus (Baillie, 1996 Teysiere, 1997).

Après quelques essais, nous proposons le modèle [1] suivant, dont les coefficients sont bien significativement différents de 0 :

Dependent Variable : D(PPB)				
Included observations : 2085 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	- 0,258801	0,101559	- 2,548288	0.0109
MA(1)	0.441608	0.094313	4.682384	0.0000
R-squared	0.034736	Mean dependent var		0.003496
Adjusted R-squared	0.034273	S.D dependent var	0.270231	
S.E of regression	0.265560	Akaike info criterion		0.187008
Sum squared resid	146.8977	Schwarz criterion		0.192421
Log likelihood	- 192.9560	Durbin-Watson stat	2.004750	

L'analyse des corrélogrammes du résidu (Figure 8.7) et du résidu au carré (Figure 8.8) indique d'une part que le résidu n'est pas un bruit blanc et d'autre part qu'il présente les caractéristiques d'une hétéroscédasticité. Cela nous suggère une hétéroscédasticité de type ARCH.

Included observations: 2085
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.003	-0.003	0.0144	
		2 -0.021	-0.021	0.9455	
		3 -0.047	-0.047	5.5137	0.019
		4 -0.023	-0.024	6.6460	0.036
		5 -0.020	-0.022	7.4834	0.058
		6 -0.023	-0.027	8.5982	0.072
		7 -0.008	-0.011	8.7241	0.121
		8 0.057	0.053	15.449	0.017
		9 0.038	0.035	18.433	0.010
		10 -0.007	-0.007	18.538	0.018

Figure 8.7 – Corrélogramme du résidu du modèle [1]

Included observations: 2085
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.082	0.082	14.104	
		2 0.081	0.075	27.752	
		3 0.171	0.160	88.709	0.000
		4 0.062	0.034	96.711	0.000
		5 0.128	0.102	130.77	0.000
		6 0.115	0.072	158.46	0.000
		7 0.025	-0.013	159.82	0.000
		8 0.057	0.010	166.55	0.000
		9 0.028	-0.015	168.15	0.000
		10 0.084	0.062	182.97	0.000

Figure 8.8 – Corrélogramme du résidu au carré du modèle [1]

Après sélection du nombre de retards à 3 à l'aide du critère d'information de Akaike, nous procédons au test ARCH. La statique de Fisher ou du *LM* indique clairement la présence d'une hétéroscédasticité de type ARCH.

ARCH Test :				
F-statistic	27.26516	Probability	0.000000	
Obs*R-squared	78.84921	Probability	0.000000	
Test Equation :				
Dependent Variable : RESID^2				
Included observations : 2082 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.050083	0.004518	11.08507	0.0000
RESID^2(-1)	0.064846	0.021651	2.995044	0.0028
RESID^2(-2)	0.062267	0.021636	2.877931	0.0040
RESID^2(-3)	0.160462	0.021633	7.417363	0.0000

Nous proposons donc le modèle [2] à erreurs GARCH. Tous les coefficients sont bien significativement différents de 0, nous allons donc analyser les résidus.

Dependent Variable : D(PPB)				
Method : ML - ARCH (Marquardt)				
Included observations : 2085 after adjusting endpoints				
	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	- 0.242000	0.119414	- 2.026559	0.0427
MA(1)	0.407333	0.112672	3.615198	0.0003
Variance Equation				
	Coefficient	Std. Error	z-Statistic	Prob.
C	0.001064	0.000248	4.284927	0.0000
ARCH(1)	0.076236	0.006535	11.66560	0.0000
GARCH(1)	0.911234	0.007885	115.5618	0.0000
R-squared	0.034295	Mean dependent var	0.003496	
Log likelihood	- 40.21315	Durbin-Watson stat	1.970330	

L'analyse du corrélogramme du résidu (Figure 8.9) indique que maintenant le résidu est un bruit blanc.

Included observations: 2085
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.012	0.012	0.2950	
		2 -0.015	-0.016	0.7895	
		3 -0.027	-0.027	2.3468	0.125
		4 0.000	0.001	2.3490	0.309
		5 0.004	0.004	2.3906	0.495
		6 0.013	0.012	2.7568	0.599
		7 -0.014	-0.014	3.1844	0.672
		8 0.037	0.038	6.0619	0.416
		9 0.023	0.022	7.1696	0.411
		10 -0.018	-0.018	7.8182	0.451
		11 -0.039	-0.036	10.949	0.279
		12 -0.016	-0.014	11.463	0.323
		13 -0.015	-0.016	11.913	0.370
		14 0.003	0.000	11.934	0.451
		15 0.018	0.017	12.611	0.478

Figure 8.9 – Corrélogramme du résidu du modèle [2]

Le modèle est donc validé, sa représentation est la suivante :

PPB_t = Prix du pétrole brut en dollars.

$$D(\text{PPB})_t = \text{PPB}_t - \text{PPB}_{t-1}$$

$$D(\text{PPB})_t = -0,24 D(\text{PPB})_{t-1} + \varepsilon_t + 0,40 \varepsilon_{t-1}$$

avec $\varepsilon_t = u_t \times h_t$ où $u_t \rightarrow N(0; 1)$

$$\text{et } h_t^2 = 0,001064 + 0,076 \varepsilon_{t-1}^2 + 0,911 h_{t-1}^2$$

Le Prix du pétrole brut en dollars est un ARIMA(1, 1, 1) à erreurs GARCH.

Exercice n° 8.7

Correction de l'exercice 5.7, le jeu des quatre erreurs

fichier C5EX7

Déterminer le processus générateur du prix du produit dérivé.

Solution synthétique

a) Détermination du processus TS ou DS ?

– Test Dickey-Fuller augmenté (retards de 0 à 6), Critères de Akaike ou Schwarz : retard = 1.

Stratégies de test :

	CALCULÉ	LU	CONCLUSION
MODÈLE 6:			
HO:	- 1.840565	- 3.430000	ACCEPTE
HO6:	1.720239	.340000	ACCEPTE
MODÈLE 5:			
HO:	- 1.841787	- 2.880000	ACCEPTE
HO5:	1.202288	4.750000	ACCEPTE
MODÈLE 4:			
HO4:	1.779579	4.630000	ACCEPTE

– Les tests de Phillips Perron pour les modèles 1, 2 et 3 font rejeter aussi l'hypothèse H0, il existe une racine unitaire.

Le prix est un processus de type DS, il convient de le stationnariser par les différences premières.

b) Est-ce une marche au hasard ou un processus ARMA ?

Corrélogramme de la série en différences premières :

Sample: 0001M10 0024M04
Included observations: 270

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.284	0.284	21.973	0.000
		2	-0.001	-0.088	21.973	0.000
		3	-0.061	-0.039	22.982	0.000
		4	-0.076	-0.051	24.583	0.000
		5	-0.102	-0.076	27.479	0.000
		6	0.047	0.102	28.105	0.000
		7	0.080	0.028	29.899	0.000
		8	0.134	0.106	34.948	0.000
		9	0.213	0.166	47.739	0.000
		10	0.150	0.064	54.081	0.000

Il ne s'agit pas d'une marche au hasard, il existe donc une représentation dans la classe des processus ARMA.

c) Recherche du processus générateur

Représentation AR(1) :

Dependent Variable: D(PRIX)				
Included observations: 269 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.283454	0.058632	4.834461	0.0000

Le test ARCH ($p = 4$) nous fait rejeter l'hypothèse H_0 , le processus est à erreurs ARCH.

ARCH Test $p = 4$:			
F-statistic	7.572414	Probability	0.000009
Obs*R-squared	27.65086	Probability	0.000015

Après différents essais (MA(1), ARCH, GARCH, ..), la représentation suivante est retenue :

Dependent Variable: D(PRIX)

Included observations: 269 after adjustments

	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.345144	0.084935	4.063618	0.0000
Variance Equation				
C	1418.869	292.4837	4.851104	0.0000
RESID(-1)^2	0.325917	0.037896	8.600343	0.0000
GARCH(-1)	0.715562	0.023636	30.27382	0.0000

Le corrélogramme des résidus indique qu'il s'agit d'un bruit blanc, d'après la statistique de Jarque-Bera la distribution n'est pas gaussienne.

Sample: 0001M12 0024M04

Included observations: 269

Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.060	0.060	0.9856	
		2 -0.019	-0.023	1.0868	0.297
		3 -0.031	-0.028	1.3426	0.511
		4 0.015	0.018	1.4059	0.704
		5 -0.019	-0.022	1.5030	0.826
		6 0.040	0.042	1.9414	0.857
		7 0.052	0.048	2.6968	0.846
		8 0.009	0.003	2.7176	0.910
		9 0.088	0.093	4.8672	0.772
		10 0.092	0.084	7.2375	0.612

Jarque-Bera = 230,12 (Probabilité critique = 0,00)

La représentation finale est donc la suivante : PRIX est un ARIMA(1, 1, 0) à effet GARCH.

$$D(\text{prix})_t = 0,345D(\text{prix})_{t-1} + \varepsilon_t$$

avec $\varepsilon_t = u_t \times h_t$, $u_t \rightarrow N(0; 1)$ et $h_t^2 = 1418,87 + 0,326 \varepsilon_t^2 + 0,715 h_t^2$

d) Calcul de la prévision

La prévision pour le mois de Mai de l'année 24 est donnée par :

$$D(\text{prix})_{\text{Mai}} = 0,345D(\text{prix})_{\text{Avril}} = 0,345 \times -173 = -59,71$$

$$\text{prix}_{\text{Mai}} = \text{prix}_{\text{Avril}} + D(\text{prix})_{\text{Mai}} = 3640 - 59,71 = 3580,29 .$$

Liste des exercices

Afin que le lecteur puisse lui-même refaire les exercices, les données utilisées (sous format Excel et Eviews), ainsi que les programmes de traitement « *Batch* » de Eviews, sont disponibles par téléchargement sur le serveur web :

<http://regisbourbonnais.dauphine.fr>

Chapitre 1

<i>Exercice n° 1.1</i> — Test de détection de la saisonnalité à partir des données du Tableau 1.1	14
<i>Exercice n° 1.2</i> — Calcul d'une fonction d'autocorrélation	18
<i>Exercice n° 1.3</i> — Désaisonnalisation selon un schéma additif	24
<i>Exercice n° 1.4</i> — Désaisonnalisation par fonction trigonométrique	27
<i>Exercice n° 1.5</i> — Désaisonnalisation à l'aide de variables dichotomiques	30
<i>Exercice n° 1.6</i> — Un exemple d'application de saisonnalité mixte (additive et multiplicative) : le cas de la téléphonie mobile	34
<i>Exercice n° 1.7</i> — Désaisonnalisation par moyenne mobile selon un schéma additif	38
<i>Exercice n° 1.8</i> — Désaisonnalisation par la méthode de Census X11	41
<i>Exercice n° 1.9</i> — Désaisonnalisation par démodulation complexe	43

Chapitre 2

<i>Exercice n° 2.1</i> — Estimation d'un modèle de diffusion de type Logistique	49
<i>Exercice n° 2.2</i> — Calcul d'une prévision par un LES	58
<i>Exercice n° 2.3</i> — Prévision d'une chronique à partir d'un LED	63
<i>Exercice n° 2.4</i> — Prévision par agrégation des composantes extra-saisonnalité et saisonnalité	69
<i>Exercice n° 2.5</i> — Prévision par le modèle de Holt-Winters	76

Chapitre 3

<i>Exercice n° 3.1</i> — Calcul des termes d'un corrélogramme et de leur intervalle de confiance à partir des éléments d'une chronique	98
--	----

<i>Exercice n° 3.2</i> — Exemples d'analyse des FAC et FAP au CAC40	94
<i>Exercice n° 3.3</i> — Recherche des conditions de stationnarité et d'inversibilité des processus	109
<i>Exercice n° 3.4</i> — Exemples de génération de processus ARMA à l'aide d'Eviews	111
<i>Exercice n° 3.5</i> — Calculs des caractéristiques de différents processus et études des propriétés d'inversibilité et de stationnarité	112

Chapitre 4

<i>Exercice n° 4.1</i> — Propriétés du filtre du Lissage Exponentiel Simple (LES)	130
<i>Exercice n° 4.2</i> — Relation entre la variable m et les valeurs j	143
<i>Exercice n° 4.3</i> — Calcul de la puissance spectrale d'un processus AR(2)	145
<i>Exercice n° 4.4</i> — Simulation graphique du spectre de différents processus ARMA	147
<i>Exercice n° 4.5</i> — Estimations de spectres de chroniques simulées	149

Chapitre 5

<i>Exercice n° 5.1</i> — Conséquence d'une mauvaise stationnarisation d'un processus	160
<i>Exercice n° 5.2</i> — Exemple d'application des tests DF et DFA aux dépenses en produits pharmaceutiques	176
<i>Exercice n° 5.3</i> — Exemple d'application des tests DF, DFA, Phillips-Perron, KPSS, ERS et Ng-Perron au CAC40	182
<i>Exercice n° 5.4</i> — Exemple d'application des tests DF, DFA et Phillips-Perron et KPSS au cours en dollars du baril de pétrole brut	189
<i>Exercice n° 5.5</i> — Application du test de Franses	196
<i>Exercice n° 5.6</i> — Relation entre lissage exponentiel simple et modèle ARIMA	200
<i>Exercice n° 5.7</i> — Le jeu des quatre erreurs	201

Chapitre 6

<i>Exercice n° 6.1</i> — Génération de processus AR et analyse des corrélogrammes	211
<i>Exercice n° 6.2</i> — Génération de processus MA et analyse des corrélogrammes	215
<i>Exercice n° 6.3</i> — Calcul des termes de la FAC d'un processus MA(2) théorique	221

<i>Exercice n° 6.4</i> — Calcul des termes de la FAC et de la FAP de processus ARMA	222
<i>Exercice n° 6.5</i> — Représentations graphiques des FAC, FAP, FACI, FAPI	230
Chapitre 7	
<i>Exercice n° 7.1</i> — Tests de bruit blanc et de normalité sur l'indice CAC40	251
<i>Exercice n° 7.2</i> — Analyse de la saisonnalité des immatriculations des voitures particulières en France	260
<i>Exercice n° 7.3</i> — Comparaison de méthodes classiques de prévision du nombre des immatriculations en France	265
<i>Exercice n° 7.4</i> — Prévision par la méthodologie de Box et Jenkins du nombre des immatriculations en France	267
<i>Exercice n° 7.5</i> — Analyse par la méthodologie de Box et Jenkins et prévision de trois processus aléatoires	271
Chapitre 8	
<i>Exercice n° 8.1</i> — Simulation et calcul des statistiques de Hurst, Lo et de GPH	290
<i>Exercice n° 8.2</i> — Applications des tests BDS, dimension de corrélation et exposant de Lyapunov	297
<i>Exercice n° 8.3</i> — Simulation d'une erreur ARCH	303
<i>Exercice n° 8.4</i> — Étude d'un processus ARCH	306
<i>Exercice n° 8.5</i> — Étude d'un processus GARCH	311
<i>Exercice n° 8.6</i> — Modélisation du cours en dollar du baril de pétrole brut	315
<i>Exercice n° 8.7</i> — Correction de l'exercice 5.7, le jeu des quatre erreurs	318

T ables statistiques

Table 1 – Distribution cumulative empirique de t_{ρ^*} , $t_{\hat{\rho}}$, $t_{\bar{\rho}}$ sous $H_0 : \rho = 0$

	Probabilité d'une valeur inférieure							
n	0,01	0,025	0,05	0,10	0,90	0,95	0,975	0,99
Modèle (1) : t_{ρ^*}								
25	-2,66	-2,26	-1,95	-1,60	0,92	1,33	1,70	2,16
50	-2,62	-2,25	-1,95	-1,61	0,91	1,31	1,66	2,08
100	-2,60	-2,24	-1,95	-1,61	0,90	1,29	1,64	2,03
250	-2,58	-2,23	-1,95	-1,62	0,89	1,29	1,63	2,01
500	-2,58	-2,23	-1,95	-1,62	0,89	1,28	1,62	2,00
∞	-2,58	-2,23	-1,95	-1,62	0,89	1,28	1,62	2,00
Modèle (2) : $t_{\hat{\rho}}$								
25	-3,75	-3,33	-3,00	-2,63	-0,37	0,00	0,34	0,72
50	-3,58	-3,22	-2,93	-2,60	-0,40	-0,03	0,29	0,66
100	-3,51	-3,17	-2,89	-2,58	-0,42	-0,05	0,26	0,63
250	-3,46	-3,14	-2,88	-2,57	-0,42	-0,06	0,24	0,62
500	-3,44	-3,13	-2,87	-2,57	-0,43	-0,07	0,24	0,61
∞	-3,43	-3,12	-2,86	-2,57	-0,44	-0,07	0,23	0,60
Modèle (3) : $t_{\bar{\rho}}$								
25	-4,38	-3,95	-3,60	-3,24	-1,14	-0,80	-0,50	-0,15
50	-4,15	-3,80	-3,50	-3,18	-1,19	-0,87	-0,58	-0,24
100	-4,04	-3,73	-3,45	-3,15	-1,22	-0,90	-0,62	-0,28
250	-3,99	-3,69	-3,43	-3,13	-1,23	-0,92	-0,64	-0,31
500	-3,98	-3,68	-3,42	-3,13	-1,24	-0,93	-0,65	-0,32
∞	-3,96	-3,66	-3,41	-3,12	-1,25	-0,94	-0,66	-0,33

Source : Fuller 1976, p. 373, Tableau 8.5.2.

Table 2 – Distribution empirique de $t_{\hat{c}}$ pour $(c, \rho) = (0, 0)$ Modèle (2)

n	Probabilité d'une valeur inférieure			
	0,90	0,95	0,975	0,99
Modèle (2) : $t_{\hat{c}}$				
25	2,20	2,61	2,97	3,41
50	2,18	2,56	2,89	3,28
100	2,17	2,54	2,86	3,22
250	2,16	2,53	2,84	3,19
500	2,16	2,52	2,83	3,18
∞	2,16	2,52	2,83	3,18

Source : Dickey et Fuller, 1981 , p.1062, Tableau I.

Table 3 – Distribution empirique de Φ_1 pour $H_0^1 : (c, \rho) = (0, 0)$ Modèle (2)

n	Probabilité d'une valeur inférieure							
	0,01	0,025	0,05	0,10	0,90	0,95	0,975	0,99
Modèle (2) $\Phi_1 : (c, \rho) = (0, 0)$								
25	0,29	0,38	0,49	0,65	4,12	5,18	6,30	7,88
50	0,29	0,39	0,50	0,66	3,94	4,86	5,80	7,06
100	0,29	0,39	0,50	0,67	3,86	4,71	5,57	6,70
250	0,30	0,39	0,51	0,67	3,81	4,63	5,45	6,52
500	0,30	0,39	0,51	0,67	3,79	4,61	5,41	6,47
∞	0,30	0,40	0,51	0,67	3,78	4,59	5,38	6,43

Source : Dickey et Fuller, 1981, p.1063, Tableau IV.

Table 4 – Distribution empirique de $t_{\hat{c}}$ pour $(c, b, \rho) = (0, 0, 0)$ Modèle (3)

	Probabilité d'une valeur inférieure			
n	0,90	0,95	0,975	0,99
Modèle (3) : $t_{\hat{c}}$				
25	2,77	3,20	3,59	4,05
50	2,75	3,14	3,47	3,87
100	2,73	3,11	3,42	3,78
250	2,73	3,09	3,39	3,74
500	2,72	3,08	3,38	3,72
∞	2,72	3,08	3,38	3,71

Source : Dickey et Fuller, 1981, p.1062, Tableau II.

Table 5 – Distribution $t_{\hat{b}}$ pour $(c, b, \rho) = (0, 0, 0)$ Modèle (3)

	Probabilité d'une valeur inférieure			
n	0,90	0,95	0,975	0,99
Modèle (3) : $t_{\hat{b}}$				
25	2,39	2,85	3,25	3,74
50	2,38	2,81	3,18	3,60
100	2,38	2,79	3,14	3,53
250	2,38	2,79	3,12	3,49
500	2,38	2,78	3,11	3,48
∞	2,38	2,78	3,11	3,46

Source : Dickey et Fuller, 1981, p.1062, Tableau III.

Table 6 – Distribution empirique de Φ_2 pour $H_0^2 : (c, b, \rho) = (0, 0, 0)$
Modèle (3)

	Probabilité d'une valeur inférieure							
n	0,01	0,025	0,05	0,10	0,90	0,95	0,975	0,99
Modèle (3) $\Phi_2 H_0^2 : (c, b, \rho) = (0, 0, 0)$								
25	0,61	0,75	0,89	1,10	4,67	5,68	6,75	8,21
50	0,62	0,77	0,91	1,12	4,31	5,13	5,94	7,02
100	0,63	0,77	0,92	1,12	4,16	4,88	5,59	6,50
250	0,63	0,77	0,92	1,13	4,07	4,75	5,40	6,22
500	0,63	0,77	0,92	1,13	4,05	4,71	5,35	6,15
∞	0,63	0,77	0,92	1,13	4,03	4,68	5,31	6,09

Source : Dickey et Fuller, 1981 , p.1063, Tableau V.

Table 7 – Distribution empirique de Φ_3 pour $H_0^3 : (c, b, \rho) = (c, 0, 0)$
Modèle (3)

	Probabilité d'une valeur inférieure							
n	0,01	0,025	0,05	0,10	0,90	0,95	0,975	0,99
Modèle (3) $\Phi_3 H_0^3 : (c, b, \rho) = (0, 0, 0)$								
25	0,74	0,90	1,08	1,33	5,91	7,24	8,65	10,61
50	0,76	0,93	1,11	1,37	5,61	6,73	7,81	9,31
100	0,76	0,94	1,12	1,38	5,47	6,49	7,44	8,73
250	0,76	0,94	1,13	1,39	5,39	6,34	7,25	8,43
500	0,76	0,94	1,13	1,39	5,36	6,30	7,20	8,34
∞	0,77	0,94	1,13	1,39	5,34	6,25	7,16	8,27

Source : Dickey et Fuller, 1981 , p.1063, Tableau VI.

Bibliographie

- ABRAHAM B., LEDOLTER J., *Statistical methods for forecasting*, John Wiley, 1983.
- AKAIKE H., A bayesian extension of the minimum AIC procedure, *Biometrika*, 66, 1979.
- ANDREWS W., LO A.W., Long term memory in stock market prices, *Econometrica*, Vol. 59, 1991.
- AZNAR A., TRIVEZ F. J., *Métodos de prediccion in economia*, Ariel, Barcelona, 1993.
- BAILLIE R. T., Long memory process and fractional cointegration in econometrics, *Journal of Econometrics*, 73, 1996.
- BARTHÉLÉMY F., Tests de racine unitaire multiples et saisonnalité, *Document de travail du Grequam*, Université de Marseille III, 1996.
- BEAULIEU J., MIRON J., Seasonal unit root in aggregate U.S. data, *Journal of Econometrics*, 55, 1993.
- BEGUIN J. M., GOURIÉROUX C., MONFORT A., *Identification of an ARIMA process : the corner method*, Times Series, Ed. T. Anderson, North Holland, 1980.
- BENSABER A., BLEUSE-TRILLON B., *Pratique des chroniques et de la prévision à court terme*, Masson, 1989.
- BERLINÉ A., Estimation the degrees of an ARMA model, *Compustat*, 1985.
- BERLINÉ A., VAN CUTSEM B., Sequence transformation : a common domain to numerical analysis and statistics, *Journées Franco-Belges de Statistiques*, 1986.
- BEVERIDGE S., NELSON C. R., A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle, *Journal of the Monetary Economics*, 7, 2, 1981.
- BHANSALI R. J., Autoregressive and window estimate of the inverse correlation function, *Biometrika*, Vol. 67, n°3, 1980.
- BHANSALI R. J., The inverse partial correlation function of a time series and its applications, *Journal of Multivariate Analysis*, Vol. 13, n°2, 1983.
- BLOOMFIELD P., *Fourier analysis of time series : an introduction*, John Wiley, New York, 1976.
- BOLLERSLEV T., CHOU R. Y., KRONER K., ARCH modelling in finance : a review of the theory and empirical evidence, *Journal of Econometrics*, Vol. 52, 1992.
- BOLLERSLEV T., ENGLE R. F., Modelling the persistence of conditional variances, *Econometric Review*, Vol. 5, 1986.
- BOLLERSLEV T., Generalised autoregressive conditional heteroskedasticity, *Journal of Econometrics*, Vol. 31, 1986.

- BOLLERSLEV T., On the correlation structure for the generalized autoregressive conditional heteroscedastic process, *Journal of Time Series Analysis*, 9, 1988.
- BOURBONNAIS R., *Econométrie : cours et exercices corrigés*, Dunod, 7 éd., 2009.
- BOURBONNAIS R., USUNIER J-C., *Prévision des ventes*, Economica, 4 éd., 2007.
- BOURBONNAIS R., VALLIN PH., The correction of the seasonal adjustment of times series with additional and multiplying effects, *Romanian Journal of Economic Forecasting*, November 2007.
- BOX G. E. P., JENKINS G. M., *Time series analysis : forecasting and control*, Holden-day, 1976.
- BOX G. E. P., PIERCE D. A., Distribution of residual autocorrelations in autoregressive moving average time series models, *Journal of the American Statistical Association*, 65, 1970.
- BRESSON G., PIROTTE A., *Econométrie des séries temporelles*, PUF, 1995.
- BREUSCH T. S., PAGAN A. R., A simple test for heteroskedasticity and random coefficient variation, *Econometrica*, Vol. 46, 1978.
- BROCK W. A., DECHERT W. D., SCHEINKMAN J., A test for independence based on the correlation dimension, *University of Wisconsin, University of Houston, University of Chicago*, 1987.
- BROCK W. A., DECHERT W. D., SCHEINKMAN J., LEBARON B., A test for independence based upon the correlation dimension, *Econometric Review*, Vol. 15, 1996.
- BROCKWELL P. J., DAVIS R., *Introduction to time series and forecasting*, Springer-Verlag, 1996.
- BROCKWELL P. J., DAVIS R., *ITSM : An introduction time series modelling package for the P C*, Springer-Verlag, 1991.
- BROCKWELL P. J., DAVIS R., *Time series : theory and method*, Springer Verlag, 1987.
- BROWN R. G., DURBIN J., EVANS J. M., Techniques for testing the constancy of the regression relationship overtime, *Journal of the Royal Statistical Society*, B, 37(2), 1975.
- BROWN R. G., *Smoothing, forecasting and prediction of discrete time series*, Prentice Hall, New Jersey, 1962.
- CHAN KH., HAYYA J.C., ORD J.K., A note on trend removal methods : the case of polynomial regression versus variate differencing, *Econometrica*, 45, 1977.
- CHATFIELD C., Inverse autocorrelation, *Journal of Royal Statistical Society*, Vol. 142, 1979.
- CHATFIELD C., The Holt-Winters forecasting procedure, *Applied Statistics*, 3, 1978.
- CHEN G., ABRAHAM B., PEIRIS M.S., Lag window estimation of the degree of differencing in fractionally integrated time series, *Journal of Time Series Analysis*, Vol. 15, 1994.
- CHUNG C. F., A generalized fractionally integrated autotregressive moving average process, *Journal of Time Series Analysis*, Vol. 17, 1996.
- CHUNG C. F., Estimating a generalized long memory process, *Journal of Econometrics*, Vol. 5, 1996.

- CLEVELAND W. S., The inverse autocorrelation of a time series and their applications, *Technometrics*, Vol. 14, n°2, 1972.
- COBLENTZ J. F., *Introduction à l'analyse de Fourier*, Eyrolles, 1988.
- COCCO F., PARUOLO P., Volatility persistence and the italian risk premium : parametric and non-parametric evaluation, *Papier présenté à la conférence sur les processus ARCH Paris, INSEE Juin 1990*, 1990.
- CROMWELL J. B., HANNAN M. J., LABYS W. C., TERRAZA M., Multivariate tests for time series models, *Sage Publications*, Vol. 100, 1994.
- CROMWELL J. B., LABYS W. C., TERRAZA M., Univariate tests for time series models, *Sage Publications*, Vol. 99, 1994.
- DAGUM E. B., Fondement des deux principaux types de méthodes de désaisonnalisation et de la méthode X11-ARMMI, *Economie Appliquée*, Vol 1, 1979.
- DARNÉ O., La désaisonnalisation des chroniques économiques – Analyse des conditions conjoncturelles, *Thèse de Doctorat*. Université de Montpellier I. 2002.
- DARNÉ O., LITAGO J., TERRAZA M., Tests de racines unitaires saisonnières pour données journalières, *Revue de Statistique appliquée*, 2002.
- DARNÉ O., TERRAZA M., Les tests de non stationnarité des séries temporelles non saisonnières en économie : une revue de la littérature. *Document de travail 2002-05*. Lameta Université de Montpellier I. 2002.
- DAVIDSON R., MACKINNON J. G., *Estimation and inference in econometrics*, Oxford University Press, 1993.
- DICKEY D., FULLER W., Distribution of the estimators for autoregressive time series with unit root, *Journal of the American Statistical Association*, 74, 366, 1979.
- DICKEY D., FULLER W., Likelihood ratio statistics for autoregressive time series with unit root, *Econometrica*, 49, 4, 1981.
- DICKEY D., HASZA D., FULLER W., Testing for unit root in seasonal time series, *Journal of the American Statistical Association*, 79, 1984.
- DICKEY D., PANTULA S., Determining the order of differencing in autoregressive processes, *Journal of Business Economic and Statistics*, 5, 1987.
- DOOB J. L., *Stochastic process*, John Wiley, 1953.
- ELLIOTT G., ROTHENBERG T., STOCK J., Efficient Tests for an Autoregressive Unit Root, *Econometrica*, Vol. 64, pp. 813-836, 1996.
- ENGLE R. F., Autoregressive conditional heteroscedasticity with estimate of the variance of U. K. inflation, *Econometrica*, Vol. 50, 1982.
- ENGLE R. F., LILIEN D. M., ROBBIN R. P., Estimating time varying risk premia in the term structure : the ARCH-M model, *Econometrica*, 55, 1987.
- ERTUR K. C., Tests de non stationnarité : application au PIB réel, *Thèse Université de Bourgogne*, 1992.
- FISHMAN G., *Spectral methods in econometrics*, Harvard V.P., Cambridge Massachussets, 1969.
- FRANSES P., Testing for seasonal unit root in monthly data, *Econometric Institute Report*, Erasmus University Rotterdam, 1990.

- FUCHS J. J., Estimation de l'ordre d'un processus ARMA à l'aide des résultats de perturbations de matrices, *IRISIA*, n°275, 1985.
- FULLER W. A., *Introduction to statistical time series*, John Wiley, 1976.
- GEWEKE J., PORTER-HUDAK S., The estimation and application of long memory time series models, *Journal of Time Series Analysis*, Vol. 4, 1983.
- GLASBEY C. A., A generalisation of partial autocorrelation useful in identification of ARMA models, *Technometrics*, Vol. 24, n°3, 1982.
- GOMEZ V., MARAVALL A., Program TRAMO and SEATS : instructions for the users, *Banque d'Espagne*. 1997.
- GOURIEROUX C., *Modèles ARCH et applications financières*, Economica, 1992.
- GOURIEROUX C., MONFORT A., *Séries temporelles et modèles dynamiques*, Economica, 1990.
- GOURIEROUX C., MONFORT A., *Cours de séries temporelles*, Economica, 1983.
- GRANGER C. W. J., *Analyse spectrale des séries temporelles en économie*, Dunod, 1969.
- GRANGER C.W.J., JOYEUX R., An introduction to long memory time series and fractional differencing, *Journal of Time Series Analysis*, Vol. 1, 1980.
- GRASSBERGER P., PROCACCIA I., Measuring the strangeness of strange attractors, *Computation statistic and analysis data*, May 1983.
- GRAY H. L., KELLEY C. D., MC INTIRRE D. D., On new approach to ARMA modeling, *Commun. Statist. Simula. Computa.*, B7(1), 1978.
- GRAY H. L., ZANG N. F., WOODWARD W., A generalized fractional process, *Journal of Time Series Analysis*, Vol. 10, 1989.
- GUEGAN D., *Séries chronologiques non linéaires à temps discret*, Economica, 1994.
- HALL A., Testing for unit root in time series with pretested data based model selection, *Mimeo, North Carolina State University*, 1990.
- HAMILTON J. D., *Time series analysis*, Princeton University Press, 1994.
- HARRISON P. J., Short-term forecasting, *Applied Statistics*, 14, 1965.
- HARVEY A. C., *Forecasting structural time series models and the kalman filter*, Cambridge University Press, 1988.
- HARVEY A. C., *The econometric analysis of time series*, Philips Allan, 1990.
- HASSLER U., Regression of spectral estimators with fractionally integrated time series, *Journal of Time Series Analysis*, Vol. 14, 1993.
- HASSLER U., Specification of long memory in seasonal time series, *Journal of Time Series Analysis*, Vol. 15, 1994.
- HASZA D. P., FULLER W. A., Testing non stationary parameter specification in seasonal time series models, *The Annals of Statistics*, 10, 1981.
- HERNAD D., MOUILLARD M., STRAUSS-KAHN D., Du bon usage de R/S, *Revue de Statistique Appliquée*, Vol. 4, 1978.
- HERNAD D., MOUILLARD M., STRAUSS-KAHN D., Forme typique du spectre et dépendance temporelle en économie, *Revue de Statistique Appliquée*, Vol. 4, 1979.
- HIMMELBLAU D. M., *Applied non linear programming*, Mc Graw Hill, 1972.

- HOLT C. *et al.*, *Planning production inventories and work force*, Prentice Hall, 1957.
- HOSKING J. R. M., Fractional differencing, *Biometrika*, Vol. 68, 1981.
- HURST H., Long term storage capacity of reservoirs, *Transaction of the American society of civil engineers*, Vol. 116, 1951.
- HURVICH C., RAY B., Estimation of the memory parameter of nonstationary or noninvertible fractionally integrated processes, *Journal of Time Series Analysis*, Vol. 16, 1995.
- HYLLEBERG S., ENGLE R., GRANGER C., YOO B., Seasonal integration and co-integration, *Journal of the Econometrics*, 44, 1990.
- HYLLEBERG S., *Modelling Seasonality*, Oxford University Press, 1992.
- HYLLEBERG S., *Seasonality in regression*, Academic Press, 1986.
- JARQUE C. M., BERA A. K., Efficient test for normality homoscedasticity and serial independance of regression residuals, *Applied Statistics*, Vol 31, n°2, 1980.
- JARQUE C. M., BERA A. K., Testing the normality assumption in limited dependant variable models, *International Economic Review*, Vol 25, n°3, 1984.
- JENKINS G.M., WATTS D., *Spectral analysis and its applications*, Holden Day, 1968.
- JOHNSTON J., DINARDO J., *Méthodes économétriques*, Economica, 4^e éd., 1999.
- KOOPMANS S. J., HARVEY A. C., DOORNIK J. A., SHEPARD N., *STAMP 5.0 : Structural Times Series Analysis, modelling and prediction*, Chapman and Hall, 1995.
- KWIATKOWSKI D., PHILLIPS P., SCHMIDT P., SHIN Y., Testing the null hypothesis of stationarity against the alternative of a unit root, *Journal of Econometrics*, 54, 159-178, 1992.
- KYRTSOU C., Hétérogénéité et chaos stochastique dans les marchés boursiers, *Thèse Sciences Economiques*, Université de Montpellier I, 2003.
- KYRTSOU C., TERRAZA M., Is it possible to study chaotic and ARCH behaviour jointly comportement economics, *Computational Economics*, Vol. 21, 2003.
- LALOIRE J. C., *Méthodes du traitement des chroniques statistiques et prévision des ventes*, Dunod, 1972.
- LARDIC S., MIGNON V., ROBERT F. ENGLE et CLIVE W.J. Granger – Prix Nobel d'économie 2003, *R.E.R.*, Vol. 114, Janvier-Février 2004.
- LARDIC S., MIGNON V., *Économétrie des séries temporelles macroéconomiques et financières*, Economica 2002.
- LARDIC S., MIGNON V., La mémoire longue en économie : une revue de la littérature, *Journal de la société française de statistique*, n° 2, 1999.
- LAROQUE G., Analyse d'une méthode de désaisonnalisation : le programme de X11 du US bureau of census, Version trimestrielle, *Annales de l'INSEE*, n° 28, 1977.
- LEYBOURNE S., MCCABE B., A consistent test for a unit root, *Journal of Business and Economic Statistics*, Vol. 12, pp. 157-186, 1994.
- LITAGO J., TERRAZA M., Tests de racines unitaires pour chroniques saisonnières, *Document de travail du Lameta*, Université de Montpellier I, 1995.
- LJUNG G. M., BOX G. E. P., On a measure of the lack of fit in time series models, *Biometrika*, 65 , 1978.

- LO A.W., Long term memory in stock market prices, *Econometrica*, Vol. 59, 1991.
- LO A. W., MC KINLAY A. C., The size and power of the variance ratio test in finite samples A Monte Carlo investigation, *Journal of econometric*, Vol. 40, pp. 203-238, 1989.
- MANDELBROT B., *Fractals : form, chances and dimension*, Freeman, San Francisco, 1977.
- MANDELBROT B., VAN NESS J.W., Fractional brownien motion, fractional noises and application, *SIAM Review*, Vol.10, 1968.
- MANDELBROT B., WALLIS J., Noah, Joseph, and operational hydrology, *Water resources research*, Vol. 4, 1968.
- MC LEOD G., *Box Jenkins in practice*, AGJP Publication, 1982.
- MÉLARD G., *Méthodes de la prévision à court terme*, Ellipses, 1990.
- MÉLARD G., Processus purement indéterminable à paramètres discrets, approches fréquentielles et temporelles, *Thèse de doctorat en sciences*, Université libre de Bruxelles, 1975.
- MIGNON V., *Marchés financiers et modélisation des rentabilités boursières*, Economica 1998.
- MILLS T. C., *Time series techniques for economists*, Cambridge University Press, 1990.
- MONTGOMERY D. C., JOHNSON L. A., *Forecasting and time series analysis*, Mc Graw Hill, 1976.
- MOODY J., WU L., Improved estimates for the relaxed range and Hurst exposant, *Neural Network in Engeneering*, World Scientific, 1996.
- NELDER J. A., MEAD R., A simplex method for function minimization, *Compact Journal*, 7, 1965.
- NELSON C. R., KANG H., Spurious periodicity in inappropriates detrended time series, *Econometrica*, 10, 1981.
- NELSON C.R., PLOSSER C., Trends and random walks in macroeconomics time series : some evidence and applications, *Journal of Monatery Economics*, 10, 1982.
- NG S., PERRON P., Lag length selection and the construction of unit root tests with good size and power, *Econometrica*, Vol. 69 (6), pp. 1519-1554, 2001.
- OULIARIS S., PARK J. Y., J. PHILLIPS P., Testing for units roots in the presence of a maintained trend, *Advanced in econometrics and modelling*, Academic Press, 1989.
- PAPARODITIS E., POLITIS N., Residual-based block bootstrap for unit root testing, *Econometrica*, Vol. 71 (3), pp. 813-855, 2003.
- PERRON P., *Testing for a random walk : a simulation experiment of power when the sampling interval is varied*, Advanced in Econometrics and Modelling, Kluwer Academic Publishers, 1989.
- PERRON P., RODRIGUEZ G., GLS Detrending, efficient unit root tests and structural change, *Journal of Econometrics*, 2003.
- PHILLIPS P., PERRON P., Testing for unit root in time series regression, *Biometrika*, 75, 1988.
- PHILIPS L., BLOMME R., VANDEN BERGHE C., DOR E., *Analyse chronologique*, De Boeck Université, Bruxelles, 3 éd., 1990.

- PORTER-HUDAK S., An application of the seasonal fractionally differenced model to the monetary aggregates, *Journal of the American Statistical association*, Vol. 85, 1990.
- PRIESTLEY M. B., *Spectral analysis and time series*, Vol. 1 et Vol. 2, Academic Press, New York, 1981.
- PRIESTLEY M. B., *Non linear and non stationary time series analysis*, Academic Press, 1988.
- PUKKILA T., Fitting of autoregressive moving average models in the frequency domain, *Report A. G.*, University of Tampere, 1977.
- QUENOUILLE M. H., The joint distribution of serial correlation coefficients, *Annals of Mathematical Statistics*, 10, 1949.
- RADIX J. C., *Introduction au filtrage numérique*, Eyrolles, 1970.
- RAO S. T., *Optimisation theory and applications*, Wiley Eastern Limited, 1979.
- RAY B., Long range forecasting of IBM product revenues using a seasonal fractionally differenced ARMA model, *International Journal of Forecasting*, Vol. 9, 1993.
- ROSENBRUCK H. H., An automatic method for finding the greatest or least value of a function, *Computer Journal*, 1960.
- SAÏD S., DICKEY D., Testing for unit root in autoregressive-moving average models of unknown order, *Biometrika*, 71, 1984.
- SALANIÉ B., Guide pratique des séries non stationnaires, *Document de travail G9814*, INSEE, 1998.
- SAMUELSON P. A., Les fondements de l'analyse économique, Tome 2, Dunod 2 éd., 1971.
- SARGAN J. D., BHARGAVA A., Testing the residuals from least squares regression for being generated by the gaussian random walk, *Econometrica*, Vol. 41, pp. 153-174, 1983.
- SCHMIDT P., PHILLIPS P., LM tests for a unit root in the presence of deterministic trends, *Oxford Bulletin of Economics and Statistics*, Vol. 54 (3), pp. 257-287, 1992.
- SCHWARZ G., Estimation the dimension of a model, *Annals of Statistics*, 6, 1978.
- SEO B., Distribution theory for unit root tests with conditional heteroskedasticity, *Journal of Econometrics*, Vol. 91, pp. 113-144, 1999.
- SHISKIN J., Electronic computers and business indicators, *National Bureau of Economic Research*, Occasional Paper, 1954.
- SIMON L., TERRAZA M., La méthode FACE d'identification des processus aléatoires, *Journées de Statistique*, La Grande Motte, 1984.
- SOWELL F., Maximum likelihood estimation of stationary univariate fractionally integrated time series models, *Journal of econometrics*, Vol. 53, 1992.
- TAKENS F., Detecting strange attractors in turbulence. In *Lecture notes in mathematics*, Vol.898. *Dynamical systems and turbulence*, Springer, 1981.

- TERRAZA M., *Economie des processus aléatoires : le cas du marché des vins de consommation courante*, Thèse doctorat d'Etat, Sciences Economiques, Université de Montpellier I., 1981.
- TERRAZA M., L'estimation des modèles ARMA par l'algorithme de la plus forte pente, *Journal de la Société Statistique de Paris*, n°3, 1981.
- TERRAZA M., Les filtres des différences d'ordre p et des différences de moyennes mobiles simples dans l'analyse spectrale, *Journal de la Société Statistiques de Paris*, n°2, 1979.
- TEYSSIÈRE G., Double long memory financial time series, *Document de travail Grequam n°97B01*, 1997.
- TONG H., *Non-linear time series*, Oxford University Press, 1990.
- TSAY R. S., TIAO G. C., Consistent estimated sample autocorrelation function for stationary and nonstationary ARMA models, *Journal of the American Statistical Association*, Vol. 79, n°385, 1984.
- VATÉ M., *Statistique chronologique et prévision*, Economica, 1993.
- WHITE H., A heteroskedasticity consistent covariance estimator and a direct test for heteroskedasticity, *Econometrica*, Vol. 48, 1980.
- WINTERS P. R., Forecasting sales by exponentially weighted moving averages, *Management Science*, Vol. 6, 1960.
- WOLD H., *A study in the analysis of stationary time series*, Almqvist-Wiksell, (première publication 1938), 1954.
- WOODWARD W. A., GRAY H. L., On the relationship between the S-array and the Box-Jenkins method of ARMA model identification, *Journal of the American Statistical Association*, Vol. 76, n°375, 1981.
- XIAO Z., Testing the null hypothesis of stationarity against an autoregressive unit root. alternative, *Journal of Time Series Analysis*, Vol. 22, pp. 87-105, 2001.
- ZIVOT E., ANDREWS D., Further evidence on the great crash, the oil price shock and the unit root hypothesis, *Journal of Business and Economic Statistics*, Vol. 10, pp. 251-270, 1992.

Index

A

Analyse harmonique, 132
Autocorrélation partielle, 89

B, C

Box-Pierce, 92
Bruit Blanc, 85, 139
Buys-Ballot, 9, 21, 261
CENSUS, 39
Coefficient
– d'autocorrélation, 15, 19, 88
– de corrélation, 15
– de détermination, 244
– de lissage, 55
– de variation, 24
Corrélogramme, 15, 88, 157
Critère
– d'estimation bayésien, 254
– d'information de Akaike, 254
– de Schwarz, 254

E

Effet Joseph, 284
Effet Noé, 284

F, G, H

Fenêtre de Parzen, 136
Fenêtre de Tukey-Hanning, 136, 137, 141
Fonction d'autocorrélation, 87
Fonction d'autocorrélation
– inverse, 227
– étendue, 231
Fonction de gain, 124
Fonction de phase du filtre, 121, 124
Fonction de transfert, 121, 122

I, J, K, L

Inversibilité, 102, 103, 109
Kurtosis, 250, 302
Lissage de Holt, 65
Lissage exponentiel, 51
Lissage exponentiel de Holt-Winters, 73
Lissage exponentiel double, 59
Lissage exponentiel généralisé, 65
Lissage exponentiel simple, 52, 130, 200
Ljung et Box, 92

M, N, O, P

Marche au hasard, 155
Méthode TRAMO-SEATS, 44
Méthologie de Box-Jenkins, 261
Modèle GARCH, 310
Modèle logistique, 49
Moindres carrés conditionnés, 241
Moyenne mobile, 37, 40, 127
Opérateur retard, 98
Processus
– aléatoire, 83
– ARIMA, 198
– ARFIMA, 82, 281, 284
– ARMA, 82, 99, 143, 215
– ARMA saisonniers, 107
– Bruit Blanc, 85
– DS, 154, 159
– SARFIMA, 282
– TS, 154, 157
Puissance spectrale, 145

R, S

Random Walk Model, 155

- Runs Test*, 246
 SARIMA, 199
 Schéma
 - additif, 20, 24, 38, 40, 75
 - multiplicatif, 20, 41, 74
 Skewness, 250
 Stationnarité, 102, 103, 109
 - faible, 85
 - forte, 84
 Statistique
 - de Lo, 288
 - « Rescaled Range », 286
 - « U » de Theil, 253
 Stratégies de tests, 179
 Stratégie des tests de Dickey-Fuller, 170
- T, V, W**
- Test
 - BDS, 293, 294
 - de Box et Pierce, 91, 245
 - de Dickey et Pantula, 179
 - de Dickey-Fuller, 163
 - de Dickey-Fuller augmentés, 169
 - de Durbin-Watson, 70, 245
 - de Elliot, Rothenberg et Stock, 181
 - de Fisher, 11
 - de Franses, 194, 196
 - de Goldfeld-Quandt, 249
 - de Hylleberg, Engle, Granger et Yoo, 193
 - de Jarque et Bera, 250
 - de Ljung et Box, 246
 - de mémoire longue, 286
 - de normalité, 250
 - de nullité de la moyenne des résidus, 244
 - de Phillips-Perron, 178
 - de racines unitaires, 162
 - de rang, 47
 - de stabilité des résidus, 247
 - de Von Neumann's, 247
 - des signes, 47
 - d'existence d'une tendance, 245
 - d'hypothèses jointes, 167, 172, 174
 - d'un coefficient d'autocorrélation, 91
 - du « CUSUM », 247
 - du multiplicateur de Lagrange, 249
 - du Turning Point, 46
 - KPSS, 180
 - Ng-Perron, 182
 Transformation Box-Cox, 241
 Variables dichotomiques, 30
 Variables indicatrices, 29
 Wold, 96
 Yule-Walker, 90, 207, 208, 209, 216

Régis Bourbonnais
Michel Terraza

ANALYSE DES SÉRIES TEMPORELLES

Applications à l'économie et à la gestion

Cette 3^e édition, mise à jour des développements les plus récents et enrichie de nouveaux exercices, traite de manière pédagogique les techniques – classiques et modernes – d'analyse des séries temporelles. Elle répond aux questions suivantes :

- Quelles sont les méthodes de prévision des ventes ?
- Que sont un lissage exponentiel et la méthodologie de Box-Jenkins ?
- Comment procéder à l'analyse spectrale ?
- Que sont les tests de racine unitaire et comment les utiliser ?
- Pourquoi recourir aux processus ARFIMA ou ARCH ?

À forte incidence mathématique, réputée technique, l'analyse des séries temporelles trouve des applications diverses en **macroéconomie, finance, marketing...**

Un **site Internet** permet au lecteur de télécharger les séries statistiques et les programmes de traitement utilisés.



3^e édition

**Manuel et
exercices
corrigés**

RÉGIS BOURBONNAIS
est Maître de Conférences à
l'Université Paris-Dauphine
et chercheur au LEDa
(Laboratoire Économie
Dauphine).

MICHEL TERRAZA
est Professeur
à l'Université Montpellier I
et chercheur au Lameta.

- L3 et Master
de Sciences
économiques
et de gestion
- Écoles de commerce
et d'ingénieurs
- Économistes
d'entreprise,
chercheurs



9 782100 549351

6690580

ISBN 978-2-10-054935-1

