

2m11.2622.3

Université de Montréal

**Modèle Autorégressif avec seuils pour séries chronologiques à
valeurs entières non-négatives**

Par

Zoubida Ben seddik

Département de mathématiques et de statistique

Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures

en vue de l'obtention du grade de Maître ès sciences (M.Sc.)

en statistique

mars 1998

© Zoubida Ben seddik, 1998



QA

3

U54

1998

V.014

Université de Montréal

Méthode d'enseignement avec scellés pour séries chronologiques
valeurs entières non-négatives

*

1998

Yves Robitaille

Département de mathématiques et de statistiques

Faculté des arts et des sciences

Mémoire présenté à la Faculté des arts et des sciences

en vue de l'obtention du grade de maître en sciences (M.Sc.)

en mathématiques

mai 1998

Yves Robitaille



Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

Modèle Autorégressif avec seuils pour séries chronologiques à valeurs entières non-négatives

présenté par

Zoubida Ben seddik

a été évalué par un jury composé des personnes suivantes:

M. Martin Bilodeau

(président du jury)

M. Serge Tardif

(directeur de recherche)

M. Alain Latour

(codirecteur de recherche)

M. Robert Cléroux

(membre du jury)

Mémoire accepté le: 15.06.1998

TABLE DES MATIÈRES

Table des matières	i
Liste des programmes, tableaux et figures	iv
Lecture des expressions par acronymes	vi
Dédicace	vii
Remerciements	viii
INTRODUCTION	1
CHAPITRE 1: Processus autorégressif à valeurs entières	4
1.1 Introduction	4
1.2 Notions préliminaires relatives au processus autorégressif à valeurs entières	7
1.2.1 Fonction génératrice des probabilités	10
1.2.2 Quelques propriétés de la fonction génératrice	11
1.2.3 Quelques propriétés de l'opérateur $\alpha \circ$ généralisé	13
1.2.4 Premier et deuxième moments	15
1.2.5 La stationnarité des modèles autorégressifs	21
CHAPITRE 2: Processus autorégressif à valeurs entières auto-excitant	27
2.1 Introduction	27
2.2 Modèles autorégressifs à seuils	28
2.3 Représentations des régimes d'un $AREAX(\ell; p, d)$	31

2.3.1	Représentation des régimes par un processus <i>GENAR</i> (p)	32
2.3.2	Représentation des régimes par un <i>AR</i> (p)	32
2.4	Convergence forte de la moyenne et des estimateurs des paramètres	38
2.4.1	Estimation de la moyenne	40
2.4.2	Estimateur des moindres carrés de $\alpha_i^{(j)}$	41
2.4.3	Convergence presque-sûre des estimateurs des paramètres $\alpha_i^{(j)}$	42
2.5	Convergence presque-sûre	44
2.5.1	Convergence presque sûre des paramètres	44
2.5.2	Convergence presque sûre de la moyenne	46
2.6	Distribution asymptotique des paramètres	46
CHAPITRE 3: Test de la linéarité		50
3.1	Introduction	50
3.2	Statistique du test	51
3.2.1	Autorégression ordonnée et résidus de prévision	51
3.2.2	Estimation par récurrence des estimateurs des moindres carrés	54
3.3	Test de la linéarité	56
3.4	Identification de la variable seuil	60
3.4.1	Sélection du paramètre de retard d	61
3.4.2	Localisation des valeurs de seuils	62

3.5 Modélisation des <i>AREAX</i>	64
CHAPITRE 4: Simulations	66
4.1 Méthodologie des simulations	66
4.1.1 Structure des programmes et simulations	67
4.1.2 Production des séries	67
4.2 Distribution des modèles simulés	68
4.3 Résultats relatifs à la distribution asymptotique	70
4.4 Simulations d'un processus <i>AREAX</i> (2;1,2)	73
4.5 Existence d'un seuil dans le modèle <i>AREAX</i> (2;1,2)	74
CHAPITRE 5: Application	76
5.1 Étude expérimentale et problématique	76
5.2 Identification du modèle <i>AR</i> (<i>p</i>)	77
5.3 Sélection du retard <i>d</i>	78
5.4 Test de non-linéarité	82
5.5 Détermination de la valeur du seuil	84
CHAPITRE 6: Conclusion	91
BIBLIOGRAPHIE	133

LISTE DES TABLEAUX ET FIGURES

ANNEXE I: Programme ayant servi aux simulations	95
Programme ayant servi à simuler un modèle autorégressif à valeurs entières généralisé	96
Programme ayant servi à simuler un modèle autorégressif à valeurs entières généralisé à seuil	101
ANNEXE II: Résultat des simulations	106
Tableau II.1: Test de l'hypothèse nulle (modèle (4.2))	107
Tableau II.2: Test de l'hypothèse nulle (modèle (4.3))	107
Tableau II.3: Test de l'hypothèse nulle (modèle (4.4))	108
Tableau II.4: Test de l'hypothèse nulle (modèle (4.5))	108
Tableau II.5: Test de l'hypothèse nulle (modèle (4.6))	109
FigureII.1: Graphique de la série simulée	110
FigureII.2: Histogramme des valeurs observées de la statistique	111
FigureII.3: Polygone de fréquence des valeurs observées de la statistique	112
ANNEXE III: Programme servant aux simulations du jeu de donnée pour l'application	113
Programme ayant servi à produire le jeu de données pour l'application	114
Programme ayant servi à l'identification du modèle autorégressif	117
Programme ayant servi à la sélection du paramètre du retard	118
Programme ayant servi au calcul des rapports standardisés	120
Programme d'ajustement de la régression du premier régime	123

Programme d'ajustement de la régression du deuxième régime	124
ANNEXE IV: Résultat de l'application	125
Tableau IV.1: Liste des autocorrélations partielles	126
Figure IV.2: Graphiques des rapports standardisés en fonction des valeurs de la variable seuil	127
Figure IV.3: Graphiques des rapports standardisés en fonction de l'indice des observations ordonnées	128
Figure IV.4: Graphiques des estimations en fonction des valeurs de la variable seuil	129
Figure IV.5: Liste des résidus de la première régression	130
Figure IV.5: Liste des résidus de la deuxième régression	132

ACRONYMES UTILISÉS

Pour plus de commodité, le nom des modèles et d'autres expressions ont été abrégés et on doit comprendre:

AR : Processus AutoRégressif.

ARMA : Processus AutoRégressif–Moyennes Mobiles, (Moving Average and Autoregressive process).

ENAR : Processus AutoRégressif à valeurs ENtières.

GENAR: Processus AutoRégressif à valeurs ENtières Généralisé.

TAR : Processus autorégressif à seuils, (Thresholds AutoRegressive models).

SETAR: Processus Autorégressif Auto–excitant, (Self Exciting–Thresholds AutoRegressive models).

TARSO: (Open loop–Threshold AutoRegressive model).

STAR : (Smooth Threshold AutoRegressive model).

AREAX: Processus AutoRégressif à valeurs Entières généralisé Auto–eXcitant.

$\alpha \bullet$: Opérateur original de Steutel et van Harn.

$\alpha \circ$: Opérateur de Steutel et van Harn généralisé.

$(A)^{\tau}$: Transposée de la matrice (ou du vecteur) A .

DÉDICACE

À la mémoire de ceux que j'aime tellement:

ma grand-mère Aïcha

qui m'a constamment soutenue
dans le meilleur et dans le pire

et

mon frère Abdellatif



REMERCIEMENT

Je tiens à témoigner mes remerciements les plus sincères à mes directeurs de recherche: monsieur Serge Tardif pour son grand professionnalisme. Ses remarques et critiques, si précises et constructives ont amélioré amplement la qualité de ce travail. Monsieur Alain Latour qui m'a supportée et conseillée par ses remarques, suggestions et critiques, sans lesquelles je n'aurais pu réaliser ce travail. Je le remercie aussi pour sa patience.

Une attention plus particulière et un grand merci à ma famille qui, tout au long de mes études, m'a apporté compréhension, support et amour. J'insiste pour accorder une attention plus particulière à mon oncle. Mes encouragements à mes soeurs et frères: Fatima-Zohra, Hourya et Khalid.

Sur le plan informatique, le professeur André Plante m'a accordé au début de cette recherche de nombreuses discussions sur le module IML de SAS. Pour cette raison, à son égard, j'éprouve une profonde reconnaissance et lui dis "Merci". Au laboratoire, le support technique et l'expertise de Gisèle Legault est à mentionner. Merci à Léonce qui m'a aidé au début de mes études à comprendre le logiciel Splus.

Un merci particulier à la famille Lavoie, spécialement Normand et sa mère Rita, ainsi qu'à la famille Trudeau, pour leur attention et leur support moral dont ils ont fait preuve à mon endroit au début de cette recherche. À leur égard, j'éprouve une profonde reconnaissance...

Je ne pourrais passer sous silence la gentillesse de tous ceux et celles qui m'ont remonté le moral dans les moments difficiles, juste par leur présence. Un grand merci à madame Lucette Gourley qui a su me soulager et reconforter dans une circonstance particulière.

Je m'en voudrais de ne pas remercier plus particulièrement une grande personne à qui je dois beaucoup et qui, par sa bonté, sa sagesse, son courage, son indépendance, ses défis et ses sacrifices pour ceux qu'elle a aimés "inconditionnellement" m'a inspirée et surtout m'a émerveillée. J'ai même envie de dire à tout le monde quelle personne extraordinaire elle était!. Cette personne est ma grand-mère que j'ai perdue pendant la rédaction de ce mémoire et qui a tant attendu mon retour. Là où elle est, je lui dis mille fois merci et m'excuse pour les hauts et les bas qu'elle a dû vivre pour nous ou à cause de nous....Les prières du monde entier ne me suffiront jamais à satisfaire le désir de vouloir lui dire merci à ma façon. Sa présence son sourire et son amitié me manquent énormément...et même son bon café et ses bons repas. Pour moi tu demeures encore "vivante" et je **t'aimerai** pour toujours "maman".

INTRODUCTION

Historiquement, les séries chronologiques servent principalement à modéliser les fluctuations d'un phénomène dont les observations ont été recueillies pendant un certain laps de temps. Puisqu'il arrive souvent que les observations suffisamment rapprochées dans le temps soient corrélées entre elles, les techniques statistiques faisant appel à l'indépendance des observations ne peuvent s'appliquer. L'approche utilisant les séries chronologiques consiste à caractériser la dépendance par modélisation; une fois que le modèle ajusté est jugé satisfaisant, en plus de permettre une meilleure compréhension du phénomène, il servira entre autres à prévoir les valeurs futures de la série.

Vers les années 70, divers problèmes ont conduit au développement d'une classe de modèles: les modèles autorégressifs–moyennes mobiles [*ARMA*], qui se révélèrent faciles d'emploi car ces modèles sont fondés sur une écriture de la valeur présente d'une série comme fonction linéaire de ses valeurs passées et de la valeur présente d'un bruit, s'interprétant comme l'innovation de la série. Cette formulation présente cependant divers inconvénients. Elle est fondamentalement linéaire, ce qui restreint automatiquement le type de dynamiques pour lesquelles elle fournit de bons ajustements. Comme elle est généralement mise en oeuvre en imposant peu de contraintes sur les paramètres autorégressifs, l'interprétation plus structurelle des phénomènes n'est pas assurée. Parmi les domaines d'application où la formulation *ARMA* classique se révèle insuffisante figurent des problèmes de santé comme l'analyse de certaines épidémies, des problèmes économiques et

financiers et d'autres où on peut voir la présence de certaines caractéristiques d'un système dynamique non-linéaire, dont la plus importante est le fait que la variabilité de la série dépend de façon importante de certaines observations de la série elle-même appelées seuils. Les modèles à seuils introduits par Tong (1978) fournissent un cadre plus adapté à ce type de problèmes.

Ce mémoire reflète implicitement l'importance accordée à ce type de modèles du point de vue de la théorie statistique et de la pratique. Il s'intéressera aux modèles non-linéaires, plus particulièrement, à des séries chronologiques à valeurs entières avec points de rupture.

Le premier chapitre introduit l'opérateur de Steutel et van Harn généralisé et présente des définitions, propositions et propriétés relatives à cet opérateur. Certains modèles faisant intervenir cet opérateur sont aussi définis. Certaines notions relatives aux modèles autorégressifs standards et généralisés sont aussi introduites dans ce chapitre.

Dans le deuxième chapitre nous introduisons les modèles à seuils. En substituant l'opérateur généralisé de Steutel et van Harn au produit par un scalaire dans le modèle autorégressif auto-excitant à valeurs réelles dénoté *SETAR* (Tong et Lim (1980)), nous définissons un processus autorégressif à valeurs entières auto-excitant que nous dénotons *AREAX* et nous examinons la structure de ce processus. Dans la littérature, celle-ci a surtout été étudiée lorsque les processus étaient à valeurs réelles. Certaines propriétés probabilistes et statistiques sont établies.

Au chapitre trois, nous allons développer un test statistique qui permettra de vérifier l'hypothèse de la linéarité locale des *AREAX*. Notre approche se veut analogue à celle préconisée par Tsay (1989) pour l'analyse des séries chronologiques à valeurs réelles.

Enfin, au chapitre 4, après avoir décrit brièvement le pourquoi de nos simulations, divers aspects techniques en vue d'effectuer l'étape expérimentale par simulation sont regroupés. On étudie la qualité des approximations asymptotiques du test proposé en comparant les résultats expérimentaux de différents modèles aux valeurs théoriques correspondantes à certains niveaux critiques.

Le chapitre 5 sera consacré à l'application de l'approche poursuivie dans cette étude en suivant en détail la procédure de Tsay (1989) afin de détecter la valeur des seuils qui est l'étape la plus importante et la plus difficile de cette méthodologie. Finalement, on termine le chapitre par une conclusion où l'on recommande l'emploi plus fréquent des séries chronologiques non-linéaires car la formulation des modèles *ARMA* classiques se révèle insuffisante dans plusieurs domaines d'application.

La lecture de ce mémoire nécessite une connaissance de base des principales méthodes d'inférence statistique (estimation et test) et des propriétés des processus autorégressifs.

CHAPITRE 1

LE PROCESSUS AUTORÉGRESSIF À VALEURS ENTIÈRES GÉNÉRALISÉ

1.1 Introduction

Un contexte d'application de la statistique souvent rencontré est celui dit des séries chronologiques, où l'on s'intéresse au développement d'une suite d'observations, prises sur une même variable, à différents instants dans le temps. Dans plusieurs champs d'activité, travailler avec des séries à l'intérieur desquelles il

y a une dépendance fait partie du quotidien. En général, on étudie l'évolution d'un phénomène dans le temps et on cherche à le comprendre et à le caractériser.

Tout au long de cette dernière décennie, des modèles tels que le modèle linéaire, le modèle de régression et le modèle *ARMA*, ont été largement développés et utilisés. Néanmoins, l'hypothèse de linéarité des données s'est avérée insuffisante à étudier, décrire et comprendre de nombreux phénomènes. Nous rappelons que les modèles linéaires sont connus par leur développement assez régulier tandis que dans la réalité, nous sommes souvent confrontés à des séries ayant des développements explosifs, comparables à celui d'un système dynamique bruité à cycle limité, avec phénomène d'oscillations. Guégan (1994) souligne que le concept de cycle limité a été introduit par Poincaré et développé par Andronov dans de nombreux exemples où l'on pouvait observer des oscillations périodiques.

C'est donc pour aborder la modélisation de cette réalité qu'en pratique, l'analyse des séries chronologiques non-linéaires a pris de l'ampleur et est devenue de plus en plus l'intérêt de plusieurs auteurs et chercheurs.

Quant aux séries chronologiques à valeurs entières, –qui font l'objet de notre recherche– elles sont utilisées dans plusieurs problèmes des domaines de la santé, des sciences économiques, de la sécurité routière et autres. Plusieurs modèles suggérés pour les séries chronologiques à valeurs entières étaient basés sur l'opérateur $\alpha \bullet$ de Steutel et van Harn (1979), voir la section 1.1 de ce chapitre. Par exemple, cet opérateur est utilisé par Al-Osh et Alzaid (1987) et Alzaid et Al-Osh (1990), par Du et Li (1991), par Dion, Gauthier et Latour (1993) et, sous une forme plus générale, par Gauthier et Latour (1994) qui ont redéfini un processus

autorégressif à valeurs entières généralisé dont la description sera donnée à la section 1.2. Dans ce dernier modèle, on dénote l'opérateur généralisé par $\alpha \circ$, au lieu de $\alpha \bullet$ utilisé pour l'opérateur original de Steutel et van Harn. On réfère au modèle faisant intervenir l'opérateur $\alpha \circ$ au moyen de l'acronyme *GENAR* (acronyme de l'expression: AutoRégressif à valeurs ENtières Généralisé).

Ce mémoire s'intéressera donc aux modèles non-linéaires, plus particulièrement, à des séries chronologiques à valeurs entières avec points de rupture. Nous allons substituer l'opérateur généralisé $\alpha \circ$ au produit par un scalaire dans le modèle autorégressif à seuils auto-excitant dénoté par *SETAR* (Tong et Lim (1980)). Le modèle *SETAR* est un modèle dérivé du modèle à seuils dénoté par *TAR*, (Tong (1978)). Nous le représenterons dans le cas des valeurs entières, par l'équation (2.1) du chapitre deux. L'objectif de notre recherche est de définir un modèle **AutoRégressif à valeurs Entières à seuils Auto-eXcitant** que nous allons dénoter par *AREAX*. La forme de ce modèle sera expliquée à la seconde section du chapitre deux. Le modèle suggéré sera une version analogue du modèle à seuils auto-excitant du type de celui considéré par Tong et Lim (1980) et par Tsay (1989) pour des séries chronologiques à valeurs réelles. Par un seuil, on entend une valeur de la série elle-même et qui produit un changement de régime lorsqu'on la franchit.

Au chapitre trois, nous allons développer un test statistique qui permettra de tester l'hypothèse de la linéarité à l'intérieur des séries chronologiques à valeurs entières à seuils. Notre approche pour tester la non-linéarité se veut analogue à celle préconisée par Tsay (1989) pour l'analyse de la non-linéarité des séries chronologiques à valeurs réelles.

L'objectif de ce test de non-linéarité est de confronter l'hypothèse de l'inexistence des seuils, c'est-à-dire :

H_0 : il n'y a pas de points de rupture,

à une contre-hypothèse stipulant l'existence d'au moins un seuil. Ce test est basé sur la régression des résidus des prévisions d'horizon un par rapport aux X_t du modèle qui sera défini par l'équation (2.3) du chapitre deux.

Ce premier chapitre sera consacré aux rappels des notions qui aideront à la compréhension des processus autorégressifs à valeurs entières. On y retrouve des définitions des processus autorégressifs à valeurs entières et des propriétés utiles à certains calculs.

1.2 Notions préliminaires relatives au processus autorégressif à valeurs entières

Dans cette section, nous allons présenter des modèles pour les séries chronologiques à valeurs entières non-négatives.

Pour amorcer cette présentation, nous allons définir l'opérateur de Steutel et van Harn (1979), ainsi que celui généralisé par Gauthier et Latour (1994) et Latour (1997).

Dans ce qui suit, X et Z dénotent des variables aléatoires à valeurs entières non-négatives, tandis que α et β sont des réels, éléments de l'intervalle $[0, 1[$.

Définition 1.2.1 L'opérateur $\alpha \bullet$ de Steutel et van Harn

Soient X une variable aléatoire à valeurs entières non-négatives et $\{Y_j\}_{j \in \mathbb{N}}$ une suite de variables de Bernoulli indépendantes et identiquement distribuées de moyenne α et indépendantes de X . L'opérateur $\alpha \bullet$ de Steutel et van Harn, appliqué à la variable aléatoire X , est défini par :

$$\alpha \bullet X = \begin{cases} \sum_{j=1}^X Y_j & \text{si } X > 0 \\ 0 & \text{si } X = 0. \end{cases}$$

La suite $\{Y_j\}_{j \in \mathbb{N}}$ est dite suite dénombrante.

Bien que la notion « \bullet » suggère une simple multiplication, il ne faut pas agir comme tel. Nous donnons plus tard dans ce chapitre, certaines propriétés importantes de cet opérateur.

Soit $\{X_t\}_{t \in \mathbb{Z}}$ une suite de variables aléatoires à valeurs entières non-négatives. Alors, selon Du et Li (1991), $\{X_t\}_{t \in \mathbb{Z}}$ est un processus autorégressif à valeurs entières, noté $ENAR(p)$, s'il satisfait l'équation suivante:

$$X_t = \sum_{j=1}^p \alpha_j \bullet X_{t-j} + \varepsilon_t, \quad t \in \mathbb{Z},$$

où $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ est une suite de variables aléatoires identiquement distribuées à valeurs entières non-négatives de moyenne μ_ε et variance σ_ε^2 .

Chacune des suites $\{Y_j\}_{j \in \mathbb{N}}$ sous-entendues dans $\alpha_k \bullet X_{t-k}$, pour $k = 1, 2, \dots, p$, sont indépendantes entre elles et indépendantes de ε_t . Cette approche permet d'obtenir une structure de corrélation de type autorégressif.

Dans la suite, nous suggérons une définition plus générale de l'opérateur de Steutel et van Harn dont les variables impliquées dans la somme intervenant dans l'application des opérateurs sont des variables à valeurs entières ayant leurs deux premiers moments finis, indépendantes entre elles et indépendantes de ε_t . Nous n'exigeons pas qu'elles soient de type Bernoulli, ce qui distingue nettement nos résultats de ceux de Du et Li (1991).

Remarque.

Nous utiliserons la notation $\alpha \circ$ au lieu de $\alpha \bullet$ afin de distinguer l'opérateur généralisé de l'opérateur original.

Définition 1.2.2 L'opérateur généralisé $\alpha \circ$

Soient X une variable aléatoire à valeurs entières non-négatives et Y une variable aléatoire de moyenne finie α , où $0 < \alpha < 1$. Soit $\{Y_j\}_{j \in \mathbb{N}}$, une suite de variables aléatoires indépendantes entre elles, indépendantes de X et distribuées selon la même loi que Y . Alors, l'opérateur généralisé $\alpha \circ$ est défini par:

$$\alpha \circ X = \begin{cases} \sum_{j=1}^X Y_j & \text{si } X > 0 \\ 0 & \text{si } X = 0. \end{cases}$$

À la définition 1.2.5, utilisant ce nouvel opérateur, nous donnerons la redéfinition d'un processus ayant une structure de corrélations identique à celle d'un processus autorégressif standard à valeurs réelles, (Gauthier et Latour (1994)).

1.2.1 Fonction génératrice des probabilités

Nous allons maintenant introduire la fonction génératrice des probabilités. Il s'agit d'un outil permettant d'établir facilement l'égalité en distribution de deux variables aléatoires. Elle permet aussi l'obtention des moments par simple dérivation. Par conséquent, elle est d'une grande utilité lorsque nous nous intéressons aux propriétés de l'opérateur et en particulier lors du calcul des moments de $\alpha \circ X$.

Définition 1.2.3 Fonction génératrice des probabilités

La fonction génératrice des probabilités de X , notée Φ_X , est définie par:

$$\Phi_X(s) = E[s^X], \quad s > 0.$$

Dans ce qui suit, nous donnons quelques propriétés de cette fonction et nous insistons sur la fonction génératrice des probabilités de la loi binomiale qui est particulièrement importante puisque si Y admet une loi de Bernoulli de paramètre α , alors, conditionnellement au fait que $X = x$, $\alpha \bullet X$ est de loi binomiale de paramètres (x, α) .

1.2.2 Quelques propriétés de la fonction génératrice

La proposition 1.1 établit le lien entre la distribution de la variable $\alpha \circ X$ et les distributions de X et de Y .

Proposition 1.1

Soient X et Y telles qu'introduites à la définition 1.2.2. Alors,

$$\Phi_{\alpha \circ X}(s) = \Phi_X(\Phi_Y(s)).$$

Démonstration

Soit $\{Y_j\}_{j \in \mathbb{N}}$, une suite de variables aléatoires indépendantes et identiquement distribuées.

$$\Phi_{\alpha \circ X}(s) = \sum_{k=0}^{\infty} s^k \Pr(\alpha \circ X = k) = \sum_{k=0}^{\infty} s^k \Pr\left(\sum_{i=1}^X Y_i = k\right).$$

En conditionnant par rapport à X ,

$$\Phi_{\alpha \circ X}(s) = \sum_{x=0}^{\infty} \left(\sum_{k=0}^{\infty} s^k \Pr\left(\sum_{i=1}^x Y_i = k\right) \right) \Pr(X = x) = \sum_{x=0}^{\infty} \Phi_{\sum_{i=1}^x Y_i}(s) \Pr(X = x).$$

Comme les Y_i sont indépendantes et identiquement distribuées,

$$\Phi_{\alpha \circ X}(s) = \sum_{x=0}^{\infty} \prod_{i=1}^x \Phi_{Y_i}(s) \Pr(X = x) = \sum_{x=0}^{\infty} (\Phi_Y(s))^x \Pr(X = x) = \Phi_X(\Phi_Y(s)).$$

cqfd.

Dans la suite, la proposition 1.2 donne la distribution de la variable $\alpha \circ X + \beta \circ X$, qui est la somme de deux variables transformées indépendamment, c'est-à-dire, que les suites de variables aléatoires intervenant dans $\alpha \circ X$ et dans $\beta \circ X$ sont indépendantes entre elles.

Proposition 1.2

Soient X une variable aléatoire à valeurs entières non-négatives, Y une variable aléatoire à valeurs entières non-négatives de moyenne finie α et de variance finie λ et W une variable aléatoire à valeurs entières non-négatives de moyenne finie β et de variance finie γ . Alors:

$$\Phi_{\alpha \circ X + \beta \circ X}(s) = \Phi_X(\Phi_Y(s)\Phi_W(s)).$$

Démonstration

En conditionnant par rapport à X ,

$$\begin{aligned} \Phi_{\alpha \circ X + \beta \circ X}(s) &= \sum_{k=0}^{\infty} s^k \Pr(\alpha \circ X + \beta \circ X = k) = \sum_{x=0}^{\infty} \Pr(X = x) \sum_{k=0}^{\infty} s^k \Pr(\alpha \circ x + \beta \circ x = k) \\ &= \sum_{x=0}^{\infty} \Pr(X = x) \Phi_{\alpha \circ x + \beta \circ x}(s). \end{aligned}$$

Puisque $\alpha \circ x$ et $\beta \circ x$ sont indépendantes, pour une valeur x fixe,

$$\Phi_{\alpha \circ x + \beta \circ x}(s) = \sum_{x=0}^{\infty} \Pr(X = x) \Phi_{\alpha \circ x}(s) \Phi_{\beta \circ x}(s) = \sum_{x=0}^{\infty} \Pr(X = x) \Phi_Y^x(s) \Phi_W^x(s),$$

car, x étant fixe, $\alpha \circ x$ est une somme de x variables Y_i et $\beta \circ x$ est une somme de x variables W_j . Finalement,

$$\Phi_{\alpha \circ X + \beta \circ X}(s) = \sum_{x=0}^{\infty} \left(\Pr(X = x) (\Phi_Y(s) \Phi_W(s))^x \right) = \Phi_X(\Phi_Y(s) \Phi_W(s)).$$

cqfd.

Notons que, si $\Phi_X(s) = \Phi_Z(s)$, pour tout $s > 0$, alors X et Z sont égales en distribution. Nous dénotons ce fait par $X \stackrel{D}{=} Z$.

1.2.3 Quelques propriétés de l'opérateur $\alpha \circ$ généralisé

La proposition suivante établit la distributivité de l'opérateur généralisé $\alpha \circ$ sur la somme de deux variables aléatoires.

Proposition 1.3

Soient X et Z deux variables aléatoires indépendantes à valeurs entières non-négatives, de moyenne et de variance finies. Alors,

$$\alpha \circ (X + Z) \stackrel{D}{=} \alpha \circ X + \alpha \circ Z.$$

Notez que les deux suites de variables indépendantes sous-entendues dans $\alpha \circ X$ et $\alpha \circ Z$ sont indépendantes entre elles.

Démonstration

On sait que $\Phi_{\alpha \circ X}(s) = \Phi_X(\Phi_Y(s))$. Puisque X et Z sont indépendantes, nous avons,

$$\Phi_{\alpha \circ (X+Z)}(s) = \Phi_{X+Z}(\Phi_Y(s)) = \Phi_X(\Phi_Y(s)) \Phi_Z(\Phi_Y(s))$$

$$= \Phi_{\alpha \circ X}(s) \Phi_{\alpha \circ Z}(s) = \Phi_{\alpha \circ X + \alpha \circ Z}(s),$$

car $\alpha \circ X$ et $\alpha \circ Z$ sont également indépendantes.

cqfd.

Notons qu'en règle générale, $\beta \circ [\alpha \circ X] \stackrel{D}{\neq} [\alpha\beta] \circ X$. Illustrons ce fait par un contre-exemple, en considérant le cas de variables de Bernoulli. Dans les sommes, conditionnellement au fait que $X = x$, le terme de droite $U = (\alpha\beta) \circ x$ est distribué selon une loi binomiale, de paramètres $(x, \alpha\beta)$. Le terme de gauche $V = \beta \circ (\alpha \circ x)$ est distribué selon une loi binomiale, de paramètres (aléatoire pour le premier) $\left(\sum_{i=1}^x Y_i, \beta \right)$. Puisque l'égalité en distribution des deux termes ne se réalise pas pour ce cas, on peut donc conclure généralement à leur inégalité en distribution. Il est clair alors que $\beta \circ [\alpha \circ X] \stackrel{D}{\neq} [\alpha\beta] \circ X$.

Notons aussi que l'opérateur $\alpha \circ$ n'est pas distributif sur l'addition:

$$\alpha \circ X + \beta \circ X \stackrel{D}{\neq} (\alpha + \beta) \circ X.$$

Nous pouvons démontrer cela en donnant un contre-exemple faisant intervenir des variables aléatoires de Bernoulli, consistant tout particulièrement à additionner deux binomiales, disons $B(1, \alpha)$ et $B(1, \beta)$. La somme ne donne pas une binomiale à moins que $\alpha = \beta$. Nous voulons donc vérifier si $\alpha \circ X + \beta \circ X \stackrel{D}{=} (\alpha + \beta) \circ X$. Pour la fonction génératrice des probabilités du terme de gauche, nous avons, en vertu de la proposition 1.2:

$$\begin{aligned}\Phi_{\alpha \circ X + \beta \circ X}(s) &= \Phi_X(\Phi_Y(s)\Phi_W(s)) \\ &= \Phi_X((1 - \alpha + \alpha s)(1 - \beta + \beta s)),\end{aligned}$$

car Y est $B(1, \alpha)$ et W est $B(1, \beta)$. La fonction génératrice du membre de droite est

$$\begin{aligned}\Phi_{(\alpha + \beta) \circ X}(s) &= \Phi_X(1 - (\alpha + \beta) + (\alpha + \beta)s) \\ &= \Phi_X(1 - \alpha - \beta + \alpha s + \beta s).\end{aligned}$$

Il est clair que, règle générale, $\Phi_{(\alpha + \beta) \circ X}(s) \neq \Phi_{\alpha \circ X + \beta \circ X}(s)$, ce qui démontre l'affirmation.

1.2.4 Premier et deuxième moments

Nous exprimons ici les deux premiers moments de la variable aléatoire $\alpha \circ X$ en fonction des moments de la variable X .

Soient X une variable aléatoire à valeurs entières non-négatives et Y une variable aléatoire de moyenne finie α et de variance finie λ . Soit $\{Y_j\}_{j \in \mathbb{N}}$, une suite de variables aléatoire indépendantes entre elles, indépendantes de X et distribuées selon la même loi que Y . Les propositions suivantes sont satisfaites.

Proposition 1.4

On a

$$E[\alpha \circ X] = \alpha E[X].$$

Démonstration

Pour toute fonction génératrice des probabilités, nous avons

$$E[X] = \Phi'_X(1) = \frac{d}{ds} \Phi_X(s) \Big|_{s=1}.$$

Puisque $(\Phi_{\alpha \circ X}(s))' = (\Phi_X(\Phi_Y(s)))' = \Phi'_X(\Phi_Y(s))\Phi'_Y(s)$, il s'ensuit que

$$E[\alpha \circ X] = \Phi'_{\alpha \circ X}(1) = \Phi'_X(\Phi_Y(1))\Phi'_Y(1) = \alpha E[X].$$

cqfd.

Proposition 1.5

On a

$$E[(\alpha \circ X)^2] = \lambda E[X] + \alpha^2 E[X^2].$$

Démonstration

Pour toute fonction génératrice des probabilités, nous avons

$$\Phi''_X(1) = \frac{d}{ds} \Phi'_X(s) \Big|_{s=1} = E[X^2] - E[X].$$

Il s'ensuit que $E[X^2] = \Phi''_X(1) + \Phi'_X(1)$. Ainsi

$$\begin{aligned} (\Phi_{\alpha \circ X}(s))'' &= \{\Phi_X(\Phi_Y(s))\}'' = \{\Phi'_X(\Phi_Y(s))\Phi'_Y(s)\}' \\ &= \Phi''_X[\Phi_Y(s)]\{\Phi'_Y(s)\}^2 + \Phi'_X(\Phi_Y(s))\Phi''_Y(s). \end{aligned}$$

Cette seconde dérivée évaluée en 1 donne:

$$\begin{aligned}
\Phi_X''(1)(\Phi_Y'(1))^2 + \Phi_X'(1)\Phi_Y''(1) &= (E[X^2] - E[X])\alpha^2 + E[X](\lambda + \alpha^2 - \alpha) \\
&= \alpha^2 E[X^2] - \alpha^2 E[X] + \lambda E[X] + \alpha^2 E[X] - \alpha E[X] \\
&= \lambda E[X] + \alpha^2 E[X^2] - \alpha E[X]
\end{aligned}$$

et finalement

$$\begin{aligned}
E[(\alpha \circ X)^2] &= \Phi_{\alpha \circ X}''(1) + E[\alpha \circ X] = \Phi_{\alpha \circ X}''(1) + \alpha E[X] \\
&= \lambda E[X] + \alpha^2 E[X^2].
\end{aligned}$$

cqfd.

Proposition 1.6

On a

$$Var(\alpha \circ X) = \lambda E[X] + \alpha^2 Var(X).$$

Démonstration

En vertu des deux propositions précédentes,

$$\begin{aligned}
Var(\alpha \circ X) &= E[(\alpha \circ X)^2] - (E[\alpha \circ X])^2 = \lambda E[X] + \alpha^2 E[X^2] - \alpha^2 (E[X])^2 \\
&= \lambda E[X] + \alpha^2 Var(X).
\end{aligned}$$

cqfd.

Proposition 1.7

Si X et Z sont deux variables aléatoires distinctes et si les deux suites de variables aléatoires indépendantes sous-entendues dans $\alpha \circ X$ et $\beta \circ Z$ sont indépendantes entre elles, alors

$$\text{Cov}(\alpha \circ X, \beta \circ Z) = \alpha\beta\text{Cov}(X, Z).$$

Démonstration

Soient $\{Y_i\}_{i \in \mathbb{N}}$ une suite de variables aléatoires indépendantes et identiquement distribuées, de moyenne α et $\{W_j\}_{j \in \mathbb{N}}$ une suite de variables aléatoires indépendantes et identiquement distribuées de moyenne β . On a

$$\begin{aligned} \text{Cov}(\alpha \circ X, \beta \circ Z) &= E[(\alpha \circ X)(\beta \circ Z)] - \alpha\beta E[X]E[Z] \\ &= E\left[\sum_{i=1}^X Y_i \sum_{j=1}^Z W_j\right] - \alpha\beta E[X]E[Z]. \end{aligned}$$

En conditionnant par rapport à X et Z simultanément, il vient

$$\begin{aligned} \text{Cov}(\alpha \circ X, \beta \circ Z) &= \sum_{x=0}^{\infty} \sum_{z=0}^{\infty} \left(E\left[\sum_{i=1}^x \sum_{j=1}^z Y_i W_j\right] \Pr(X=x \text{ et } Z=z) - \alpha\beta E[X]E[Z] \right) \\ &= \alpha\beta E[XZ] - \alpha\beta E[X]E[Z] = \alpha\beta\text{Cov}(X, Z). \end{aligned}$$

cqfd.

Remarque.

La formule $Var(\alpha \circ X) = Cov(\alpha \circ X, \alpha \circ X)$ n'est pas un cas particulier de la proposition précédente car dans ce cas, les deux suites $\{Y_i\}_{i \in N}$ et $\{W_j\}_{j \in N}$ ne sont pas indépendantes.

Nous attirons l'attention du lecteur sur le fait que les résultats démontrés dans Brillinger (1969) permettent de retrouver ces résultats et fournissent une méthode de calcul basée sur le calcul des cumulants au moyen des cumulants conditionnels.

Dans ce qui suit, nous donnons la définition du modèle utilisant l'opérateur de Steutel et van Harn (1979).

Définition 1.2.4 Processus $ENAR(p)$

Soient $\{X_t\}_{t \in \mathbb{Z}}$ une suite de variables aléatoires à valeurs non-négatives, $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ une suite de variables aléatoires indépendantes et identiquement distribuées à valeurs entières non-négatives, $p \in \mathbb{N}$ et $\{\alpha_k\}_{k \in \{1, 2, \dots, p\}}$ un ensemble de constantes telles que $\forall k \in \{1, 2, \dots, p-1\}$, $\alpha_k \in [0, 1[$, $\alpha_p \in]0, 1[$ et $\sum_{k=1}^p \alpha_k < 1$. Alors $\{X_t\}_{t \in \mathbb{Z}}$ est un processus $ENAR(p)$, (acronyme de l'expression: AutoRégressif à valeurs ENtières), s'il satisfait l'équation suivante :

$$X_t = \sum_{k=1}^p \alpha_k \bullet X_{t-k} + \varepsilon_t. \quad (1.1)$$

Toutes les suites $\{Y_i\}_{i \in \mathbb{N}}$ sous-entendues dans les termes $\alpha_k \bullet X_{t-k}$, $k = 1, 2, \dots, p$, sont indépendantes entre elles et indépendantes de ε_t .

L'existence de ce processus et ses propriétés de base ont été mises en évidence par Du et Li (1991).

Une extension naturelle est donc obtenue en substituant l'opérateur $\alpha \circ$ à l'opérateur $\alpha \bullet$ dans l'équation (1.1).

En utilisant le nouvel opérateur $\alpha \circ$, Gauthier et Latour (1994) ont défini un processus autorégressif à valeurs entières ayant une structure de corrélation identique à celle d'un processus autorégressif standard à valeurs réelles. Un tel processus est dit processus autorégressif à valeurs entières généralisé, dénoté par $GENAR(p)$.

Définition 1.2.5 Processus $GENAR(p)$

Soient $\{X_t\}_{t \in \mathbb{Z}}$ une suite de variables aléatoires à valeurs entières non-négatives, $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ une suite de variables aléatoires indépendantes et identiquement distribuées à valeurs entières non-négatives, de moyenne finie μ_ε et de variance finie σ_ε^2 , $p \in \mathbb{N}$ et $\{Y_i^{(k)}\}_{i \in \mathbb{N}}$, $k = 1, \dots, p$, des suites indépendantes de variables aléatoires indépendantes de moyenne α_k et de variance λ_k . Supposons $\{\alpha_k\}_{k \in \{1, 2, \dots, p\}}$ un ensemble de constantes telles que, pour tout $k \in \{1, 2, \dots, p-1\}$, $\alpha_k \in [0, 1[$, $\alpha_p \in]0, 1[$ et $\sum_{k=1}^p \alpha_k < 1$. Alors, $\{X_t\}_{t \in \mathbb{Z}}$ est un $GENAR(p)$ s'il satisfait l'équation suivante:

$$X_t = \sum_{k=1}^p \alpha_k \circ X_{t-k} + \varepsilon_t. \quad (1.2)$$

Il va de soi que toutes les suites $\{Y_i\}_{i \in \mathbb{N}}$ sous-entendues dans $\alpha_k \circ X_{t-k}$ pour $k = 1, 2, \dots, p$, sont indépendantes entre elles et indépendantes de ε_t . Elles sont de moyenne finie α_k et de variance finie λ_k .

Notons que la preuve de l'existence des processus $GENAR(p)$ est donnée par le lemme 3.1 de Latour (1997).

1.2.5 Stationnarité des modèles autorégressifs

La stationnarité d'un processus est une propriété importante lorsque nous effectuons de l'inférence statistique. Plusieurs résultats théoriques sont effectivement basés sur cette hypothèse. Nous distinguons deux types de stationnarité: la stationnarité au sens large et la stationnarité au sens strict.

Définition 1.2.6 La stationnarité au sens strict

Un processus $\{X_t\}_{t \in \mathbb{Z}}$ est dit stationnaire au sens strict si, peu importe $n \in \{1, 2, \dots\}$ et t_1, t_2, \dots, t_n dans \mathbb{Z} , la distribution conjointe de $X_{t_1+k}, X_{t_2+k}, \dots, X_{t_n+k}$ est indépendante de $k, k \in \mathbb{Z}$.

Définition 1.2.7 La stationnarité au sens large

Un processus $\{X_t\}_{t \in \mathbb{Z}}$ est dit stationnaire au sens large si pour tout $t \in \mathbb{Z}$ et $k \in \mathbb{Z}$,

i) $E[X_t] = \mu_X < \infty$,

ii) $Var(X_t) = \sigma_X^2 < \infty$ et

$$\text{iii) } \text{Cov}(X_t, X_{t-k}) = \gamma(k).$$

La stationnarité au sens large requiert donc que les deux premiers moments du processus $\{X_t\}_{t \in \mathbb{Z}}$ soient finis et indépendants de l'instant t . La fonction $\gamma(k)$ est dite fonction d'autocovariance du processus $\{X_t\}_{t \in \mathbb{Z}}$. Enfin, mentionnons que, si un processus est stationnaire au sens strict et si $E[X_t^2] < \infty$ pour tout $t \in \mathbb{Z}$, il est également stationnaire au sens large.

Tout au long de ce mémoire, lorsque nous parlons d'un processus stationnaire, nous faisons référence à la stationnarité au sens large à moins qu'il en soit spécifié autrement dans le texte.

La stationnarité au sens large définie précédemment est aussi parfois désignée dans la littérature sous le nom de la stationnarité du second ordre.

Proposition 1.8

Soit $\{X_t\}_{t \in \mathbb{Z}}$ un processus $GENAR(p)$ stationnaire au sens large. Alors:

$$E[X_t] = \mu_X = \mu_\varepsilon \left[1 - \sum_{k=1}^p \alpha_k \right]^{-1}.$$

Démonstration

En utilisant la proposition 1.5 et la stationnarité au sens large du processus, on trouve:

$$\mu_X = E \left[\sum_{k=1}^p \alpha_k \circ X_{t-k} + \varepsilon_t \right]$$

$$= \sum_{k=1}^p \alpha_k E[X_{t-k}] + \mu_\varepsilon.$$

Comme notre processus est stationnaire au sens large, nous pouvons donc écrire en isolant μ_X ,

$$\begin{aligned} \mu_X &= \mu_X \sum_{k=1}^p \alpha_k + \mu_\varepsilon \\ \mu_X &= \mu_\varepsilon \left[1 - \sum_{k=1}^p \alpha_k \right]^{-1}. \end{aligned}$$

cqfd.

Puisque $E[X_t]$ et μ_ε sont non-négatives, nous remarquons que la condition $\sum_{k=1}^p \alpha_k < 1$ est nécessaire pour que le processus soit stationnaire au sens large.

Proposition 1.9

Soit $\{X_t\}_{t \in \mathbb{Z}}$ un processus $GENAR(p)$ stationnaire au sens large et ayant pour fonction d'autocovariance $\gamma(\cdot)$. Alors, si $k \geq 1$, l'autocovariance de délai k satisfait l'équation suivante:

$$\gamma(k) = \sum_{i=1}^p \alpha_i \gamma(k-i).$$

Démonstration

Pour $k \geq 1$ nous avons

$$E[X_t X_{t-k}] = E \left[\left(\sum_{i=1}^p \alpha_i \circ X_{t-i} + \varepsilon_t \right) X_{t-k} \right]$$

$$\begin{aligned}
&= \sum_{i=1}^p E[(\alpha_i \circ X_{t-i})X_{t-k}] - (E[X_t])^2 \sum_{i=1}^p \alpha_i + (E[X_t])^2 \sum_{i=1}^p \alpha_i + E[\varepsilon_t X_{t-k}] \\
&= \sum_{i=1}^p \alpha_i \text{Cov}(X_{t-i}, X_{t-k}) + (E[X_t])^2 \sum_{i=1}^p \alpha_i + \mu_\varepsilon E[X_t]
\end{aligned}$$

Nous avons donc

$$\begin{aligned}
\text{Cov}(X_t, X_{t-k}) &= E[X_t X_{t-k}] - (E[X_t])^2 \\
&= \sum_{i=1}^p \alpha_i \text{Cov}(X_t, X_{t-k+i}) + (E[X_t])^2 \sum_{i=1}^p \alpha_i + \mu_\varepsilon E[X_t] - (E[X_t])^2 \\
&= \sum_{i=1}^p \alpha_i \gamma(k-i) + (E[X_t]) \left[E[X_t] \left(\sum_{i=1}^p \alpha_i - 1 \right) + \mu_\varepsilon \right]
\end{aligned}$$

En substituant $E[X_t]$ par $\mu_\varepsilon \left(1 - \sum_{i=1}^p \alpha_i \right)^{-1}$, on obtient alors

$$\gamma(k) = \sum_{i=1}^p \alpha_i \gamma(k-i).$$

cqfd.

Proposition 1.10

Soit $\{X_t\}_{t \in \mathbb{Z}}$ un processus $GENAR(p)$ stationnaire au sens large. Alors la variance de X_t est donnée par:

$$\gamma(0) = \mu_X \sum_{i=1}^p \lambda_i + \sum_{i=1}^p \alpha_i \gamma(i) + \sigma_\varepsilon^2.$$

Démonstration

$$\gamma(0) = \text{Var} \left(\sum_{i=1}^p \alpha_i \circ X_{t-i} \right) + \sigma_\varepsilon^2$$

$$= \sum_{i=1}^p \text{Var}(\alpha_i \circ X_{t-i}) + 2 \sum_{i=1}^{p-1} \sum_{j=i+1}^p \text{Cov}(\alpha_i \circ X_{t-i}, \alpha_j \circ X_{t-j}) + \sigma_\varepsilon^2$$

En utilisant les propositions 1.6 et 1.7,

$$\begin{aligned} \gamma(0) &= \sum_{i=1}^p [\lambda_i E[X_{t-i}] + \alpha_i^2 \text{Var}(X_{t-i})] + 2 \sum_{i=1}^{p-1} \sum_{j=i+1}^p \alpha_i \alpha_j \gamma(j-i) + \sigma_\varepsilon^2 \\ &= \mu_X \sum_{i=1}^p \lambda_i + \gamma(0) \sum_{i=1}^p \alpha_i^2 + 2 \sum_{i=1}^{p-1} \sum_{j=i+1}^p \alpha_i \alpha_j \gamma(j-i) + \sigma_\varepsilon^2. \end{aligned}$$

Par ailleurs, la proposition 1.9 entraîne

$$\gamma(0) \sum_{i=1}^p \alpha_i^2 + 2 \sum_{i=1}^{p-1} \sum_{j=i+1}^p \alpha_i \alpha_j \gamma(j-i) = \sum_{j=1}^p \sum_{i=1}^p \alpha_j \alpha_i \gamma(j-i) = \sum_{j=1}^p \alpha_j \gamma(j) = \sum_{i=1}^p \alpha_i \gamma(i).$$

On peut donc conclure que,

$$\gamma(0) = \mu_X \sum_{i=1}^p \lambda_i + \sum_{i=1}^p \alpha_i \gamma(i) + \sigma_\varepsilon^2.$$

cqfd.

L'autocorrélation de délai k , notée $\rho(k)$, d'un processus $\{X_t\}_{t \in \mathbb{Z}}$ $GENAR(p)$ stationnaire au sens large est donnée par:

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)}.$$

Notons que, pour un processus stationnaire au sens large, nous avons

$$\rho(X_t, X_{t+k}) = \rho(-k) = \rho(k), \text{ pour } k \in \mathbb{Z},$$

car $\gamma(-k) = \text{Cov}(X_t, X_{t+k}) = \text{Cov}(X_{t+k}, X_t) = \gamma(k)$.

La fonction d'autocorrélation satisfait:

$$\rho(k) = \begin{cases} 1 & \text{si } k = 0, \\ \sum_{i=1}^p \alpha_i \rho(k-i) & \text{si } k \geq 1. \end{cases}$$

Remarque: il est à noter que $\gamma(-k) = \gamma(k)$ entraîne $\rho(-k) = \rho(k)$.

Les autocorrélations satisfont les équations de Yule-Walker. Cette structure de corrélation est identique à celle du modèle $AR(p)$ à valeurs réelles.

Conclusion

Dans ce chapitre, nous avons introduit l'opérateur généralisé. Nous avons fait un rappel relativement à certaines propriétés de cet opérateur et aux modèles autorégressifs à valeurs entières. Il est clair que nous pouvons substituer l'opérateur généralisé au produit par un scalaire dans un modèle autorégressif à seuils auto-excitants classique afin d'obtenir un modèle autorégressif à valeurs entières à seuils. Ce modèle fera l'objet de notre recherche dans les chapitres qui suivent.

Quelles sont les relations possibles entre les modèles $GENAR(p)$ lorsqu'on travaille avec un modèle avec seuils?. Le chapitre 2 se propose de répondre à cette question. Nous allons y définir la notion de processus autorégressif à valeurs entières à seuils. Ainsi, nous allons développer certaines propriétés statistiques et probabilistes de ces modèles.

CHAPITRE 2

PROCESSUS AUTORÉGRESSIF À VALEURS ENTIÈRES AUTO-EXCITANT

2.1 Introduction

Les modèles autorégressifs à valeurs entières auto-excitants, dénotés *AREAX*, sont des modèles *GENAR* par morceaux, munis de seuils d'excitation les faisant basculer d'un régime à l'autre. Chaque régime est régi par un modèle *GENAR* dont la structure d'autocorrélation est bien connue.

Par une série auto-excitante, on entend une série où les changements de régimes sont déterminés par les valeurs passées de la série elle-même.

Notons que le problème considéré ici est relié à celui des points de rupture. Plusieurs références sont disponibles: Quandt (1960), Shaban (1980), Pole et Smith (1985) et Siegmund (1988). La différence ici est que les données sont sériellement corrélées.

Avant de présenter le modèle autorégressif à valeurs entières auto-excitants, nous introduisons la classe des modèles à seuils.

2.2 Modèles autorégressifs à seuils

Les modèles autorégressifs à seuils, dénotés *TAR* (Treshold autoregressive models), ont été proposés par Tong (1978) dans le cas des séries à valeurs réelles. Ce type de modèle ne fait pas partie de la famille des modèles linéaires. Il se caractérise souvent par des systèmes particuliers, à savoir, les cycles limités, les fréquences à amplitudes dépendantes entre elles et les phénomènes d'oscillations.

Le modèle autorégressif auto-excitant, dénoté *SETAR* (Self-Exciting TAR) est un modèle dérivé des *TAR*. Il a été introduit par Tong et Lim (1980). Il a aussi été traité par Tong (1990). Il s'agit d'un modèle dont l'utilisation a été faite dans de nombreuses applications présentant des phénomènes cycliques avec transition de phases ou changement de régimes.

À part le modèle *SETAR*, beaucoup d'autres modèles dérivés des modèles autorégressifs à seuils ont vu le jour. Les modèles *TARSO*, *TARCS*, *EMA* et *TMA*

en sont quelques exemples. Quant aux applications de certains de ces modèles, nous proposons au lecteur de voir Ozaki (1981), Chan (1988), Chan et Tong (1986) et Hili (1992). On peut noter en particulier le modèle à seuils non-linéaire introduit par Ozaki (1981) qui considère que, sur des parties du modèle, au lieu de modéliser un processus autorégressif, on peut trouver n'importe quelle fonction non-linéaire.

Comme nous l'avons déjà mentionné dans l'introduction du premier chapitre, en substituant l'opérateur généralisé $\alpha \circ$ défini au chapitre 1 au produit par un scalaire dans un modèle *SETAR* classique, nous définissons un modèle **AutoRégressif à valeurs Entières Auto-eXcitant** que nous dénotons par *AREAX*.

Définition 2.2.1 Le processus *AREAX*

Soit $(r_0, r_1, \dots, r_j, \dots, r_\ell)$ un ensemble de nombres entiers positifs et ordonnés, $0 = r_0 < r_1 < \dots < r_j < \dots < r_\ell = \infty$, et considérons les intervalles $R_j = [r_{j-1}, r_j[$, $j = 1, \dots, \ell$, déterminant une partition de N , l'ensemble des entiers non-négatifs.

Soit $\{X_t\}_{t \in \mathbb{Z}}$ une suite de variables aléatoires à valeurs entières non-négatives, satisfaisant

$$X_t = \sum_{j=1}^{\ell} I_{R_j}(X_{t-d}) \left\{ \sum_{i=1}^p \alpha_i^{(j)} \circ X_{t-i} + \varepsilon_t^{(j)} \right\}, \quad (2.1)$$

où chaque suite $\{\varepsilon_t^{(j)}\}_{t \in \mathbb{Z}}$, $j = 1, \dots, \ell$, en est une de variables entières non-négatives de moyenne finie $\mu_{\varepsilon^{(j)}}$ et de variance finie $\sigma_{\varepsilon^{(j)}}^2$, $p \in N$, et où $\{\alpha_k^{(j)}\}$, $k = 1, 2, \dots, p$, $j = 1, \dots, \ell$ est un ensemble de constantes tel que pour tout $k \leq p-1$, $\alpha_k^{(j)} \in [0, 1[$, $\alpha_p^{(j)} \in]0, 1[$ et $\sum_{k=1}^p \alpha_k^{(j)} < 1$, $j = 1, \dots, \ell$. Conformément à Tong (1978), le paramètre

d qui apparaît dans l'indice temporel de l'équation (2.1) sera dit: «**paramètre de retard**», puisqu'il est le délai nécessaire avant que le système bascule dans un nouveau régime. On suppose donc qu'à l'intérieur de chaque régime du modèle d'équation (2.1), le processus suit un modèle autorégressif à valeurs entières généralisé d'ordre p . Les nombres r_j , $j = 1, \dots, \ell$, sont appelés les seuils du modèle. Ce sont les valeurs de la série elle-même qui déterminent le passage d'un régime à l'autre. Par commodité, nous attribuons au modèle (2.1) l'acronyme $AREAX(\ell; p, d)$, où ℓ est le nombre de régimes séparés par $\ell - 1$ seuils distincts r_j , p dénote l'ordre autorégressif $GENAR$ et d est le paramètre de retard. Notons que $\alpha_k^{(j)}$, $j = 1, \dots, \ell$, est la moyenne des variables intervenant dans les opérateurs de l'équation (2.1). La fonction $I_{R_j}(X_{t-d})$ est la fonction indicatrice de l'ensemble R_j à savoir:

$$I_{R_j}(x) = \begin{cases} 1 & \text{si } x \in R_j \\ 0 & \text{si } x \notin R_j. \end{cases}$$

Dans notre étude, nous nous limiterons souvent au cas où $\ell = 2$, c'est-à-dire au cas où il n'y a que deux régimes et un seul seuil. Le processus $\{X_t\}_{t \in \mathbb{Z}}$ sera donc défini de la manière suivante:

$$X_t = \begin{cases} \sum_{i=1}^p \alpha_i^{(1)} \circ X_{t-i} + \varepsilon_t^{(1)} & \text{si } X_{t-d} < r, \\ \sum_{i=1}^p \alpha_i^{(2)} \circ X_{t-i} + \varepsilon_t^{(2)} & \text{si } X_{t-d} \geq r, \end{cases} \quad (2.2)$$

ou encore

$$X_t = \sum_{j=1}^2 I_{R_j}(X_{t-d}) \left[\sum_{i=1}^p \alpha_i^{(j)} \circ X_{t-i} + \varepsilon_t^{(j)} \right],$$

où $R_1 : \{x \in \mathbb{N}; x < r\}$ et $R_2 : \{x \in \mathbb{N}; x \geq r\}$.

Dans le cas des séries chronologiques à valeurs réelles, Tsay (1989) fait remarquer qu'en pratique, les modèles autorégressifs à seuils n'ont pas été couramment utilisés. Cela est dû, d'une part, à la difficulté d'identifier les variables seuils et d'estimer les valeurs associées et, d'autre part, au fait qu'une procédure de modélisation simple et valable n'était pas vraiment disponible: celle proposée par Tong et Lim (1980), était assez complexe.

2.3 Représentations des régimes d'un $AREAX(\ell; p, d)$

Pour commencer l'étude du modèle $AREAX$, nous allons tout d'abord faire une description de quelques propriétés relatives à ce modèle.

Soit $\{X_t\}_{t \in \mathbb{Z}}$ un processus autorégressif à valeurs entières à seuils tel que défini à la définition 2.2.1 par l'équation (2.1). Ce processus fait intervenir ℓ régimes à l'intérieur desquels les observations sont régies – comme nous le verrons plus tard – par un processus autorégressif «standard». Notons que le processus $\{X_t\}_{t \in \mathbb{Z}}$ est dit non-linéaire s'il existe au moins deux régimes dont les modèles linéaires sont différents.

2.3.1 Représentation des régimes par un processus $GENAR(p)$

Nous donnons tout d'abord des conditions d'existence d'un processus $GENAR(p)$ à l'intérieur de chaque régime. Considérons un processus $\{X_t\}_{t \in \mathbb{Z}}$ auto-excitant à valeurs entières positives, localement stationnaire. Le $j^{\text{ème}}$ régime du processus $\{X_t\}_{t \in \mathbb{Z}}$ est un $GENAR(p)$ si et seulement si, il existe une série $\{\varepsilon_t^{(j)}\}_{t \in \mathbb{Z}}$ de variables aléatoires à valeurs entières non-négatives, de moyenne $\mu_{\varepsilon^{(j)}} > 0$ et de variance $\sigma_{\varepsilon^{(j)}}^2 > 0$ et p opérateurs de Steutel et van Harn généralisés $\alpha_i^{(j)}$, $i = 1, \dots, p$, mutuellement indépendants tels que

$$X_t = \sum_{i=1}^p \alpha_i^{(j)} \circ X_{t-i} + \varepsilon_t^{(j)}, \quad j = 1, \dots, \ell. \quad (2.3)$$

Notons que pour un processus $AREAX$, le $j^{\text{ème}}$ régime a une moyenne constante $\mu_X^{(j)} = \mu_{\varepsilon^{(j)}} \left(1 - \sum_{i=1}^p \alpha_i^{(j)}\right)^{-1} > 0$ car $\sum_{k=1}^p \alpha_k^{(j)} < 1$, $j = 1, \dots, \ell$.

2.3.2 Représentation des régimes par un $AR(p)$

Considérons un régime donné, c'est-à-dire, fixons un entier j , $1 \leq j \leq \ell$, écrivons α_i et λ_i , $i = 1, \dots, p$, et σ_{ε}^2 au lieu de $\alpha_i^{(j)}$, $\lambda_i^{(j)}$, $j = 1, \dots, p$, et $\sigma_{\varepsilon^{(j)}}^2$ et intéressons-nous au processus autorégressif (pas nécessairement à valeurs entières) Y_t satisfaisant l'équation:

$$Y_t - \mu_Y = \sum_{k=1}^p \alpha_k (Y_{t-k} - \mu_Y) + \tilde{\varepsilon}_t, \quad t \in \mathbb{Z}, \quad (2.4)$$

où $\{\tilde{\varepsilon}_t\}_{t \in \mathbb{Z}}$ est un processus de bruit blanc de variance

$$\sigma_e^2 = \mu_Y \sum_{i=1}^p \lambda_i + \sigma_\varepsilon^2 = 1, 2, \dots, p.$$

où λ_i est la variance d'innovation et σ_ε^2 est la variance de la série $\{\varepsilon_t^{(j)}\}_{t \in \mathbb{Z}}$ associée au $j^{\text{ème}}$ régime. Dans son lemme 2.2, Latour (1997) a démontré que $\{Y_t\}_{t \in \mathbb{Z}}$ est un processus $AR(p)$ stationnaire et causal.

Soient $\gamma_k, k = 0, \dots, p$, les $p + 1$ premières autocovariances de ce processus. Les équations de Yule–Walker donnent la relation entre les paramètres et les autocovariances du modèle à savoir:

$$\gamma_k = \begin{cases} \sum_{i=1}^p \alpha_i \gamma_i + \sigma_e^2, & k = 0 \\ \sum_{i=1}^p \alpha_i \gamma_{k-i}, & 1 \leq k \leq p. \end{cases} \quad (2.5)$$

Dans certains cas, on peut exprimer les paramètres $\alpha_k, k = 1, \dots, p$, et σ_e^2 du modèle (2.4) en fonction des $\gamma_k, k = 0, \dots, p$. Si on suppose que $\gamma_0 > 0$, la matrice des coefficients des p dernières équations du système (2.5) est alors définie positive (voir Brockwell et Davis (1991), p. 167). Donc, la solution de ce sous-système est unique et en utilisant le système (2.5) avec $k = 0$, nous pouvons trouver σ_e^2 . Dans d'autres cas, on peut exprimer de manière unique les autocovariances $\gamma_k, k = 0, \dots, p$, en fonction des paramètres $\alpha_k, k = 1, \dots, p$, et σ_e^2 . Le lemme suivant nous donne un critère simple et suffisant concernant l'existence d'une solution unique du système d'équations (2.5).

Lemme 2.1

Soient α_i , $1 \leq i \leq p$, des nombres tels que $0 \leq \alpha_i < 1$ et tels que $\sum_{i=1}^p \alpha_i < 1$.

Le système d'équations (2.5), où les γ_k , $k = 0, \dots, p$, sont les inconnues, admet une solution unique en fonction des paramètres α_i et de σ_e^2 .

La démonstration de ce lemme est donnée par Latour (1997).

Proposition 2.2

Supposons que $\{Y_t\}$ et $\{X_t\}$ soient deux processus stationnaires de même moyenne non-nulle et de même fonction d'autocovariance $\gamma(\cdot)$ et supposons que $\{X_t\}$ soit un processus $ARMA(p, q)$. Alors, $\{Y_t\}$ est aussi un processus $ARMA(p, q)$.

La démonstration se trouve dans l'article de Latour (1997) qui s'est lui-même inspiré de l'exercice 3.19, p. 112 de Brockwell et Davis (1991).

Supposons que pour le $j^{\text{ème}}$ régime, $j = 1, \dots, \ell$, les $\alpha_i^{(j)}$, $i = 1, \dots, p$, soient des nombres réels non-négatifs tels que $\sum_{k=1}^p \alpha_k^{(j)} < 1$ et que $\lambda_i^{(j)}$, $i = 1, \dots, p$, et $\sigma_e^{2(j)}$

soient des nombres réels positifs. Nous pouvons aisément vérifier qu'à l'intérieur du $j^{\text{ème}}$ régime, nous avons une représentation $AR(p)$. En effet, si le processus $\{X_t\}_{t \in \mathbb{Z}}$ satisfait au modèle de l'équation (2.1), où chacun des régimes $j = 1, 2, \dots, \ell$, est un processus $GENAR(p)$ stationnaire, et si $k \in \mathbb{N}$, $k \neq 0$, alors, en vertu de la proposition 1.9 du premier chapitre, l'autocovariance de délai k satisfait:

$$\gamma_k^{(j)} = \sum_{i=1}^p \alpha_i^{(j)} \gamma_{k-i}^{(j)}.$$

Les équations de Yule–Walker nous donnent alors la relation entre les paramètres et les autocovariances à savoir:

$$\gamma_k^{(j)} = \begin{cases} \sum_{i=1}^p \alpha_i^{(j)} \gamma_i^{(j)} + \mu_X^{(j)} \sum_{i=1}^p \lambda_i^{(j)} + \sigma_\varepsilon^2, & k = 0 \\ \sum_{i=1}^p \alpha_i^{(j)} \gamma_{k-i}^{(j)}, & 1 \leq k \leq p. \end{cases} \quad (2.6)$$

Pour $k = 1, \dots, p$, les valeurs des paramètres $\alpha_k^{(j)}$ sont solutions du système d'équations (2.6).

Le lemme suivant nous donne un critère simple et suffisant concernant l'existence d'une solution unique du système d'équations (2.6).

Lemme 2.2

Pour chaque $j = 1, 2, \dots, \ell$, soient $\alpha_i^{(j)}$, $1 \leq i \leq p$, des nombres tels que $0 \leq \alpha_i^{(j)} < 1$ et tels que $\sum_{i=1}^p \alpha_i^{(j)} < 1$. Le système d'équations (2.6) où les autocovariances $\gamma_k^{(j)}$, $k = 0, \dots, p$, sont les inconnues, admet une solution unique en fonction des paramètres $\alpha_i^{(j)}$ et de σ_ε^2 .

Par le lemme 2.2, la solution du système d'équations (2.6) est unique. Donc, la fonction des autocovariances du processus $GENAR(p)$ représentant le $j^{\text{ème}}$ régime, est identique à la fonction d'autocovariance d'un processus $AR(p)$ classique.

Corollaire 2.1

Soit le processus $\{X_t\}_{t \in \mathbb{Z}}$ satisfaisant le modèle de l'équation (2.1) et étant régi par un $GENAR(p)$ stationnaire dans chaque régime $j = 1, 2, \dots, \ell$. Alors, chacun de ces régimes est un processus $AR(p)$ qu'on peut écrire comme suit:

$$X_t - \mu_X^{(j)} = \sum_{k=1}^p \alpha_k^{(j)} (X_{t-k} - \mu_X^{(j)}) + \tilde{z}_t^{(j)}, \quad t \in \mathbb{Z}. \quad (2.7)$$

où $\{\tilde{z}_t^{(j)}\}_{t \in \mathbb{Z}}$ est un bruit blanc de variance, $\sigma_j^2 = \mu_X^{(j)} \sum_{i=1}^p \lambda_i^{(j)} + \sigma_{\varepsilon}^{2(j)}$.

Démonstration

Puisque le processus $\{X_t\}_{t \in \mathbb{Z}}$ est régi par un processus $GENAR(p)$ représentant les régimes $j = 1, 2, \dots, \ell$, nous avons $\alpha_k^{(j)} \in [0, 1[$, pour $k = 1, \dots, p-1$, $\alpha_p^{(j)} \in]0, 1[$ et $\sum_{k=1}^p \alpha_k^{(j)} < 1$. Soit $\sigma_j^2 = \mu_X^{(j)} \sum_{i=1}^p \lambda_i^{(j)} + \sigma_{\varepsilon}^{2(j)}$. Chacun des régimes du processus $\{X_t\}_{t \in \mathbb{Z}}$ satisfaisant l'équation (2.7) est un processus $AR(p)$ stationnaire. Par les propositions 3.2 et 3.3 de Gauthier et Latour (1994), la fonction d'autocovariance du processus $GENAR(p)$ du $j^{\text{ème}}$ régime, satisfait l'équation (2.6). Par le lemme 2.2, la solution de ce système est unique. Donc, la fonction d'autocovariance d'un processus $GENAR(p)$ est identique à celle du processus Y_t , $AR(p)$. Par la proposition 2.2, nous pouvons donc conclure que chacun des régimes $j = 1, 2, \dots, \ell$ du modèle $AREAX(\ell; p, d)$, régis par un $GENAR(p)$, est un processus $AR(p)$ satisfaisant l'équation (2.7).

cqfd.

En d'autres termes, le processus $AREAX(\ell; p, d)$ est un processus $AR(p)$ par morceaux. Il est alors naturel d'appliquer la théorie et les méthodes développées pour le processus $AR(p)$ aux régimes régis par un processus $GENAR(p)$. Puisque $\mu_{\varepsilon^{(j)}} = \left(1 - \sum_{i=1}^p \alpha_i^{(j)}\right) \mu_X^{(j)}$, nous pouvons aussi écrire l'équation (2.7) sous la forme:

$$X_t = \sum_{k=1}^p \alpha_k^{(j)} X_{t-k} + \mu_{\varepsilon^{(j)}} + \tilde{\varepsilon}_t^{(j)}, \quad t \in \mathbb{Z}.$$

Définition 2.3.1 Coefficient de corrélation partielle $\phi_k^{(j)}$

Soit $\{X_t\}_{t \in \mathbb{Z}}$ un processus $GENAR(p)$ stationnaire au sens large, représentant les régimes du modèle $AREAX(\ell; p, d)$. Considérons, pour chaque j , $j = 1, \dots, \ell$, la matrice d'autocorrélations,

$$\mathbf{P}_k^{(j)} = \begin{pmatrix} 1 & \rho_1^{(j)} & \rho_2^{(j)} & \dots & \rho_{k-1}^{(j)} \\ \rho_1^{(j)} & 1 & \rho_1^{(j)} & \dots & \rho_{k-2}^{(j)} \\ \rho_2^{(j)} & \rho_1^{(j)} & 1 & \dots & \rho_{k-3}^{(j)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1}^{(j)} & \rho_{k-2}^{(j)} & \rho_{k-3}^{(j)} & \dots & 1 \end{pmatrix},$$

le vecteur d'autocorrélations $\mathbf{R}_{\sim k}^{(j)} = (\rho_1^{(j)}, \dots, \rho_k^{(j)})^\tau$ et un vecteur inconnu $\mathbf{\Phi}_{\sim k}^{(j)} = (\phi_{k1}^{(j)}, \dots, \phi_{kk}^{(j)})^\tau$. Le coefficient d'autocorrélation partielle de délai k , $\phi_{kk}^{(j)}$, $k \in N$, du $j^{\text{ème}}$ régime est alors la dernière composante du vecteur $\mathbf{\Phi}_{\sim k}^{(j)}$ qui est solution du système d'équations

$$\mathbf{P}_k^{(j)} \mathbf{\Phi}_{\sim k}^{(j)} = \mathbf{R}_{\sim k}^{(j)}.$$

N.B: nous désignerons dorénavant ce coefficient d'autocorrélation partielle $\phi_k^{(j)}$ au lieu de $\phi_{kk}^{(j)}$.

Remarque

Le coefficient de corrélation partielle de délai k d'un processus stationnaire représente le $k^{\text{ème}}$ coefficient à l'intérieur d'un processus $GENAR(k)$ dans le cas où nous faisons l'approximation du processus stationnaire original par un processus autorégressif d'ordre k . Une conséquence de la proposition 1.9 est que si la représentation des régimes $j = 1, 2, \dots, \ell$ est sous la forme d'un $GENAR(p)$, alors $\phi_k^{(j)} = 0$ lorsque $k > p$. Puisque le coefficient d'autocorrélation partielle de délai k est une certaine fonction des k premières autocorrélations, c'est-à-dire, $\phi_k^{(j)} = f(\rho_1^{(j)}, \dots, \rho_k^{(j)})$, disons, on définit tout naturellement son estimateur comme étant $\hat{\phi}_k^{(j)} = f(\hat{\rho}_1^{(j)}, \dots, \hat{\rho}_k^{(j)})$, où $\hat{\rho}_1^{(j)}, \dots, \hat{\rho}_k^{(j)}$ sont les estimateurs respectifs de $\rho_1^{(j)}, \dots, \rho_k^{(j)}$.

Puisqu'il y a convergence forte des $\hat{\rho}_i^{(j)}$ vers leur valeur théorique respective et que la solution du système d'équations est une fonction des $\hat{\rho}_i^{(j)}$ mesurable au sens de Borel, nous avons aussi la convergence forte des estimateurs des coefficients d'autocorrélation partielle.

2.4 Convergence forte de la moyenne et des estimateurs des paramètres

Nous présentons dans cette section certains résultats relatifs à l'estimation des paramètres. Notons que l'étude probabiliste des modèles à seuils a été entreprise par Chan (1988), Chan et Tong (1986) et Hili (1992). Des résultats concernant l'estimation des paramètres ont été obtenus par Petrucelli et Woolford (1984), Chan, Petrucelli, Tong et Woolford (1985), Hilli (1991) et par Pham Chan, et Tong (1991). Nous venons de voir dans les sections précédentes que le modèle

$AREAX(\ell; p, d)$ est localement linéaire. Par conséquent, les techniques bien connues des moindres carrés peuvent s'appliquer pour estimer les principaux paramètres.

Afin de comprendre un phénomène dans le temps et de le caractériser, on sait que la moyenne du processus est souvent utilisée: elle donne le niveau autour duquel oscillent les valeurs d'un processus. C'est pourquoi il est intéressant de l'estimer. Il est aussi possible de démontrer la convergence forte de la moyenne échantillonnale des régimes.

Si l'on observe un échantillon de taille n du processus, nous dénotons dans ce qui suit, par N_j le nombre d'observations réalisées sous le contrôle du $j^{\text{ème}}$ régime, où $j = 1, \dots, \ell$. Il va de soi que N_1, \dots, N_ℓ sont des variables aléatoires telles que $n = \sum_{j=1}^{\ell} N_j$. Nous allons de plus supposer que

$$\frac{N_j}{n} \xrightarrow[n \rightarrow \infty]{P} c_j, \quad j = 1, \dots, \ell, \quad (2.8)$$

où c_1, \dots, c_ℓ sont les fractions positives telles que $\sum_{j=1}^{\ell} c_j = 1$. Il en résulte qu'à la limite, le processus est tel qu'on revient infiniment souvent à chacun des ℓ régimes.

2.4.1 Estimation de la moyenne

Comme estimateur de la moyenne de chaque régime, il est naturel de considérer la moyenne arithmétique des observations faites à l'intérieur du $j^{\text{ème}}$ régime, c'est-à-dire,

$$\bar{X}^{(j)} = \frac{1}{N_j} \sum_{t=1}^n \mathbf{1}_{R_j}(X_t) X_t, \quad j = 1, \dots, \ell.$$

Évidemment, à l'intérieur du $j^{\text{ème}}$ régime, $\bar{X}^{(j)}$ est un estimateur asymptotiquement sans biais de $\mu_X^{(j)}$. Le théorème 2.1 qui suit en établit la convergence presque sûre vers $\mu_X^{(j)}$.

Théorème 2.1

Si $\sum_{k=0}^{\infty} |\gamma_k^{(j)}| < \infty$, alors $\bar{X}^{(j)}$ est un estimateur fortement convergent de $\mu_X^{(j)}$.

La démonstration de ce théorème est donnée par Gauthier et Latour (1994). Il est à noter que ce résultat est une conséquence directe de la convergence presque sûre de $\bar{X}^{(j)}$ dans le cas où les observations sont indépendantes, résultat démontré par Grimmett et Stirzaker (1983).

Dans ce qui suit, nous donnons une présentation simple concernant l'estimation classique d'un processus stochastique stationnaire.

2.4.2 Estimateur des moindres carrés de $\alpha_i^{(j)}$

Les composantes du vecteur des paramètres du modèle *AREAX* sont des paramètres inconnus qui doivent être estimés. Dans le contexte du modèle de type autorégressif, la méthode d'estimation la plus couramment utilisée est la méthode des moindres carrés qui, pour chaque $j = 1, \dots, \ell$, consiste à trouver le point $\hat{\alpha}^{(j)} = (\hat{\alpha}_1^{(j)}, \dots, \hat{\alpha}_p^{(j)}, \hat{\mu}_{\varepsilon^{(j)}})^{\tau}$ minimisant la somme des carrés des erreurs de prévision d'horizon 1. Puisque le modèle *AREAX* est un modèle localement linéaire, les techniques ordinaires des moindres carrés sont utilisées pour l'analyse de ce processus.

Cette sous-section a pour principal but d'établir la convergence forte des estimateurs des moindres carrés des paramètres $\alpha_i^{(j)}$, $i = 1, \dots, p$, et $\mu_{\varepsilon^{(j)}}$ à l'intérieur du modèle *AREAX*($\ell; p, d$), c'est-à-dire, dans chacun de ses régimes.

La démonstration de la convergence forte des estimateurs des autocovariances échantillonnales, de même que celle des autocorrélations échantillonnales d'un processus autorégressif à valeurs entières généralisé sont données par Gauthier et Latour (1994). Des résultats semblables apparaissent dans Du et Li (1991) pour le modèle utilisant l'opérateur original de Steutel et van Harn. De la convergence presque sûre de ces statistiques, nous pouvons déduire finalement la convergence presque sûre des estimateurs des paramètres $\alpha_i^{(j)}$ et de la moyenne de la «composante bruit» des régimes du processus *AREAX*($\ell; p, d$). Puisque la moyenne, les autocovariances et les autocorrélations de chacun des régimes sont reliées entre elles par des relations identiques à celles d'un modèle

$AR(p)$ standard, le lecteur peut tout simplement utiliser les estimateurs classiques des moindres carrés.

Nous commentons ici deux hypothèses concernant l'autocovariance.

1- Nous savons que l'hypothèse $\sum_{k=0}^{\infty} |\gamma_k| < \infty$ est très courante dans le domaine des séries chronologiques. Si elle est satisfaite, cela implique que l'autocovariance de délai k tend vers zéro au fur et à mesure que k croît, ce qui signifie que les observations X_t et X_{t+k} ne sont à peu près pas corrélées lorsqu'elles sont suffisamment éloignées.

2- Quant à l'hypothèse $\sum_{h=0}^{\infty} |\varphi_k(h)| < \infty$, où $\varphi_k(h) = Cov(X_t X_{t+k}, X_{t+h} X_{t+k+h})$, elle demande que les produits $X_t X_{t+k}$ et $X_{t+h} X_{t+k+h}$ ne soient à peu près pas corrélés lorsqu'ils sont suffisamment distants l'un de l'autre.

Dans la sous-section qui suit, nous présentons des résultats relatifs à l'estimation par la méthode des moindres carrés conditionnels de certains paramètres à l'intérieur du modèle $AREAX(\ell; p, d)$.

2.4.3 Convergence forte des paramètres $\alpha_1^{(j)}, \dots, \alpha_p^{(j)}$ et $\mu_{\varepsilon^{(j)}}$

Cette sous-section a pour principal but de traiter la convergence forte des estimateurs des moindres carrés des paramètres $\alpha_i^{(j)}$ dans le contexte du modèle $AREAX(p)$. La démonstration est basée sur les résultats de Klimko et Nelson (1978).

On suppose tout d'abord que les seuils, le nombre ℓ de seuils, l'ordre p et le retard d sont connus. Puisque la présentation de chaque régime du modèle autorégressif auto-excitant $AREAX(\ell; p, d)$ est sous la forme d'un $GENAR(p)$, on s'intéresse donc à l'estimation des paramètres des modèles $GENAR(p)$.

Considérons le vecteur $\underline{\Theta} = (\alpha_1^{(1)}, \dots, \alpha_p^{(1)}, \mu_{\varepsilon^{(1)}}; \dots; \alpha_1^{(\ell)}, \dots, \alpha_p^{(\ell)}, \mu_{\varepsilon^{(\ell)}})$ et, pour $t \geq p+1$, l'expression

$$\begin{aligned} g(\underline{\Theta}, X_1, X_2, \dots, X_{t-1}) &= E[X_t | X_1, X_2, \dots, X_{t-1}] \\ &= \sum_{j=1}^{\ell} \mathbf{I}_{R_j}(X_{t-d}) \left(\sum_{i=1}^p \alpha_i^{(j)} X_{t-i} + \mu_{\varepsilon^{(j)}} \right), \end{aligned}$$

car $E[\alpha \circ X | X] = \alpha X$. Considérons enfin la somme des carrés conditionnels

$$Q_n(\underline{\Theta}) = \sum_{t=p+1}^n \left(X_t - g(\underline{\Theta}, X_1, X_2, \dots, X_{t-1}) \right)^2. \quad (2.9)$$

Les estimateurs des moindres carrés conditionnels des paramètres sont donc obtenus en minimisant par rapport à $\underline{\Theta}$ la somme des carrés de l'équation (2.9).

L'estimateur $\hat{\underline{\Theta}}_n$ de $\underline{\Theta}$ est donc la solution du système à $\ell \times (p+1)$ équations

suivant:

$$\nabla Q_n(\underline{\Theta}) = \underline{0},$$

où ∇Q_n est le gradient de Q_n et $\underline{0}$ est un vecteur ne contenant que des zéros. Le résultat principal concernant ces estimateurs est fourni par le théorème suivant.

Théorème 2.2

Si $\gamma_0^{(j)} > 0$ et si $\sum_{h=0}^{\infty} |\varphi_k^{(j)}(h)| < \infty$, $j = 1, \dots, \ell$, alors la solution de $\nabla Q_n(\underline{\Theta}) = \underline{0}$ donne lieu à des estimateurs fortement convergents des composantes de $\underline{\Theta} = (\alpha_1^{(1)}, \dots, \alpha_p^{(1)}, \mu_{\varepsilon^{(1)}}; \dots; \alpha_1^{(\ell)}, \dots, \alpha_p^{(\ell)}, \mu_{\varepsilon^{(\ell)}})$.

Ce théorème se démontre en utilisant les lemmes 5.2 et 5.5 ainsi que le corollaire 5.1 de Gauthier et Latour (1994).

2.5 Convergence presque sûre

Nous allons ici préciser quelques résultats concernant la convergence en probabilité qui sont nécessaires pour l'inférence statistique.

2.5.1 Convergence presque sûre des paramètres

Petrucelli et Woolford (1984) ont fourni les estimateurs des moindres carrés des paramètres des modèles à seuils à valeurs réelles. La convergence presque sûre de ces estimateurs vers les vraies valeurs y fut aussi démontrée.

Considérons le modèle $AREAX(\ell; p, d)$ de l'équation (2.1) et, pour un échantillon de taille n , les ensembles $T_j = \{t: X_t \text{ est régie par le } j^{\text{ème}} \text{ régime}\}$, $j = 1, \dots, \ell$. Il va de soi que $\text{Card}(T_j) = N_j$. Désignons alors par $t_1^{(j)} < \dots < t_{N_j}^{(j)}$ l'énumération en ordre croissant des éléments de T_j . Notons enfin la matrice de design associée au $j^{\text{ème}}$ régime par $A^{(j)}$. En fait,

$$\mathbf{A}^{(j)} = \begin{pmatrix} X_{t_1-1} & \cdots & X_{t_1-p} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ X_{t_{N_j}-1} & \cdots & X_{t_{N_j}-p} & 1 \end{pmatrix}.$$

Par conséquent, la matrice associée au $j^{\text{ème}}$ régime est $(\mathbf{A}^{(j)})^\tau (\mathbf{A}^{(j)})$. Dénotons également les valeurs propres de cette matrice par $\lambda_1^{(j)} \geq \lambda_2^{(j)} \geq \dots \geq \lambda_{p+1}^{(j)}$. De plus, nous supposons qu'à l'intérieur du $j^{\text{ème}}$ régime,

$$\lambda_{p+1}^{(j)} \rightarrow \infty \text{ p. s. et } \ln \lambda_1^{(j)} = o(\lambda_{p+1}^{(j)}) \text{ p. s. ,} \quad (2.10)$$

quand $n \rightarrow \infty$. Sous le contrôle du $j^{\text{ème}}$ régime, $j = 1, \dots, \ell$, nous dénotons les estimateurs des moindres carrés de $(\alpha_1^{(j)}, \dots, \alpha_p^{(j)}, \mu_{\varepsilon^{(j)}})$ par $(\hat{\alpha}_1^{(j)}, \dots, \hat{\alpha}_p^{(j)}, \hat{\mu}_{\varepsilon^{(j)}})$.

Citons tout d'abord le théorème de Tsay (1989) relatif à la convergence presque sûre des paramètres d'un modèle $SETAR(p)$.

Théorème 2.3 (Tsay, 1989)

Supposons que $\{X_t\}_{t \in \mathbb{Z}}$ suive un modèle $SETAR(\ell; p, d)$, où les tailles N_j et les matrices $(\mathbf{A}^{(j)})^\tau (\mathbf{A}^{(j)})$ satisfont les conditions (2.8) et (2.10) respectivement. Pour ℓ et d donnés, ainsi que les valeurs de seuils fixées, les estimateurs des moindres carrés $\hat{\alpha}_i^{(j)}$, $i = 1, \dots, p$, et $\hat{\mu}_{\varepsilon^{(j)}}$ convergent presque sûrement vers $\alpha_i^{(j)}$, $i = 1, \dots, p$, et $\mu_{\varepsilon^{(j)}}$.

Démonstration

Dans le cas du modèle $SETAR(\ell; p, d)$, sous les conditions (2.8) et (2.10), le $j^{\text{ème}}$ régime de $\{X_t\}_{t \in \mathbb{Z}}$ est un processus autorégressif linéaire. La démonstration

se trouve dans l'article de Lai et Wei (1982), où la convergence presque sûre des estimateurs des moindres carrés est établie pour un processus stochastique.

Dans le cas du modèle $AREAX(\ell; p, d)$, la convergence presque sûre demeure vérifiée comme le stipule le prochain théorème.

Théorème 2.4

Supposons que $\{X_t\}_{t \in \mathbb{Z}}$ suive un modèle $AREAX(\ell; p, d)$ défini à l'équation (2.1) où les tailles N_j et les matrices $(\mathbf{A}^{(j)})^\tau (\mathbf{A}^{(j)})$ satisfont les conditions (2.8) et (2.10) respectivement. Supposons de plus que le nombre de régimes ℓ , le paramètre de retard d ainsi que les valeurs des seuils soient tous connus. Les estimateurs des moindres carrés $\hat{\alpha}_i^{(j)}$, $i = 1, \dots, p$, et $\hat{\mu}_{\varepsilon^{(j)}}$ convergent presque sûrement vers $\alpha_i^{(j)}$, $i = 1, \dots, p$, et $\mu_{\varepsilon^{(j)}}$.

2.5.2 Convergence presque sûre de la moyenne

On a vu à la sous-section 2.4.1 la convergence forte de la moyenne de chacun des régimes du modèle $AREAX(p)$ régissant les différents processus autorégressifs à valeurs entières généralisés. Selon Latour (1997), la fonction de densité spectrale est continue au voisinage de zéro et selon Hall et Heyde (1980), p. 194, $\bar{X}^{(j)}$ converge donc presque sûrement vers $\mu_X^{(j)}$.

2.6 Distribution asymptotique des paramètres

Grâce aux notations de la section 2.5, si nous posons $\underline{Y}^{(j)} = \left(X_{t_1^{(j)}}, \dots, X_{t_{N_j}^{(j)}} \right)^\tau$, nous pouvons écrire le modèle sous la forme matricielle

$$\underline{Y}^{(j)} = \mathbf{A}^{(j)} \underline{\alpha}^{(j)} + \underline{e}^{(j)}, \quad (2.11)$$

où $\underline{\alpha}^{(j)}$ est le vecteur des paramètres inconnus et $\underline{e}^{(j)}$ est un vecteur de bruit blanc. L'estimateur des moindres carrés conditionnels de $\underline{\alpha}^{(j)}$ est donné par:

$$\hat{\underline{\alpha}}^{(j)} = \left(\left(\mathbf{A}^{(j)} \right)^\tau \mathbf{A}^{(j)} \right)^{-1} \left(\mathbf{A}^{(j)} \right)^\tau \underline{Y}^{(j)}$$

et la matrice des covariances de $\hat{\underline{\alpha}}^{(j)}$ est $\sigma_j^2 \left(\left(\mathbf{A}^{(j)} \right)^\tau \left(\mathbf{A}^{(j)} \right) \right)^{-1}$. Les éléments de la matrice de schéma $\mathbf{A}^{(j)}$ de dimensions $N_j \times (p+1)$ et de plein rang, où $p+1 \leq N_j$, sont fixes. Lord (1990, pp. 41–42) a montré que

$$\frac{\left(\mathbf{A}^{(j)} \right)^\tau \left(\mathbf{A}^{(j)} \right)}{N_j} \xrightarrow[n \rightarrow \infty]{P} V_j, \quad (2.12)$$

où V_j est une matrice de dimensions $(p+1) \times (p+1)$ définie positive dont les éléments sont finis. Par conséquent, il s'ensuit que

$$\sqrt{N_j} \left(\hat{\underline{\alpha}}^{(j)} - \underline{\alpha}^{(j)} \right) \xrightarrow[n \rightarrow \infty]{D} \mathbf{N} \left(\underline{0}, \sigma_j^2 V_j^{-1} \right)$$

et

$$\frac{\left(\left(\mathbf{A}^{(j)} \right)^\tau \left(\mathbf{A}^{(j)} \right) \right)^{\frac{1}{2}} \left(\hat{\underline{\alpha}}^{(j)} - \underline{\alpha}^{(j)} \right)}{\sigma_j} \xrightarrow[n \rightarrow \infty]{D} \mathbf{N} \left(\underline{0}, \mathbf{I}_{p+1} \right), \quad (2.13)$$

(Freedman, 1981), où $\underline{0} = (0, \dots, 0)^\tau$ un vecteur de dimension $(p+1)$. Il est à noter que, si l'on diagonalise la matrice $\left(\mathbf{A}^{(j)} \right)^\tau \left(\mathbf{A}^{(j)} \right)$, c'est-à-dire, si l'on écrit $\left(\mathbf{A}^{(j)} \right)^\tau \left(\mathbf{A}^{(j)} \right) = \mathbf{P}^\tau \mathbf{\Lambda} \mathbf{P}$, où $\mathbf{\Lambda}$ est la matrice diagonale des valeurs propres associées à $\left(\mathbf{A}^{(j)} \right)^\tau \left(\mathbf{A}^{(j)} \right)$ et \mathbf{P} est la matrice des vecteurs propres normés, alors $\left[\left(\mathbf{A}^{(j)} \right)^\tau \left(\mathbf{A}^{(j)} \right) \right]^{\frac{1}{2}}$

est définie comme étant la matrice $P^\tau \Lambda_2^{-1} P$, où Λ_2^{-1} est la matrice diagonale des racines carrées des valeurs propres.

Il est bien connu que la variation résiduelle

$$S_j^2 = \frac{\left(\underline{Y}^{(j)} - \underline{A}^{(j)} \underline{\hat{\alpha}}^{(j)} \right)^\tau \left(\underline{Y}^{(j)} - \underline{A}^{(j)} \underline{\hat{\alpha}}^{(j)} \right)}{N_j - p - 1}$$

est un estimateur convergent de σ_j^2 . Ce dernier résultat et (2.13) entraînent alors que:

$$\frac{\left(\underline{\hat{\alpha}}^{(j)} - \underline{\alpha}^{(j)} \right)^\tau \left(\underline{A}^{(j)} \right)^\tau \left(\underline{A}^{(j)} \right) \left(\underline{\hat{\alpha}}^{(j)} - \underline{\alpha}^{(j)} \right)}{S_j^2} \xrightarrow[n \rightarrow \infty]{D} \chi_{p+1}^2,$$

ce qui permet de déterminer une région de confiance de niveau approximatif $100(1 - \alpha)\%$ pour $\underline{\alpha}^{(j)}$. Cette région est donnée par l'ellipsoïde:

$$R = \left\{ \underline{\alpha}^{(j)} : \left(\underline{\hat{\alpha}}^{(j)} - \underline{\alpha}^{(j)} \right)^\tau \left(\underline{A}^{(j)} \right)^\tau \underline{A}^{(j)} \left(\underline{\hat{\alpha}}^{(j)} - \underline{\alpha}^{(j)} \right) \leq S_j^2 \chi_{p+1; 1-\alpha}^2 \right\},$$

où $\chi_{v; 1-\alpha}^2$ dénote le quantile d'ordre $(1 - \alpha)$ d'une loi du khi-deux centrée avec v degrés de liberté. Si l'on dénote par $h_{ii}^{(j)}$ le $i^{\text{ème}}$ terme de la diagonale de la matrice $\left(\left(\underline{A}^{(j)} \right)^\tau \underline{A}^{(j)} \right)^{-1}$, un intervalle de confiance de niveau approximatif $100(1 - \alpha)\%$ pour une composante $\alpha_i^{(j)}$ de $\underline{\alpha}^{(j)}$ est donné par:

$$\hat{\alpha}_i^{(j)} \pm z_{1-\alpha/2} S_j \sqrt{h_{ii}^{(j)}},$$

où $S_j = \sqrt{S_j^2}$ et $z_{1-\alpha/2}$ dénote le quantile d'ordre $1 - \alpha/2$ d'une loi normale centrée réduite.

Conclusion

Jusqu'à présent, nous avons exposé le modèle $AREAX(\ell; p, d)$ et ses caractéristiques probabilistes et statistiques. Par les résultats obtenus dans ce chapitre, il est alors possible de faire de l'inférence et des tests d'hypothèses concernant la variable observée. Maintenant il serait intéressant de pousser l'étude encore plus loin en développant les propriétés relatives au test considéré par l'approche de Tsay (1989).

Au chapitre 3, nous explicitons alors la méthode d'estimation des paramètres. Plus spécifiquement, nous calculons la statistique permettant d'effectuer le test de la linéarité à l'intérieur des régimes d'un modèle $AREAX(\ell; p, d)$.

CHAPITRE 3

TEST DE NON-LINÉARITÉ

3.1 Introduction

Étant donné l'observation du processus $\{X_t\}_{t \in \mathbb{Z}}$ pour $t = 1, \dots, n$, un des premiers problèmes à résoudre est celui de l'ajustement d'un modèle. Cet ajustement soulève un certain nombre de difficultés, à savoir: l'identification d'un modèle adéquat dans lequel interviennent des hypothèses de linéarité ou de non-linéarité, l'estimation des paramètres et par conséquent, l'élaboration des tests d'adéquation.

Plusieurs types de tests ont été développés dans le contexte des modèles à seuils: les tests de Keenan (1985), Tsay (1986), basés sur les développements en série de Volterra et le test de McLeod et Li (1983), basé sur le test Portmanteau, en sont quelques exemples.

Dans ce chapitre, nous considérons le modèle *AREAX*. Notre approche se veut analogue à celle préconisée par Tsay (1989) pour tester la linéarité à l'intérieur des séries chronologiques à seuils à valeurs réelles. Nous allons donc préciser la statistique sur laquelle est basé le test. Les outils présentés dans ce chapitre sont utiles au moment de l'identification d'un modèle.

3.2 Statistique du test

L'approche suggérée repose sur l'estimation des paramètres par la méthode des moindres carrés. Des formules de récurrence pour obtenir les estimateurs des moindres carrés seront données ultérieurement. Ensuite, les résidus des prévisions d'horizon 1 sont calculés. Ils interviennent dans l'expression de la statistique F du test, dont la définition sera donnée à la section 3.3. Cette statistique est basée sur la somme des carrés des résidus de la prévision. Elle est l'outil principal pour l'identification de différentes linéarités à l'intérieur des régimes.

3.2.1 Autorégression ordonnée et les résidus de la prévision

Soit $\{X_t\}_{t \in \mathbb{Z}}$ un $GENAR(p)$ dont on a prélevé un échantillon de taille n . Ayant considéré la représentation d'un modèle autorégressif standard du processus $GENAR(p)$, nous avons la forme d'autorégression suivante,

$$X_t = (1, X_{t-1}, \dots, X_{t-p}) \underline{\alpha} + e_t.$$

pour $t = p+1, \dots, n$, où $\underline{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_p)^\tau$ est le vecteur $(p+1) \times 1$ des coefficients, α_0 étant le terme constant, et où e_t est un bruit blanc. Le vecteur ligne $(X_t, 1, X_{t-1}, \dots, X_{t-p})$ sera dit: «**unité de régression**».

Afin de simplifier l'exposé, considérons le modèle $AREAX(2; p, d)$ ne comportant qu'un seul seuil, donc deux régimes, et n observations (modèle de l'équation (2.2)). La variable seuil X_{t-d} peut prendre les valeurs dans l'ensemble des entiers positifs $H = \{X_h, \dots, X_{n-d}\}$, où d est le paramètre de retard et $h = \max\{1, p+1-d\}$. Dénotons par π_i l'indice temporel de la $i^{\text{ème}}$ plus petite observation de l'ensemble H . Soient $\underline{X} = (X_{\pi_1+d}, \dots, X_{\pi_{n-d-h+1}+d})^\tau$, $\underline{X}_v = (X_{\pi_1+d-v}, \dots, X_{\pi_{n-d-h+1}+d-v})^\tau$, où $v = 1, \dots, p$, et $\underline{M} = \begin{pmatrix} 1 & X_1 & \dots & X_p \\ \vdots & \vdots & \dots & \vdots \end{pmatrix}$ une matrice $(n-d-h+1) \times (p+1)$ dont la première colonne est constituée de 1. Ainsi, à la valeur X_{π_i+d} du vecteur \underline{X} correspond le vecteur ligne $(1, X_{\pi_i+d-1}, \dots, X_{\pi_i+d-p})$ de la matrice \underline{M} . Soit $\underline{\varepsilon}$ le vecteur colonne dont les éléments sont les bruits e_{π_i+d} . Nous aurons alors la forme matricielle suivante:

$$\underline{X} = \underline{M} \underline{\alpha} + \underline{\varepsilon}. \quad (3.1)$$

À partir de maintenant, nous supposons que l'ordre p du processus et le retard d sont fixés. Nous pouvons donc réécrire le modèle des observations ordonnées du modèle $AREAX(2; p, d)$ comme suit: pour $i = 1, \dots, n-d-h+1$,

$$X_{\pi_i+d} = \begin{cases} \alpha_0^{(1)} + \sum_{v=1}^p \alpha_v^{(1)} X_{\pi_i+d-v} + e_{\pi_i+d}^{(1)} & \text{si } i < s, \\ \alpha_0^{(2)} + \sum_{v=1}^p \alpha_v^{(2)} X_{\pi_i+d-v} + e_{\pi_i+d}^{(2)} & \text{si } i \geq s, \end{cases} \quad (3.2)$$

où s satisfait la condition: $X_{\pi_s} < r_1 \leq X_{\pi_{s+1}}$. Il s'agit des unités de régression ordonnées où les $s-1$ premières unités sont dans le premier régime. Quant aux autres, elles sont dans le second régime. Lorsque les unités de régression sont triées selon l'ordre croissant de la variable seuil X_{i-d} , nous dirons que l'équation (3.2) est une «**autorégression ordonnée**». Notons que les moyennes peuvent être différentes d'un régime à l'autre selon les valeurs des paramètres dans chaque régime. La moyenne de chacun des régimes est donnée selon le chapitre 1 comme suit:

$$\frac{\mu_{\varepsilon}^{(j)}}{1 - \sum_{v=1}^p \alpha_v^{(j)}}, \quad j = 1, 2 .$$

En triant les unités de régression selon les valeurs de la variable seuil, on insère dans un même groupe les observations d'un même régime. Il est évident que toutes les observations d'un régime suivent le même modèle $GENAR(p)$. Par contre, le modèle peut être différent d'un régime à l'autre. Notons que seulement le nombre d'observations dans chaque groupe dépend du seuil. En pratique, cette séparation ne nécessite pas la connaissance exacte de ce dernier. Si le seuil n'est pas connu, les observations seront toujours réparties en deux groupes. La valeur de s sera cependant inconnue et elle dépendra bien sûr du seuil.

Afin d'illustrer l'utilisation potentielle de cette autorégression ordonnée dans l'étude des modèles *AREAX*, nous allons donner l'heuristique du test proposé.

Considérons le système d'équations (3.2). Si nous connaissions la valeur du seuil r_1 , l'estimation convergente des paramètres s'obtiendrait facilement. Mais puisque cette valeur est inconnue, nous procédons d'une manière séquentielle. Supposons que le nombre d'observations N_1 du 1^{er} régime soit suffisamment grand pour que la théorie asymptotique vue au chapitre 2 soit à toutes fins pratiques vérifiée. Alors, les estimateurs des moindres carrés du premier régime, $\hat{\alpha}_v^{(1)}$, $v = 0, \dots, p$, sont convergents. Pour les observations appartenant au premier régime, les résidus de la prévision sont asymptotiquement un bruit blanc et sont, par conséquent, orthogonaux aux variables explicatives $\{X_{\pi_i+d-v} \mid i = 1, \dots, n-d-h+1; v = 1, \dots, p\}$. D'autre part, dès que i atteint ou excède s , nous nous retrouvons dans le deuxième régime. Les résidus de la prévision sont alors biaisés à cause du changement du régime au temps $\pi_{s+1} + d$. Il est facile ici de constater qu'alors, les résidus de la prévision sont une fonction de $\{X_{\pi_i+d-v} \mid i = 1, \dots, n-d-h+1; v = 1, \dots, p\}$. Par conséquent, il n'y a plus d'orthogonalité entre les résidus de la prévision de ce groupe d'observations et les variables explicatives X_{π_i+d-v} , $i = 1, \dots, n-d-h+1$, $v = 1, \dots, p$, car l'autorégression s'applique ici au groupe d'observations excédant la valeur du seuil r_1 .

3.2.2 Estimation par récurrence des estimateurs des moindres carrés

Nous allons maintenant décrire explicitement le caractère récursif de la procédure suggérée. Dans ce problème de régression, nous devons débiter l'estimation des paramètres au moyen d'un certain nombre d'observations b . Tsay

(1989) a choisi $b = \left[\frac{n}{10} \right] + p$, où $[z]$ désigne la partie entière du réel z . C'est cette valeur que nous avons retenue dans nos simulations.

Considérons l'autorégression ordonnée décrite par le système d'équations (3.2). Nous désignerons par $\hat{\alpha}_{\sim}^{[m]}$, $m \geq b$, les vecteurs des estimateurs des moindres carrés basés sur les m premières unités de régression. Avant de donner les équations de récurrence permettant le calcul de $\hat{\alpha}_{\sim}^{[m+1]}$ en fonction de $\hat{\alpha}_{\sim}^{[m]}$, fixons certaines notations. Dénotons par P_m l'inverse de la matrice de design associée aux m premières observations, c'est-à-dire $P_m = (M_m^T M_m)^{-1}$, où M_m est la matrice obtenue de M en ne considérant que les m premières lignes. Désignons par $x_{\sim m+1} = (1, X_{\pi_{m+1}+d-1}, \dots, X_{\pi_{m+1}+d-p})$ la $(m+1)^{\text{ème}}$ ligne de la matrice M . Alors, les estimateurs des moindres carrés s'obtiennent de manière récursive à l'aide des formules suivantes:

$$D_{m+1} = 1 + x_{\sim m+1} P_m x_{\sim m+1}^T,$$

$$K_{m+1} = P_m x_{\sim m+1}^T / D_{m+1},$$

$$\hat{\alpha}_{\sim}^{[m+1]} = \hat{\alpha}_{\sim}^{[m]} + K_{\sim m+1} \left[X_{\pi_{m+1}+d} - x_{\sim m+1} \hat{\alpha}_{\sim}^{[m]} \right]$$

et

$$P_{m+1} = \left(I - P_m \frac{x_{\sim m+1}^T x_{\sim m+1}}{D_{m+1}} \right) P_m.$$

(Voir Ertel et Fowlkes 1976; Goodwin et Payne 1977). Les résidus de la prévision et les résidus standardisés de la prévision sont respectivement

$$\hat{a}_{\pi_{m+1}+d} = X_{\pi_{m+1}+d} - x_{\sim m+1} \hat{\alpha}_{\sim}^{[m]}, \quad m = b+1, \dots, n - \hat{d} - h + 1, \quad (3.3)$$

et

$$\hat{e}_{\pi_{m+1}+d} = \frac{\hat{a}_{\pi_{m+1}+d}}{\sqrt{D_{m+1}}}. \quad (3.4)$$

Les résidus de la prévision peuvent aussi être utilisés pour déterminer la variable seuil. Notons que le problème abordé est relié au point de changement (switching-regression problem), pour lequel plusieurs références sont disponibles. Quandt (1960), Shaban (1980), Pole et Smith (1985) et Siegmund (1988) en sont quelques exemples. La différence ici est que les données sont sériellement corrélées.

3.3 Test de non-linéarité

Dans cette section, nous donnons les détails concernant le test de la non-linéarité du modèle.

Basée sur les considérations prémentionnées, une façon de tester la non-linéarité du processus *AREAX* est d'effectuer la régression des résidus de la prévision de l'autorégression ordonnée, à savoir, \hat{e}_{π_i+d} du modèle d'équation (3.2) par rapport aux variables explicatives X_{π_i+d-v} , $i = b+1, \dots, n-d-h+1$ et $v = 1, \dots, p$:

$$\hat{e}_{\pi_i+d}^{[b]} = \omega_0 + \sum_{v=1}^p \omega_v X_{\pi_i+d-v} + \varepsilon_{\pi_i+d}^{[b]}. \quad (3.5)$$

Si l'ordre p et le paramètre de retard d sont fixés, le nombre réel d'observations dans l'autorégression ordonnée est $n-d-h+1$, où $h = \max\{1, p+1-d\}$.

Suite à cette procédure, nous pouvons expliciter la statistique permettant de tester l'hypothèse nulle:

$$H_0 : \omega_0 = \omega_1 = \dots = \omega_p = 0.$$

Le test de linéarité du modèle (3.5) que nous allons suggérer est basé sur l'idée que, sous l'hypothèse H_0 , les résidus de la prévision d'horizon un sont asymptotiquement un bruit blanc et qu'ils sont orthogonaux aux variables explicatives $X_{\pi_i+d-\nu}$, $i = b+1, \dots, n-d-h+1$, $\nu = 1, \dots, p$. L'équation (3.5) exprime les résidus de la prévision comme une fonction linéaire des variables explicatives $X_{\pi_i+d-\nu}$, $i = b+1, \dots, n-d-h+1$, $\nu = 1, \dots, p$, où les paramètres $\omega_0, \omega_1, \dots, \omega_p$ sont inconnus. Si nous supposons que l'autorégression récursive commence par b observations (où l'entier b est tel que défini auparavant), nous aurons alors $n-d-b-h+1$ résidus de prévision disponibles. Regroupons l'ensemble des perturbations aléatoires, ainsi que l'ensemble des $(p+1)$ paramètres de l'équation (3.5) sous la forme de vecteurs que nous dénoterons respectivement par $\underline{\varepsilon}^{[b]} = (\varepsilon_{\pi_{b+1}+d}^{[b]}, \dots, \varepsilon_{\pi_{n-d-h+1}+d}^{[b]})^\tau$ et $\underline{\omega} = (\omega_0, \dots, \omega_p)^\tau$. Le modèle reliant les \hat{e}_{π_i+d} , $i = b+1, \dots, n-d-h+1$, aux variables explicatives $X_{\pi_i+d-\nu}$, $i = b+1, \dots, n-d-h+1$ et $\nu = 1, \dots, p$, s'écrit alors sous la forme matricielle

$$\underline{\hat{e}}^{[b]} = \mathbf{X} \underline{\omega} + \underline{\varepsilon}^{[b]}.$$

où $(\hat{e}_{\pi_{b+1}+d}, \dots, \hat{e}_{\pi_{n-d-h+1}+d})^\tau$ est le vecteur des valeurs de $\hat{e}^{[b]}$ et \mathbf{X} est la matrice de dimensions $((n-d-b-h+1) \times (p+1))$.

L'estimation des paramètres de $\underline{\omega}$ s'effectue par la méthode des moindres carrés consistant à minimiser par rapport à $\underline{\omega}$, la somme des carrés des erreurs notée *SCE*:

$$SCE(\underline{\omega}) = \left(\underline{\varepsilon}^{[b]} \right)^\tau \left(\underline{\varepsilon}^{[b]} \right) = \left(\hat{\underline{e}}^{[b]} - X \underline{\omega} \right)^\tau \left(\hat{\underline{e}}^{[b]} - X \underline{\omega} \right).$$

On sait que $\hat{\underline{\omega}} = (X^\tau X)^{-1} X^\tau \hat{\underline{e}}^{[b]}$ et il en résulte que la somme des carrés des résidus est

$$\left(\hat{\underline{\varepsilon}}^{[b]} \right)^\tau \left(\hat{\underline{\varepsilon}}^{[b]} \right) = SCE(\hat{\underline{\omega}}) = \left(\hat{\underline{e}}^{[b]} \right)^\tau \left(I - X(X^\tau X)^{-1} X^\tau \right) \hat{\underline{e}}^{[b]}.$$

La statistique permettant d'effectuer le test est calculée de la manière suivante:

$$\hat{F}_{(p,d)} = \frac{\left(\left(\hat{\underline{e}}^{[b]} \right)^\tau \left(\hat{\underline{e}}^{[b]} \right) - \left(\hat{\underline{\varepsilon}}^{[b]} \right)^\tau \left(\hat{\underline{\varepsilon}}^{[b]} \right) \right) / (p+1)}{\left(\hat{\underline{\varepsilon}}^{[b]} \right)^\tau \left(\hat{\underline{\varepsilon}}^{[b]} \right) / (n-d-b-p-h)}. \quad (3.6)$$

Les indices (p, d) de $\hat{F}_{(p,d)}$ sont utilisés pour indiquer que la statistique dépend des paramètres p et d . Cette formule constitue l'outil principal permettant de vérifier la linéarité du modèle d'équation (3.5) ainsi que la linéarité locale du modèle *AREAX*.

Théorème 3.1 de Tsay (1989)

Soit $\{X\}_{t \in Z}$ un processus autorégressif stationnaire d'ordre p satisfaisant l'équation (2.2) où $\ell = 1$. Alors, pour une taille échantillonnale n assez grande, la statistique $\hat{F}_{(p,d)}$, définie par l'équation (3.6), suit approximativement une loi de Fisher avec $p+1$ et $n-d-b-p-h$ degrés de liberté. De plus, lorsque n tend

vers l'infini, $(p+1)\hat{F}_{(p,d)}$ est asymptotiquement distribuée selon une loi du *khi-deux* avec $p+1$ degrés de liberté.

Ce théorème peut être démontré en utilisant les techniques de Tsay (1986, théorème 3.1) et le lemme 3.1 de Keenan (1985). Basé sur le théorème précédent, la statistique $\hat{F}_{(p,d)}$ permet donc de vérifier l'hypothèse nulle suivante:

$$H_0: \text{Il n'y a qu'un seul régime.}$$

Selon cette hypothèse, si nous considérons le cas d'un seul régime ($\ell = 1$), nous nous donnons alors le processus stationnaire $GENAR(p)$ suivant:

$$X_t = \alpha_1 \circ X_{t-1} + \alpha_2 \circ X_{t-2} + \dots + \alpha_p \circ X_{t-p} + \varepsilon_t,$$

où $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ est une suite de variables entières non-négatives de moyenne finie μ_ε et de variance finie σ_ε^2 .

Dans l'analyse de régression linéaire, sous H_0 , la statistique du test est distribuée selon une loi de Fisher telle que $\hat{F}_{(p,d)} \sim F_{v_1, v_2}$ avec $v_1 = (p+1)$ et $v_2 = (n - d - b - p - h)$ degrés de libertés. Ce résultat approximatif nous permet donc l'utilisation de la table de Fisher afin de vérifier l'hypothèse H_0 . Elle consiste à rejeter l'hypothèse nulle au niveau α si la valeur observée de la statistique $\hat{F}_{(p,d)}$ excède $F_{v_1, v_2; 1-\alpha}$, où $F_{v_1, v_2; 1-\alpha}$ dénote le quantile d'ordre $(1-\alpha)$ de la distribution F_{v_1, v_2} .

Notons que, si nous remplaçons les résidus standardisés de la prévision $\hat{\varepsilon}_t$ de l'équation (3.4) par les résidus ordinaires \hat{a}_t de l'équation (3.3), la distribution

asymptotique de la statistique $\hat{F}_{(p,d)}$ demeure la même. Selon Tsay (1989), il est préférable de recourir aux résidus standardisés dans le cas d'un échantillon de taille petite. Dans le cas contraire, les résidus ordinaires suffisent.

Afin d'évaluer la puissance du test, Tsay (1989) l'a comparé au test Portmanteau, proposé par Davies et Petruccelli (1986). Nous pouvons justifier l'intérêt de cette comparaison par la popularité du test Portmanteau basé sur le fait que, sous l'hypothèse qu'un modèle est correctement ajusté à une série donnée, les autocorrélations des résidus de prévision d'horizon 1 sont approximativement indépendantes entre elles et suivent une loi normale.

3.4 Identification de la variable seuil

Selon Tsay (1989), la difficulté majeure dans l'étude des modèles à seuils réside dans l'identification de la valeur du seuil r_1 jouant un rôle fondamental puisqu'elle est déterminante dans l'identification de la nature non-linéaire du modèle *AREAX*.

Dans l'objectif de déterminer la valeur du seuil, nous avons tout d'abord besoin d'identifier la variable seuil X_{t-d} . Cette identification équivaut à la détermination du paramètre de retard d . Pour ce faire, nous allons exposer dans la sous-section suivante la procédure utilisée par Tsay (1989), que nous avons suivie dans ce mémoire afin de déterminer la linéarité par morceaux du modèle *AREAX*. Dans le but de résoudre cette étape de la modélisation des modèles à seuils, quelques autres méthodologies ont été expérimentées. Celles proposées par Tong et Lim (1980) en sont des exemples.

3.4.1 Sélection du paramètre de retard d

Une première approche possible permettant de sélectionner le paramètre de retard d (après avoir choisi tous les autres paramètres) est suggérée par Tong et Lim (1980). Ils ont utilisé le critère d'information d'Akaike noté CIA , (Akaike (1974)). Rappelons tout d'abord rapidement le principe. Étant données des observations (X_1, \dots, X_n) et un modèle ajusté à ces données, soit $\hat{\varepsilon}_t$ les résidus obtenus après ajustement du modèle. Si l'on note SCR la somme des carrés des résidus et k le nombre de paramètres du modèle, alors, le critère d'Akaike est défini par:

$$CIA = n \log\left(\frac{SCR}{n}\right) + 2k.$$

Soient le processus autorégressif à valeurs entières auto-excitant $AREAX$ d'ordre p donné et la suite S telle que $S = \{\text{entiers positifs prédéterminés, candidats pour le retard } d\}$. Nous choisissons un estimateur du paramètre du retard d en nous basant sur la valeur maximale de la statistique de l'équation (3.6); en d'autres termes, on détermine l'élément $v \in S$, noté \hat{d} , tel que

$$\hat{F}_{(p, \hat{d})} = \max_{v \in S} \{\hat{F}_{(p, v)}\}. \quad (3.7)$$

N.B: si \hat{d} n'est pas unique, on prend la plus petite valeur rencontrée. L'estimateur \hat{d} peut dépendre de p puisque la statistique $\hat{F}_{(p, v)}$ en dépend.

Selon Tsay, puisque les degrés de liberté sont différents, nous devons calculer les niveaux critiques des $\hat{F}_{(p, v)}$, $v \in S$. Nous retenons la valeur \hat{d} correspondante à la valeur minimale des niveaux critiques associées aux statistiques $\hat{F}_{(p, v)}$, $v \in S$. Cela

est une solution de rechange à la recherche du maximum des statistiques $\hat{F}_{(p, \nu)}$, $\nu \in S$, de la formule (3.7).

3.4.2 Localisation des valeurs de seuils

Nous savons que le fait d'avoir k seuils implique l'existence de $k + 1$ régimes et, par conséquent, de $k + 1$ modèles linéaires dont au moins deux différents l'un de l'autre. Pour simplifier, considérons le cas de deux régimes et un seul seuil (sans perte de généralité). Soit le modèle *AREAX* de l'équation (3.2). Supposons que la vraie valeur du seuil r_1 satisfait la condition, $X_{\pi_s} \leq r_1 \leq X_{\pi_{s+1}}$. Donc, toutes les valeurs entières de l'intervalle $[X_{\pi_s}, X_{\pi_{s+1}}]$ représentent aussi de bons candidats estimant le seuil r_1 car chacun de ces entiers conduit au même résultat d'ajustement du modèle à seuils spécifié. Supposons que la valeur du seuil que nous recherchons soit égale à r_1 . Ainsi, si la variable seuil $X_{t-\hat{d}}$ est inférieure à r_1 , nous nous retrouvons dans le premier régime. Par contre, si $X_{t-\hat{d}}$ est supérieure ou égale à r_1 , nous nous retrouvons dans le deuxième régime. La variable seuil $X_{t-\hat{d}}$ est telle que déjà vue à la sous-section 3.2.1 de ce chapitre. Elle peut prendre ses valeurs dans l'ensemble des entiers positifs $H = \{X_h, \dots, X_{n-\hat{d}}\}$ exprimés en ordre croissant, c'est-à-dire $X_{\pi_1-\hat{d}}, \dots, X_{\pi_{n-\hat{d}-h+1}-\hat{d}}$, où π_i dénote l'indice temporel de la $i^{\text{ème}}$ plus petite observation de l'ensemble H , où $h = \max\{1, p + 1 - \hat{d}\}$, comme déjà vu à la sous-section 3.2.1 de ce chapitre.

Dans le but de localiser la valeur de seuil la plus appropriée parmi les valeurs de l'intervalle $[X_{\pi_s}, X_{\pi_{s+1}}]$, on utilise les graphiques des rapports standardisés (Tratios) de l'estimation récursive des coefficients d'un *GENAR* en fonction de la variable seuil $X_{t-\hat{d}}$, où le paramètre \hat{d} retenu ici est celui déterminé selon le critère

donné à l'équation (3.7) de la sous-section 3.4.1. L'utilité de tels rapports est double: d'abord, ils évaluent la signification des coefficients d'un *GENAR* particulier et ensuite, quand les coefficients sont significatifs, ceux-ci changent graduellement, permettant ainsi des bifurcations. Par conséquent, nous pouvons localiser directement la valeur de seuil déterminant le point de rupture entre les régimes. Ainsi, le seuil constitue la frontière des différents modèles linéaires de chacun des régimes.

Les rapports standardisés sont définis par:

$$t_i^{[m]} = \frac{\hat{\alpha}_i^{[m]}}{\sqrt{\hat{\Sigma}_i^{[m]}}}, \quad m = b + 1, \dots, n - \hat{d} - h + 1, \quad i = 1, \dots, p, \quad (3.8)$$

où

$$\hat{\Sigma}_i^{[m]} = \sqrt{\hat{Var}(\hat{\alpha}_i^{[m]})} = (M_m^T M_m)^{-1/2} \sigma_\varepsilon.$$

3.5 Modélisation des *AREAX*

Dans cette section, nous utilisons les résultats des sections précédentes afin de proposer une procédure de modélisation des *AREAX*. La procédure est constituée de plusieurs étapes que nous exposons ci-dessous.

Première étape. Choisir l'ordre p du processus *GENAR* ainsi que l'ensemble S des entiers positifs représentant les valeurs possibles du retard \hat{d} .

Dans cette étape, l'ordre p du processus *GENAR* devrait être choisi en se basant sur la fonction des autocorrélations partielles, dénotée par *FACP*, ou bien par un critère d'information tel que celui d'Akaike. (Sauf qu'il est préférable selon Tsay de considérer la *FACP*, puisqu'elle nous fournit des valeurs raisonnables de p).

Deuxième étape. Ajuster les autorégressions ordonnées pour une valeur de l'ordre p et du paramètre d donnée. Calculer et évaluer la statistique $\hat{F}_{(p,d)}$. Advenant le rejet de l'hypothèse de linéarité du processus, choisir le paramètre de délai \hat{d} le plus approprié selon la méthode décrite à la sous-section 3.4.1.

Troisième étape. Une fois p et \hat{d} sont déterminés, localiser les valeurs des seuils X_{t-d} en utilisant les graphiques des rapports standardisés introduits à la sous-section 3.4.2. Notons que, tant que les différents coefficients du processus *GENAR* sont significatifs, les graphiques des rapports standardisés de ces coefficients peuvent être examinés. Par contre, si les coefficients sont non significatifs, ces graphiques ne donnent pas d'information.

Quatrième étape. Si nécessaire, raffiner dans chaque régime l'ordre p du processus *GENAR* ainsi que les valeurs des seuils. Puisque l'ordre p assez grand d'un modèle *AR* peut fournir une approximation raisonnable d'un modèle non-linéaire, la sélection du paramètre p à l'étape 1 peut donc refléter une telle approximation. Selon Tsay, la *FACP* est en mesure de mieux nous révéler cette propriété que ne le ferait le *CIA* utilisée par Tong et Lim (1980). La raison est simple: la *FACP* n'impose pas de pénalités sur les plus grands termes de l'ordre p contrairement au *CIA* qui a tendance à pénaliser les termes d'ordre supérieur.

Conclusion

Jusqu'à présent, sous H_0 , nous avons exposé notre approche concernant le test de la linéarité locale des *AREAX*, ainsi que le comportement asymptotique de la statistique du test relativement à cette hypothèse nulle. À la lumière de l'approche décrite dans ce chapitre, nous allons maintenant procéder à une étude expérimentale vérifiant la crédibilité et la validité de certaines théories par simulation. Les résultats étant asymptotiques, il importe de s'assurer qu'ils peuvent être utilisés dans le contexte d'une série relativement courte car, en pratique, on a souvent des séries de longueur relativement limitée. Nous élaborerons sur les détails importants des simulations, nous identifierons aussi les résultats approximatifs que nous analyserons pour répondre aux objectifs de l'expérience.

CHAPITRE 4

SIMULATIONS

4.1 Méthodologie des simulations

Dans le but de tirer des conclusions éclairées vérifiant l'application des résultats asymptotiques dans le cas des séries de tailles finies, nous avons procédé par simulations des séries.

Notons que pour toutes les simulations effectuées, nous retrouvons à l'annexe I les programmes ayant servi à la réalisation des simulations et à l'annexe

II les tableaux des résultats et les représentations graphiques relatives à la distribution asymptotique de la statistique du test.

4.1.1 Structure des programmes et simulations

L'algorithme décrit dans ce qui suit s'appuie sur la méthode de Tsay (1989). Il a été utilisé dans le programme informatique de l'annexe I afin de calculer la statistique du test considérée. Une fois cette statistique calculée, elle est reportée dans un fichier de données utilisé par ce programme afin d'étudier sa distribution. Dans le programme informatique, nous considérons deux points critiques et nous calculons la probabilité de rejet de l'hypothèse nulle décrite à la section (3.1.3) du chapitre 3. L'objectif de notre programmation est de déterminer si la valeur observée de la statistique tabulée avec $p+1$ et $n-d-b-p-h$ degrés de liberté conduit à un taux de rejet raisonnable. La procédure «Univariate» de SAS utilisée dans le programme fournit les principales caractéristiques marginales d'un ensemble de données, à savoir: la moyenne, l'écart type, les quantiles et elle trace les histogrammes.

Avant de présenter nos résultats, voici tout d'abord quelques détails techniques concernant la réalisation des simulations. Dans la sous-section suivante, nous exposons la méthode ayant permis de produire les séries.

4.1.2 Production des séries

Les séries ont été produites à l'aide de programmes écrits en utilisant le module SAS/IML du progiciel SAS. Ces programmes sont construits de telle sorte qu'avant chaque soumission, l'utilisateur doit spécifier les valeurs des paramètres du

modèle. Le traitement des sorties tirées des programmes a été fait avec le chiffrier électronique EXCEL sur Macintosh.

La production d'une série de longueur n selon un processus autorégressif d'ordre p exige p valeurs initiales. Pour chacune des simulations, nous avons donc initialisé le processus en générant p observations indépendantes, notées $X_0, X_{-1}, \dots, X_{1-p}$ (dans notre cas $p=2$), toutes de loi de Poisson. Cette façon d'initialiser le processus donne lieu à une série asymptotiquement stationnaire. C'est pourquoi, afin d'obtenir une série simulée $X_t = (X_1, \dots, X_n)$ de longueur n , nous avons dû simuler $n+n'$ observations du modèle théorique, où n' est un nombre fixe de données supplémentaires en vue d'atténuer l'effet des valeurs initiales. Nous avons pris $n=200$ et $n'=100$ et avons considéré $N=1000$ réalisations pour produire deux types de modèle: $AREAX(1; 1, 2)$ et $AREAX(2; 1, 2)$. Le germe considéré pour produire ces modèles fut 12345. Pour chaque série simulée, nous avons estimé les paramètres du modèle et nous avons calculé la statistique $\hat{F}_{(2, 1)}$.

4.2 Distribution des modèles simulés

Dans le but de vérifier l'application des résultats relatifs à la distribution asymptotique à des séries de tailles finies, nous nous limitons ici au modèle (2.2) du chapitre 2, où $\ell = 1$. En d'autres termes, nous étudions le cas d'un seul régime. Il est clair que le modèle $AREAX(1; 1, 2)$ associé n'est qu'un $GENAR(2)$. Chaque processus à valeurs entières fait intervenir des termes d'erreur d'une loi de Poisson de moyenne $\lambda = 1$. Les différentes valeurs des paramètres que nous avons considérées pour ce modèle $AREAX(1; 1, 2)$ sont:

$(\alpha_1, \alpha_2) = (0,1; 0,8), (0,2; 0,5), (0,3; 0,4), (0,1; 0,5)$ et $(0,7; 0,1)$. Ces points sont représentés dans l'espace paramétrique illustré à la figure 4.1 par les nombres 1, 2, 3, 4 et 5 respectivement.

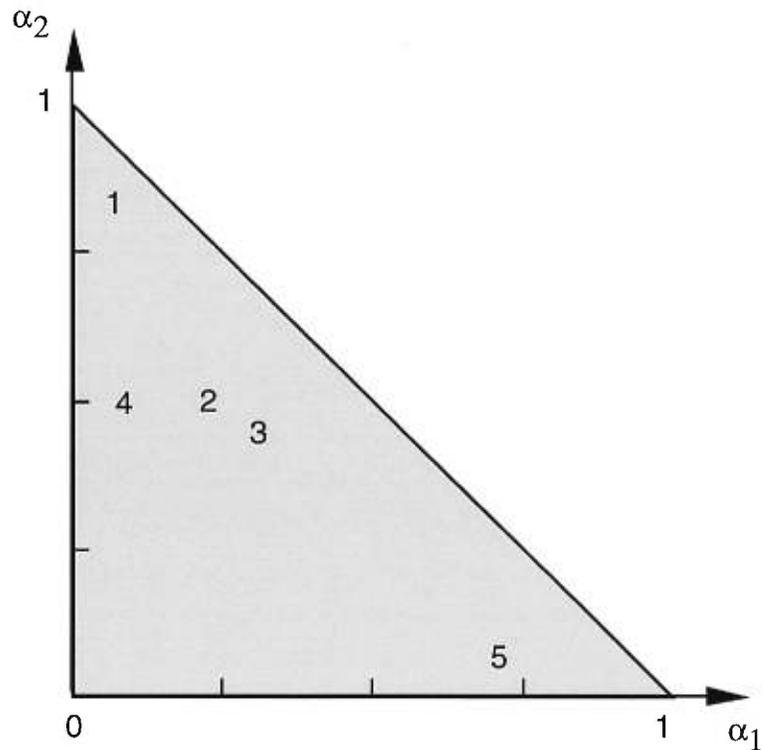


Figure 4.1: Position des paramètres des différents modèles $AREAX(1; 1, 2)$

Les modèles associés sont donc les suivants:

$$X_t = 0,1 \circ X_{t-1} + 0,8 \circ X_{t-2} + \varepsilon_t \quad (4.1)$$

$$X_t = 0,2 \circ X_{t-1} + 0,5 \circ X_{t-2} + \varepsilon_t \quad (4.2)$$

$$X_t = 0,3 \circ X_{t-1} + 0,4 \circ X_{t-2} + \varepsilon_t \quad (4.3)$$

$$X_t = 0,1 \circ X_{t-1} + 0,5 \circ X_{t-2} + \varepsilon_t \quad (4.4)$$

$$X_t = 0,7 \circ X_{t-1} + 0,1 \circ X_{t-2} + \varepsilon_t. \quad (4.5)$$

Dans le but d'évaluer l'impact de la longueur de la série sur la distribution de la statistique $\hat{F}_{(2,1)}$, nous avons effectué notre expérience sur deux tailles échantillonales distinctes, à savoir, $n = 100$ et $n=200$ observations.

Le tableau 4.1, donné à la section 4.3, présente les résultats relatifs au modèle (4.1) tandis que les tableaux II.1 à II.4 de l'annexe II présentent les résultats obtenus relativement aux modèles 4.2 à 4.5. Ces tableaux fournissent le taux de rejet de l'hypothèse nulle selon les longueurs des séries considérées et le niveau adopté.

4.3 Résultats relatifs à la distribution asymptotique

Comparons maintenant la valeur calculée de la statistique du test à la valeur théorique appropriée. Notons que pour une taille échantillonnale $n=100$, nous avons $v_1 = p + 1 = 3$ degrés de liberté au numérateur et $v_2 = n - d - b - p - h = 83$ degrés de liberté au dénominateur car le nombre d'observations initiales est $b=12$, le retard $d=1$ et $h = \max(1, p + 1 - d) = 2$.

Pour $n=200$, nous avons respectivement $v_1=3$ et $v_2=173$ degrés de liberté au numérateur et au dénominateur puisque $b=22$, $d=1$ et $h = \max(1, p + 1 - d) = 2$.

Le test consiste à rejeter l'hypothèse nulle si la valeur observée de la statistique $\hat{F}_{(2,1)}$ excède $F_{v_1, v_2; 1-\alpha}$, où $F_{v_1, v_2; 1-\alpha}$ dénote le quantile d'ordre $(1 - \alpha)$ d'une loi de Fisher à v_1 et v_2 degrés de liberté.

Notons que le taux de rejet \hat{P} est obtenu en divisant le nombre de fois où la statistique calculée est supérieure à la valeur théorique $F_{v_1, v_2; 1-\alpha}$ par $N=1000$; c'est-à-dire, le nombre de séries simulées:

$$\hat{P} = \frac{\text{nombre de rejets}}{1000}.$$

L'intervalle de confiance correspondant au niveau de confiance 95 % est:

$$\hat{P} \pm 1,96 \sqrt{\frac{\hat{P}\hat{Q}}{N}},$$

où $\hat{Q} = 1 - \hat{P}$.

Passons maintenant à l'analyse des résultats. Le tableau 4.1 présente les résultats du modèle (4.1). Pour $n = 200$ observations, la moyenne échantillonnale de la statistique $\hat{F}_{(2,1)}$ est 0,8971. Les valeurs théoriques correspondantes, au niveau significatif de 0,10 est $F_{3, 173; 0,90}=2,1157$ et au niveau significatif de 0,05, elle est $F_{3, 173; 0,95}=2,6568$. Quant aux taux de rejet, nous remarquons qu'au niveau nominal de 10 %, on rejette l'hypothèse dans 7,5 % des cas. Au niveau de signification de 5 %, on ne rejette cette hypothèse que dans 3,7 % des cas.

Pour $n = 100$ observations, le rejet de l'hypothèse nulle survient dans 8,4 % des cas au niveau nominal de 10 % et dans 4,5 % des cas au niveau nominal de 5 %. La statistique $\hat{F}_{(2,1)}$ donne lieu à des tests conservateurs aux niveaux nominaux de 5 % et 10 %. Néanmoins, il semble raisonnable de recourir à la loi de Fisher pour des séries de longueurs finies telles que 100 ou 200.

L'impact provoqué par un changement de taille sur le taux de rejet reste relativement faible. Le niveau du test ne semble donc pas être trop affecté par la

longueur de la série car le taux de rejet de l'hypothèse nulle reste assez proche du taux nominal et se comporte sensiblement de la même manière même lorsqu'on change la taille de l'échantillon. Par ailleurs, on peut remarquer que le taux de rejet dépend fortement des paramètres du modèle. On peut expliquer ces résultats par la nature autorégressive du processus en cause.

$\alpha_1=0,1$ et $\alpha_2=0,8$	<u>statistique</u>	<u>valeur</u>	<u>valeur</u>	<u>taux de rejet</u>	<u>taux de rejet</u>
	<u>moyenne</u>	<u>théorique</u> à 10%	<u>théorique</u> à 5%	à 10 %	à 5 %
$n = 200$	0,8971	2,1157	2,6568	$7,5 \pm 1,632$	$3,7 \pm 1,169$
$n = 100$	0,8348	2,1509	2,7145	$8,4 \pm 1,718$	$4,5 \pm 1,283$

Tableau 4.1: Test de l'hypothèse nulle (modèle (4.1))

Notons que les mêmes remarques s'appliquent relativement aux autres modèles. La conclusion est à l'effet que, sous H_0 , nous pouvons utiliser la loi de Fisher à titre d'approximation pour la distribution de la statistique $\hat{F}_{(2,1)}$. Les figures II.2 et II.3 de l'annexe II donnent un aperçu de l'histogramme et du polygone de fréquence des valeurs observées de $\hat{F}_{(2,1)}$ lorsque $n=200$. Le cas de $n=100$ n'a pas été reproduit car les résultats étaient similaires.

4.4 Simulations d'un processus $AREAX(2; 1, 2)$

Nous avons simulé $N=1000$ séries de $n=200$ observations, chacune d'elles à partir du modèle $AREAX(2; 1, 2)$. Pour les raisons mentionnées précédemment, nous avons dû simuler plus d'observations que nécessaire (300 observations). Les 200 dernières observations forment alors notre échantillon. Pour chacune des simulations, nous avons initialisé le processus en générant relativement à chaque régime, deux observations indépendantes, toutes deux de loi de Poisson. Nous voulons confronter l'hypothèse nulle

$$H_0: \text{il n'y a pas de seuil}$$

et la contre-hypothèse H_1 stipulant l'existence d'un seuil dans le modèle $AREAX(2; 1, 2)$ que nous avons prédéterminé de la manière suivante:

$$X_t = \begin{cases} 0,3 \circ X_{t-1} + 0,2 \circ X_{t-2} + \varepsilon_t^{(1)} & \text{si } X_{t-1} < 3 \\ 0,4 \circ X_{t-1} + 0,1 \circ X_{t-2} + \varepsilon_t^{(2)} & \text{si } X_{t-1} \geq 3. \end{cases} \quad (4.6)$$

Remarquons que chaque régime est un processus autorégressif à valeurs entières généralisé d'ordre deux. Les deux régimes sont déterminés par un seuil $r=3$, choisi a priori; par ailleurs, les termes d'erreur $\varepsilon_t^{(j)}$, $j=1, 2$, sont distribués selon une loi de Poisson de paramètre λ_j , où $\lambda_1 = 2,8$ pour le premier régime et $\lambda_2 = 1$ pour le deuxième régime.

Rappelons tout d'abord la définition de la loi de Poisson ainsi que celle de la loi binomiale qui est particulièrement importante puisque, conditionnellement

au fait que $X = x$, $\alpha \bullet X$ admet une loi binomiale de paramètres (x, α) si l'opérateur $\alpha \bullet$ fait intervenir des variables de Bernoulli.

Définition 4.4.1 (loi de Poisson)

On dit que la variable X est distribuée selon une loi de Poisson de paramètre λ ($\lambda > 0$) si et seulement si sa fonction de masse est de la forme:

$$\Pr(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, \dots$$

Définition 4.4.2 (loi binomiale)

On dit que la variable X est distribuée selon une loi binomiale de paramètres n et θ , (n entier et $0 < \theta < 1$) si et seulement si sa fonction de masse est de la forme:

$$\Pr(X = k) = \binom{n}{k} \theta^k (1 - \theta)^{n-k}, \quad k = 0, 1, \dots, n.$$

4.5 Existence d'un seuil dans le modèle $AREAX(2; 1, 2)$

Pour chacune des 1000 séries simulées, nous avons estimé les paramètres du modèle et nous avons calculé la statistique du test. Comme précédemment, nous comparons la statistique calculée à la valeur théorique associée afin d'évaluer le taux de rejet de l'hypothèse nulle. Ce taux de rejet est calculé pour chacune des deux valeurs critiques mentionnées auparavant. Les résultats relatifs à cette simulation sont présentés au tableau II.5 de l'annexe II.

Pour $n = 200$, nous constatons que le rejet de l'hypothèse nulle survient dans plus de 60 % des cas.

Conclusion

Ce chapitre a été consacré à l'expérimentation de l'approche de Tsay (1989) par simulation. L'objectif a été de vérifier si sa méthodologie dans l'étude des modèles à seuils nous permet d'appliquer des résultats asymptotiques dans le cas de séries de tailles finies.

Dans le chapitre 5, nous allons analyser un jeu de données selon la technique de Tsay (1989), déjà utilisée dans l'analyse de plusieurs données réelles.

CHAPITRE 5

APPLICATION

5.1 Etude expérimentale et problématique

Les différents programmes utilisés pour obtenir les résultats de ce chapitre sont présentés dans l'annexe III et les graphiques correspondants se trouvent à l'annexe IV.

Nous voulions appliquer notre approche théorique à un phénomène réel tel qu'une épidémie ou encore les données relatives aux nouveaux immigrants

investisseurs arrivés au Canada. Mais comme nous n'avons pas pu obtenir un tel jeu de données, nous avons procédé par simulation. Nous avons donc simulé une série de $n=200$ observations puisqu'en pratique, nous devons assez fréquemment traiter des séries de longueur similaire.

Le modèle simulé est:

$$X_t = \begin{cases} 1/2 \circ X_{t-1} + 1/6 \circ X_{t-2} + \varepsilon_t^{(1)} & \text{si } X_{t-1} \leq 4 \\ 1/5 \circ X_{t-1} + 1/2 \circ X_{t-2} + \varepsilon_t^{(2)} & \text{si } X_{t-1} > 4, \end{cases} \quad (5.1)$$

où $\varepsilon_t^{(1)} \sim \text{Poisson}(4/3)$, $\varepsilon_t^{(2)} \sim \text{Poisson}(3/2)$, $p=2$ et $d=1$.

Le problème consiste en premier lieu à vérifier si le modèle de la série étudiée est non-linéaire et à détecter ensuite la valeur de seuil.

NB: le fait de considérer que la valeur de seuil appartient au premier plutôt qu'au deuxième régime ne change pas la théorie déjà vue puisque cette valeur est déterminée par les observations de la série elle-même.

5.2 Identification du modèle $AR(p)$

Afin d'identifier le modèle autorégressif, tout comme la méthodologie de Box et Jenkins (1976), Tsay (1989) a suggéré d'examiner le graphique de la fonction d'autocorrélation partielle. Nous savons que, pour un modèle autorégressif d'ordre p , les coefficients d'autorégression partielle devraient tous être sensiblement égaux à zéro à partir du délai $p+1$ tandis que les autocorrélations devraient décroître de façon exponentielle. À partir du délai deux

(voir le tableau IV.1 de l'annexe IV), nous constatons que les estimations $\hat{\phi}_i$, $i \geq 3$, sont sensiblement égales à zéro. La structure de la fonction d'autocorrélation partielle d'une telle série correspond à celle d'un processus $AR(2)$. Nous pouvons donc conclure qu'un modèle autorégressif d'ordre deux peut être adéquat.

5.3 Sélection du retard \hat{d}

La sélection du paramètre de retard se fait en déterminant l'indice d , noté \hat{d} , donnant lieu à la valeur maximale des statistiques $\hat{F}_{(2, d)}$, $d \in S$, de la formule (3.7). Nous ferons varier l'indice d dans l'ensemble des entiers positifs $S = \{1, 2, 3, 4, 5\}$. Puisque les degrés de liberté varient avec p , il est plus aisé de sélectionner le paramètre de retard associé à la valeur minimale des niveaux critiques (p -value). La valeur de \hat{d} trouvée a un rôle important dans la détermination de la valeur de seuil.

Dans son article, Tsay (1989) a limité la taille de l'ensemble S à $S = \{1, \dots, p\}$ et a effectué l'estimation récursive à partir du nombre initial $b = \left\lceil \frac{n}{10} \right\rceil + p$, comme on l'a déjà vu au chapitre 3. Le tableau 5.1 résume cette approche. D'après ce résultat, le paramètre de retard à retenir est $\hat{d} = 1$ qui correspond au niveau critique le plus faible.

Test	$d=1$	$d=2$
statistique $\hat{F}_{(2, d)}$	1,545986	0,795
niveau critique	0,2160311	0,4532239

Tableau 5.1: Sélection du paramètre \hat{d} selon l'approche de Tsay.

À cette étape de la modélisation, nous avons constaté que dans son algorithme, la suggestion de Tsay (1989) limitant S à l'ensemble $\{1, 2\}$, où la valeur 2 correspond à l'ordre autorégressif, a l'avantage d'être plus simple mais elle exclut la possibilité d'obtenir un niveau critique plus important quand d excède cet ordre autorégressif, c'est-à-dire quand $d > p$. C'est pourquoi nous avons plutôt utilisé $S = \{1, 2, 3, 4, 5\}$ comme ensemble de candidats possibles au retard \hat{d} au lieu de $S = \{1, 2\}$. De plus, comme notre jeu de données est à valeurs entières, il y a possibilité de produire certaines valeurs maintes fois. Un tel phénomène est plus susceptible de se produire dans le cas des séries à valeurs entières que dans celui des séries à valeurs réelles. C'est à cause de cela que nous avons effectué l'estimation récursive en se basant sur des blocs d'unités de régression plutôt relatifs au nombre de fois que les valeurs sont rencontrées dans l'échantillon et exprimées en ordre croissant. Cette technique est explicitée dans le paragraphe suivant.

À l'examen des données, il s'avère que l'ensemble des valeurs distinctes prises par notre échantillon de taille 200 se limite à la valeur 0 et aux dix premiers entiers 1, ..., 10. Regroupons dans le bloc B_g le sous-ensemble des valeurs identiques g rencontrées dans l'échantillon et désignons son cardinal par k_g , $g=0, 1, \dots, 10$. L'énumération en ordre croissant des observations de notre échantillon est alors donnée par B_0, B_1, \dots, B_{10} et il va de soi que $\sum_{g=0}^{10} k_g = 200$. On peut remarquer par exemple que la valeur 0 ne s'est produite qu'une seule fois, c'est-à-dire, $k_0=1$. Parce que cela s'avérera plus utile pour nous afin d'avoir un nombre initial d'observations pour effectuer l'estimation des moindres carrés, nous avons regroupé les blocs B_0, B_1 et B_2 en un seul bloc que nous pouvons désigner par B_2^* . Son cardinal est bien sûr donné par $k_0 + k_1 + k_2$. Quant aux autres blocs, ils demeurent inchangés. L'ensemble des blocs n'est à vrai dire que l'ensemble des valeurs susceptibles d'être prises par la variable seuil $X_{\pi_{t-j}}$ (voir la section 3.4.2 du chapitre 3).

C'est au moyen de ces blocs que nous avons effectué l'estimation des paramètres. Le caractère récursif de cette estimation est basée sur les mêmes formules conçues à ces fins, comme on l'a déjà vu à la sous-section 3.2.2 du chapitre 3, sauf que le nombre initial b ici est $b = \sum_{h=0}^2 k_h$ au lieu de $b = \left\lceil \frac{n}{10} \right\rceil + p$, tel que Tsay (1989) l'avait suggéré dans le cas des valeurs réelles. De plus, si l'on pose $K_g = \sum_{h=0}^g k_h$, $2 \leq g \leq 10$, la procédure récursive ne sera pas faite par pas d'une unité, c'est-à-dire $m = K_2, K_2 + 1, \dots, 200$, mais plutôt par pas correspondant au cumul des cardinaux des blocs, c'est-à-dire $m = K_2, K_3, \dots, K_{10}$, où $K_{10} = 200$.

En procédant ainsi, on s'assure que la procédure récursive est invariante quant à toute permutation des éléments d'un même bloc et ce, pour tout bloc.

Considérons l'autorégression ordonnée décrite par le système d'équations (3.2) du chapitre 3. Nous désignons par $\hat{\alpha}^{[m]} = (\hat{\alpha}_0^{[m]}, \hat{\alpha}_1^{[m]}, \hat{\alpha}_2^{[m]})$, $m = K_2, \dots, K_{10}$, les vecteurs des estimations des moindres carrés basés sur les m premiers cumuls de régression et à l'aide des équations de récurrence de la sous-section 3.2.2 du chapitre 3, nous pouvons calculer $\hat{\alpha}^{[K_{s+1}]}$ en fonction de $\hat{\alpha}^{[K_s]}$, où $g = 2, \dots, 9$. Basée sur les considérations prémentionnées, une façon de tester la non-linéarité du processus produit est d'effectuer la régression des résidus de la prévision de l'autorégression ordonnée, à savoir $\hat{e}_{\pi_i+d}^{[b]}$, où $b = K_2 = 30$, par rapport aux variables explicatives $X_{\pi_i+d-\nu}$, $i = 31, \dots, 200 - d - h + 1$, où $h = \max\{1, 2 + 1 - d\}$ et $\nu = 1, 2$:

$$\hat{e}_{\pi_i+d}^{[30]} = \omega_0 + \sum_{\nu=1}^2 \omega_\nu X_{\pi_i+d-\nu} + \varepsilon_{\pi_i+d}^{[30]}.$$

Rappelons que la statistique permettant d'effectuer le test de $H_0 : \omega_1 = \omega_2 = 0$ est donnée par

$$\begin{aligned} \hat{F}_{(2,d)} &= \frac{\left(\left(\hat{e}_{\sim}^{[b]} \right)^\tau \left(\hat{e}_{\sim}^{[b]} \right) - \left(\hat{\varepsilon}_{\sim}^{[b]} \right)^\tau \left(\hat{\varepsilon}_{\sim}^{[b]} \right) \right) / [p+1]}{\left(\hat{\varepsilon}_{\sim}^{[b]} \right)^\tau \left(\hat{\varepsilon}_{\sim}^{[b]} \right) / [n-d-b-p-h]} \\ &= \frac{\left(\left(\hat{e}_{\sim}^{[b]} \right)^\tau \left(\hat{e}_{\sim}^{[b]} \right) - \left(\hat{\varepsilon}_{\sim}^{[b]} \right)^\tau \left(\hat{\varepsilon}_{\sim}^{[b]} \right) \right) / 3}{\left(\hat{\varepsilon}_{\sim}^{[b]} \right)^\tau \left(\hat{\varepsilon}_{\sim}^{[b]} \right) / [168-d-h]} \end{aligned}$$

car $n=200$, $p=2$ et $b=30$.

Dans ce qui suit nous donnons le résultat obtenu.

La valeur calculée de $\max_{d \in S} \hat{F}_{(2, d)}$ est 4,19596, où $S = \{1, 2, \dots, 5\}$. Elle correspond au niveau critique le plus faible, soit 0,006773, lorsque $d=1$. Le paramètre de retard qui doit être retenu est alors $\hat{d}=1$, comme le montre le tableau 5.2 ci-dessous.

Test	$d=1$	$d=2$	$d=3$	$d=4$	$d=5$
statistique $\hat{F}_{(2, d)}$	4,195963	1,173648	0,479236	0,832836	1,321557
niveau critique	0,006773	0,32134	0,697148	0,477537	0,269075

Tableau 5.2: Sélection du paramètre \hat{d} .

5.4 Test de non-linéarité

La première étape pour réaliser le test de non-linéarité est la mise en forme de la régression ordonnée et la seconde consiste à procéder à l'estimation. Nous avons déjà vu au chapitre 3 que la variable seuil $X_{\pi_{r-\hat{d}}}$ peut prendre en ordre croissant les valeurs $\{X_{\pi_h}, \dots, X_{\pi_{n-1}}\} = \{X_{\pi_2}, \dots, X_{\pi_{199}}\}$, où $h = \max\{1, 3 - \hat{d}\} = 2$ et \hat{d} est le paramètre de retard retenu. Nous effectuons la régression des résidus des

prévisions d'horizon 1, $\hat{e}_{\pi_i+1}^{[b]}$, où $b=30$, par rapport aux variables explicatives

$X_{\pi_i+1-\nu}$:

$$\hat{e}_{\pi_i+1}^{[30]} = \omega_0 + \sum_{\nu=1}^2 \omega_\nu X_{\pi_i+1-\nu} + \varepsilon_{\pi_i+1}^{[30]}, \quad i = 31, \dots, 198; \nu = 1, 2,$$

où π_i est l'indice temporel de la $i^{\text{ème}}$ plus petite observation de l'échantillon. Les résidus des prévisions obtenus sont nécessaires –comme on l'a déjà vu– pour le calcul de la statistique du test

$$\hat{F}_{(2,1)} = \frac{\left(\left(\hat{\underline{e}}^{[30]} \right)^\tau \left(\hat{\underline{e}}^{[30]} \right) - \left(\hat{\underline{\varepsilon}}^{[30]} \right)^\tau \left(\hat{\underline{\varepsilon}}^{[30]} \right) \right) / 3}{\left(\hat{\underline{\varepsilon}}^{[30]} \right)^\tau \left(\hat{\underline{\varepsilon}}^{[30]} \right) / 165}.$$

Nous testons la linéarité du modèle de régression précédent visant à vérifier si on doit rejeter l'hypothèse nulle $H_0 : \omega_1 = \omega_2 = 0$ en faveur de la contre-hypothèse:

$$H_1 : \exists \text{ au moins un } i \in \{1, 2\} \text{ tel que } \omega_i \neq 0 .$$

À l'aide des résultats obtenus (voir tableau 5.3), nous pouvons facilement conclure que le modèle simulé est non-linéaire puisque, sous l'hypothèse nulle, on rejette H_0 à un niveau critique de 0,00677. Nous résumons le résultat associé à ce test dans le tableau 5.3 ci-dessous.

Test de non-linéarité	Paramètre de retard $\hat{d}=1$
statistique $\hat{F}_{(2,1)}$	4,195963
niveau critique	0,006773

Tableau 5.3: Test de la non-linéarité.

5.5 Détermination de la valeur du seuil

Tsay (1989) a mentionné que la phase la plus difficile dans l'étude des modèles à seuils consiste à trouver la valeur du seuil la plus adéquate. Cette étape constitue en même temps l'idée fondamentale de sa technique.

Étant donné le retard sélectionné à la section précédente, $\hat{d}=1$, nous retenons la variable seuil d'indice temporel $t-\hat{d}$: $X_{\pi_t-\hat{d}} = X_{\pi_t-1}$. Le principe consiste à localiser la valeur de seuil qui constitue une frontière entre deux groupes d'observations formant ainsi les deux modèles autorégressifs à l'intérieur du modèle étudié d'équation (5.1). Pour ce faire, la technique de Tsay (1989) repose sur l'examen des graphiques décrits à la sous-section 3.3.2 du troisième chapitre. Ces graphiques ont trait aux rapports standardisés et sont définis par

$$t_i^{[m]} = \frac{\hat{\alpha}_i^{[m]}}{\sqrt{\hat{\Sigma}_i^{[m]}}, m = K_3, \dots, K_{10}, i = 0, 1, 2,$$

où

$$\hat{\Sigma}_i^{[m]} = \sqrt{\hat{Var}(\hat{\alpha}_i^{[m]})} = (M_m^T M_m)^{-1/2} \sigma_\varepsilon.$$

Les tableaux 5.4, 5.5 et 5.6 qui suivent fournissent les chiffres devant servir à tracer les graphes donnés à l'annexe IV.

g	K_g	$\hat{\alpha}_0^{[K_g]}$	rapports standardisés	régimes
3	61	2,35677	2,3468	I
4	98	1,36980	1,1473	I
5	134	1,08978	0,9485	II
6	160	1,43859	1,2991	II
7	183	1,34733	1,2322	II
8	195	1,61704	1,4477	II
9	197	1,75145	1,6781	II
10	200	1,72068	1,6768	II

Tableau 5.4: Estimations $\hat{\alpha}_0^{[K_g]}$ et des rapports standardisés correspondants en fonction des valeurs g susceptibles d'être prises par la variable seuil

g	K_g	$\hat{\alpha}_1^{[K_g]}$	rapports standardisés	régimes
3	61	0,02771	0,0692	I
4	98	0,42209	1,67291	I
5	134	0,64977	4,7166	II
6	160	0,36682	4,3647	II
7	183	0,41311	6,6862	II
8	195	0,32198	6,6544	II
9	197	0,28887	7,1614	II
10	200	0,29619	7,7157	II

Tableau 5.5: Estimations $\hat{\alpha}_1^{[K_g]}$ et des rapports standardisés correspondants en fonction des valeurs g susceptibles d'être prises par la variable seuil

g	K_g	$\hat{\alpha}_2^{[K_g]}$	rapports standardisés	régimes
3	61	0,11758	2,6674	I
4	98	0,30030	5,8215	I
5	134	0,25167	5,9836	II
6	160	0,34449	10,20989	II
7	183	0,33403	10,1568	II
8	195	0,33919	9,5660	II
9	197	0,33468	9,5335	II
10	200	0,33550	9,5711	II

Tableau 5.6: Estimations $\hat{\alpha}_2^{[K_g]}$ et des rapports standardisés correspondants en fonction des valeurs g susceptibles d'être prises par la variable seuil

Ces graphiques représentent les rapports standardisés en fonction des valeurs distinctes susceptibles d'être prises par la variable seuil $X_{\pi_{i-1}}$. Ceux donnés par les figures IV.2 et IV.3 de l'annexe IV s'appliquent à nos données: la figure IV.2 représente le graphique des rapports standardisés en fonction de ces valeurs distinctes: nous pouvons remarquer la tendance ou changement de ces rapports et,

par conséquent, là où il y a rupture. Quant à la figure IV.3, elle représente le graphique des rapports standardisés en fonction des positions ou, tout simplement, des indices des observations ordonnées. Ceci nous permet de localiser la valeur associée à cette position où la bifurcation a eu lieu. Quant au graphique IV.4, il nous indique le changement au niveau de la tendance de ces estimations avant d'avoir une certaine tendance stable. Si on avait un seul régime, ces estimations devraient se stabiliser après un certain moment sans changer de direction.

D'après le graphique de la figure IV.3, nous pouvons constater que le point de rupture est situé à la 99^{ème} observation ordonnée de la série ou encore à la 99^{ème} position. Cette position correspond à la valeur 5. C'est également au niveau de cette valeur qu'on constate une rupture dans le graphique de la figure IV.2. Par conséquent, nous retenons la valeur du seuil comme étant la $\pi_{(t-1)}^{ème} = 98^{ème}$ observation ordonnée, soit 4. Ce résultat nous permet ainsi de distinguer les observations de chacun des deux régimes: le premier contient $n_1 = 98$ observations inférieures ou égales à la valeur de seuil 4. Les observations du premier régime forment ainsi la première régression. Par contre, les observations restantes, de taille $n_2 = 102$, ont toutes des valeurs strictement supérieures à la valeur de seuil 4 et forment ainsi la deuxième régression.

Sous la contrainte de la valeur de seuil trouvée, nous sommes maintenant en mesure d'estimer le modèle linéaire associé à chacun des deux régimes. Pour effectuer le test de linéarité de chacun de ces deux régimes, la procédure «proc reg» de SAS nous a permis la réalisation de ce test ainsi que l'estimation des paramètres associés aux premier et deuxième régimes (le tableau 5.7). Nous constatons que le test de linéarité est très significatif puisqu'on rejette l'hypothèse

nulle à un niveau critique de 0,0001 ($P - value = 0,0001$) dans le cas des deux régressions, comme le montre le tableau 5.7.

estimation des para- mètres	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	statistique F	niveau critique
régime I	1,08977	0,64977	0,25167	15,061	0,0001
régime II	2,16540	0,19339	0,35651	10,094	0,0001

Tableau 5.7: Test de linéarité à l'intérieur des deux régimes.

Les programmes conçus aux fins du test d'ajustement du modèle linéaire associé à chacune des deux régressions retenues sont présentés à l'annexe III.

À la lumière des informations obtenues jusqu'à présent, nous sommes en mesure de construire sous la contrainte de la valeur de seuil trouvée précédemment, les modèles de régression correspondants à chacun des deux régimes:

$$X_t = \begin{cases} 1,090 + 0,650X_{t-1} + 0,252X_{t-2} + \epsilon_t^{(1)} & \text{si } X_{t-1} \leq 4 \\ 2,165 + 0,193X_{t-1} + 0,357X_{t-2} + \epsilon_t^{(2)} & \text{si } X_{t-1} > 4. \end{cases}$$

Relativement à chacun des régimes, nous avons jugé la qualité d'ajustement des modèles par l'analyse des résidus car, selon l'approche classique, on sait que, si le modèle choisi est adéquat, les résidus devraient être sensiblement nuls. Ce résultat

est conforme aux sorties de la procédure «proc reg» du logiciel SAS utilisée dans les programmes conçus à ces fins. Nous donnons aux tableaux IV.5 et IV.6 de l'annexe IV les résultats de ces deux analyses des résidus.

CHAPITRE 6

CONCLUSION

Si l'analyse des données est un des objectifs fondamentaux de la statistique, la modélisation est une approche possible. Dans ce mémoire, nous avons explicité les étapes nécessaires, selon l'approche de Tsay (1989), pour modéliser et étudier un modèle à seuil. Étant l'étape la plus importante selon Tsay, la sélection du paramètre de retard nous permet l'identification de la variable seuil. À ce propos, nous avons fait remarquer que la procédure de Tsay peut manquer de performance car, malgré qu'elle a l'avantage d'être plus simple, elle exclut la possibilité d'obtenir de meilleurs résultats en imposant comme borne supérieure de l'ensemble S des valeurs du paramètre de

retard l'ordre autorégressif p . De plus, nous avons fait remarquer précédemment que la procédure récursive de Tsay appliquée par pas d'une unité à un processus à valeurs entières, où l'on est susceptible de rencontrer des ex aequo, est dépendante de l'ordre d'énumération des observations de même valeur. En effet, une permutation de celles-ci donnerait un résultat différent. C'est pourquoi nous avons visé à effectuer cette étape avec plus de performance en proposant une solution de rechange basée sur la notion de blocs d'observations de même valeur: en effectuant la procédure récursive par pas basés sur les cumuls des cardinaux des blocs, nous sommes alors assurés d'obtenir un résultat qui est invariant quant à toute permutation des observations à l'intérieur de tout bloc.

Bien qu'au début de l'étude des modèles à seuils, la distinction entre les concepts de linéarité et non-linéarité ait été difficile à concevoir, nous avons vu que l'approche de Tsay permet de s'orienter vers l'analyse de ces modèles. De plus, dans sa technique, l'utilisation des représentations graphiques constitue un outil important et direct d'identification des valeurs de seuil.

L'approche non-linéaire des séries chronologiques est une avenue riche en développements autant pour les théoriciens que pour les praticiens. Les problèmes que l'on y rencontre sont multiples et ils peuvent trouver dans ce domaine des points de vue complémentaires aux approches jusqu'à présent envisagées. Les champs d'application sont multiples et peuvent inciter les chercheurs à s'orienter dans cette voie de recherche. Par exemple, la construction de tests permettant de vérifier si un modèle à seuil est préférable à un autre modèle non-linéaire reste à faire. Les puissances des différents tests proposés ne sont pas connues en général. La comparaison des tests existants avec certaines approches non paramétriques, par

exemple les tests développés par Diebolt (1990) et Diebolt et Laib (1991) devrait donner un éclairage intéressant quant à l'utilisation de ces modèles.

Ce domaine de recherche n'est évidemment pas fermé. Au contraire, il en est aux premiers stades et mérite d'être poursuivi.

ANNEXE I

Programmes ayant servi aux simulations

**/*Programme ayant servi à simuler un modèle de type
AREAX(1;1,2)=GENAR(2)*;/**

```

LIBNAME ssd '~';
options linesize=95 PS=72;
proc IML;
start;
n=200;
p=2;
d=1;
h=max(1,p+1-d);
m=FLOOR((n/10)+p);
print m;
A=J(n-d-h+1,2);
Yh=J(n-d-h+1,1);
ih=J(n-d-h+1,1);
mat1=J(n-d-h+1,2);
mat2=J(n-d-h+1,2);
Y_pi=J(n-d-h+1,1);
pi=J(n-d-h+1,1);
Aord=J(n-d-h+1,4);

NoSimule=1000;
alpha_1=0.1;
alpha_2=0.8;
lambda_1=1;
F=J(1,1);
Fstat=J(NoSimule,1);
DO k=1 TO NoSimule ;
  X_init=J(n+100,1) ;
  Y=J(n,1);
  X_init(11,1)=RANPOI(0,lambda_1) ;
  X_init(12,1)=RANPOI(0,lambda_1) ;
DO t=3 TO n+100 ;
  IF X_init(t-1,1)=0
  THEN tempo_1=0;
  ELSE
  tempo_1=RANBIN(0,X_init(t-1,1),alpha_1) ;
  IF X_init(t-2,1)=0
  THEN tempo_2=0;
  ELSE

```

```

        tempo_2=RANBIN(0,X_init(lt-2,1l),alpha_2) ;
        X_init(lt,1l)=tempo_1+tempo_2+RANPOI(0,lambda_1) ;
        END ;
Y(l1:n,1l)=X_init(l101:n+100,1l);
simser =Y ;
DO i=1 TO n-d-h+1;
    ih[i]=i+h-1;
    Yh[i]=Y[ih[i]];
    END;
mat1=ih||Yh;
ind=RANK(Yh);
DO i=1 TO n-d-h+1;
    Y_pi[ind[i]]=Yh[i];
    pi[ind[i]]=ih[i];
    END;
mat2=pi||Y_pi;
DO i=1 TO n-d-h+1;
    Aord[i,1]=Y[pi[i]+d];
    Aord[i,2]=1;
    DO j=1 TO p;
        Aord[i,j+2]=Y[pi[i]+d-j];
    END;
    END;
Pm=J(1,1) ;
Pm=GINV(Aord(l1:m,2:p+2l)*Aord(l1:m,2:p+2l));
beta_m=Pm*(Aord(l1:m,2:p+2l))*Aord(l1:m,1l);
resid=J(NROW(Aord)-m,1);
res=J(1,1);
j=1;
DO i=m+1 TO NROW(Aord);
    D_m1=1.0+Aord(li,2:p+2l)*Pm*(Aord(li,2:p+2l))';
    K_m1=(Pm*(Aord(li,2:p+2l)))/D_m1 ;
    P_m1=(I(p+1)-Pm*((Aord(li,2:p+2l))*Aord(li,2:p+2l)/D_m1))*Pm ;
    beta_m1=beta_m+K_m1*(Aord(li,1l)-Aord(li,2:p+2l)*beta_m) ;
    a_m1=Aord(li,1l)-Aord(li,2:p+2l)*beta_m ;
    res=a_m1/SQRT(D_m1);
    resid(lj,l)=res;
    j=j+1;
    Pm=P_m1;
    beta_m=beta_m1 ;
    END;

```

```

Pe=INV((Aord(lm+1:NROW(Aord),2:p+2l))*Aord(lm+1:NROW(Aord),2:p+2l));
omega_m1=J(p+1,1);
omega_m1=Pe*(Aord(lm+1:NROW(Aord),2:p+2l))*resid ;
Sepsilca=J(1,1);
epsilon=J(NROW(Aord)-m,1);
epsilon=resid(11:NROW(resid),1l)-Aord(lm+1:NROW(Aord),2:p+2l)*omega_m1 ;
/* print epsilon */;
Sepsilca=(epsilon(11:NROW(epsilon),1l))*epsilon(11:NROW(epsilon),1l) ;
/* print Sepsilca */;
/* Sepsilca=(resid(11:NROW(resid),1l)-
Aord(11:NROW(Aord),2:p+2l)*omega_m1)*
(resid(11:NROW(resid),1l)-Aord(11:NROW(Aord),2:p+2l)*omega_m1) */;
/* PRINT Sepsilca */;
Sresidca=(resid(11:NROW(resid),1l))*resid(11:NROW(resid),1l);
/* print Sresidca */;
F=((Sresidca-Sepsilca)/(p+1))/(Sepsilca/(n-d-m-p-h));
Fstat(lk,l)=F;
END;
test=Fstat;
alfa1=0.10;
alfa2=0.5;
F1=FINV(1-alfa1,p+1,n-d-m-p-h);
F2=FINV(1-alfa2,p+1,n-d-m-p-h);
compt1=0;
compt2=0;
DO i=1 TO NROW(Fstat);
  IF Fstat[i] > F1 THEN compt1=compt1+1;
  IF Fstat[i] > F2 THEN compt2=compt2+1;
  END;
print F1;
print F2;
title2 'le nombre de fois a rejeter l,hypothese nulle au niveau alfa1=0.10';
PRINT compt1;
taux1=compt1/NoSimule;
PRINT taux1;
title3 'le nombre de fois a rejeter l,hypothese nulle au niveau alfa2=0.5';
PRINT compt2;
taux2=compt2/NoSimule;
PRINT taux2;
FBAR=J(1,1);
FBAR=SUM(Fstat)/NROW(Fstat);

```

```

PRINT FBAR;
sigma=(Fstat[1]-FBAR)**2;
DO i=2 TO NoSimule ;
  sigma=sigma+(Fstat[i]-FBAR)**2;
  END;
sigma=sqrt(sigma/(NoSimule-1));
PRINT sigma;
Nbclasse=20;
Fmin=0;
Fmax=FBAR+(2.5*sigma);
etendue=Fmax-Fmin;
largeur=etendue/Nbclasse;
compteur=J(Nbclasse+1,1);
DO i=1 TO Nbclasse+1;
  compteur[i]=0;
  END;
DO i=1 TO NoSimule ;
  Noclasse=FLOOR((Fstat[i]-Fmin)/etendue*Nbclasse)+1;
  IF Noclasse>Nbclasse THEN Noclasse=Nbclasse+1;
  compteur[Noclasse]=compteur[Noclasse]+1;
  END;
BINF=J(Nbclasse,1);
BSUP=J(Nbclasse,1);
DO i= 1 TO Nbclasse;
  BINF[i]=Fmin+(i-1)*largeur;
  BSUP[i]=Fmin+i*largeur;
  END;
THEO=J(Nbclasse,1);
DO i= 1 TO Nbclasse ;
  BINF[i]=Fmin+(i-1)*largeur ;
  BSUP[i]=Fmin+i*largeur ;
  THEO[i]=NoSimule*(PROBF(BSUP[i],p+1,n-d-m-p-h)-
  PROBF(BINF[i],p+1,d-m-p-h));
  END;
borninf=J(Nbclasse+1,1);
bornsup=J(Nbclasse+1,1);
lastBINF=Fmin+Nbclasse*largeur;
borninf=BINF//lastBINF;
bornsup=BSUP;
lastTHEO=NoSimule*(1-PROBF(lastBINF,p+1,n-d-m-p-h));
theoric=THEO//lastTHEO;

```

```
efftheo=theoric;
PRINT borninf bornsup theoric compteur;
Chi2=(compteur[1]-THEO[1])**2/THEO[1];
DO i=2 TO Nbclasse ;
  Chi2=Chi2+((compteur[i]-THEO[i])**2/THEO[i]);
END;
CREATE ssd.donnee FROM simser (lcolname={Y}l) ;
APPEND FROM simser ;
CREATE ssd.element FROM test (lcolname={Fstat}l) ;
APPEND FROM test ;
CREATE ssd.obstheo FROM efftheo (lcolname={theoric}l) ;
APPEND FROM efftheo;
FINISH ;
RUN;

PROC CHART data=ssd.element;
  HBAR Fstat ;
run ;

PROC TIMEPLOT data=ssd.donnee;
  title 'la serie Y';
  plot Y='*' ;
RUN ;

PROC UNIVARIATE data=ssd.element plot normal;
  title ' La statistique F ' ;
RUN;

PROC UNIVARIATE data=ssd.obstheo plot normal;
  title 'les effectifs theoriques ' ;
RUN;
```

```
/* Programme ayant servi à simuler un modèle de type AREAX(2;1,2)*;
```

```
LIBNAME ssd '~';
options linesize=95 PS=72;
proc IML;
start;
n=200;
p=2;
r=3;
d=1;

NoSimule=1000;
alpha_1=0.3;
alpha_2=0.2;
alpha_3=0.4 ;
alpha_4=0.1 ;
lambda_1=2.8;
lambda_2=1;
pr=0.5;
h=max(1,p+1-d);
m=FLOOR((n/10)+p);

A=J(n-d-h+1,2);
Yh=J(n-d-h+1,1);
ih=J(n-d-h+1,1);
mat1=J(n-d-h+1,2);
mat2=J(n-d-h+1,2);
Fstat=J(NoSimule,1);
Y_pi=J(n-d-h+1,1);
pi=J(n-d-h+1,1);
Aord=J(n-d-h+1,4);
DO k=1 TO NoSimule ;
X_init=J(n+100,1) ;
Y=J(n,1);
IF RANUNI(12345)<=pr THEN DO;
  X_init(l1,1)=RANPOI(0,lambda_1/(1-alpha_1-alpha_2)) ;
  X_init(l2,1)=RANPOI(0,lambda_1/(1-alpha_1-alpha_2)) ;
  regime1=1;
END;
ELSE DO ;
```

```

X_init(11,11)=RANPOI(0,lambda_2/(1-alpha_3-alpha_4));
X_init(12,11)=RANPOI(0,lambda_2/(1-alpha_3-alpha_4));
END;
DO t=3 TO n+100;
  IF X_init(t-d,11)<= r THEN DO;
    IF X_init(t-1,11)=0
      THEN tempo_1=0;
    ELSE tempo_1=RANBIN(0,X_init(t-1,11),alpha_1);
    IF X_init(t-2,11)=0
      THEN tempo_2=0;
    ELSE tempo_2=RANBIN(0,X_init(t-2,11),alpha_2);
    X_init(t,11)=tempo_1+tempo_2+RANPOI(0,lambda_1);
    END;
  ELSE DO;
    IF X_init(t-1,11)=0
      THEN tempo_1=0;
    ELSE tempo_1=RANBIN(0,X_init(t-1,11),alpha_3);
    IF X_init(t-2,11)=0
      THEN tempo_2=0;
    ELSE tempo_2=RANBIN(0,X_init(t-2,11),alpha_4);
    X_init(t,11)=tempo_1+tempo_2+RANPOI(0,lambda_2);
    END;
  END;
END;
END;
Y=X_init(1101:n+100,11);
simser =Y;
DO i=1 TO n-d-h+1;
  ih[i]=i+h-1;
  Yh[i]=Y[ih[i]];
  END;
mat1=ih||Yh;
ind=RANK(Yh);
DO i=1 TO n-d-h+1;
  Y_pi[ind[i]]=Yh[i];
  pi[ind[i]]=ih[i];
  END;
mat2=pi||Y_pi;
DO i=1 TO n-d-h+1;
  Aord[i,1]=Y[pi[i]+d];
  Aord[i,2]=1;

```

```

DO j=1 TO p;
  Aord[i,j+2]=Y[pi[i]+d-j];
END;
END;

Pm=J(1,1);
Pm=GINV(Aord(11:m,2:p+2l)*Aord(11:m,2:p+2l));
beta_m=Pm*(Aord(11:m,2:p+2l))*Aord(11:m,1l);
resid=J(NROW(Aord)-m,1);
res=J(1,1);
j=1;
DO i=m+1 TO NROW(Aord);
  D_m1=1.0+Aord(li,2:p+2l)*Pm*(Aord(li,2:p+2l));
  K_m1=(Pm*(Aord(li,2:p+2l))^2/D_m1);
  P_m1=(1-(p+1)-Pm*(Aord(li,2:p+2l))*Aord(li,2:p+2l)/D_m1)*Pm;
  beta_m1=beta_m+K_m1*(Aord(li,1l)-Aord(li,2:p+2l)*beta_m);
  a_m1=Aord(li,1l)-Aord(li,2:p+2l)*beta_m;
  res=a_m1/SQRT(D_m1);
  resid(lj,l)=res;
  j=j+1;
  Pm=P_m1;
  beta_m=beta_m1;
END;

Pe=INV((Aord(lm+1:NROW(Aord),2:p+2l))*Aord(lm+1:NROW(Aord),2:p+2l));
omega_m1=J(p+1,1);
omega_m1=Pe*(Aord(lm+1:NROW(Aord),2:p+2l))*resid;
Sepsilca=J(1,1);
epsilon=J(NROW(Aord)-m,1);
epsilon=resid(11:NROW(resid),1l)-Aord(lm+1:NROW(Aord),2:p+2l)*omega_m1;
Sepsilca=(epsilon(11:NROW(epsilon),1l))*epsilon(11:NROW(epsilon),1l);
Sresidca=(resid(11:NROW(resid),1l))*resid(11:NROW(resid),1l);
F=((Sresidca-Sepsilca)/(p+1))/(Sepsilca/(n-d-m-p-h));
Fstat(lk,l)=F;
END;

test=Fstat;
alfa=0.10;
q=FINV(1-alfa,p+1,n-d-m-p-h);
compt=0;

```

```

DO i=1 TO NROW(Fstat);
  IF Fstat[i] > q THEN compt=compt+1;
  END;
PRINT compt;
taux=compt/NoSimule;
PRINT taux;
FBAR=SUM(Fstat)/NROW(Fstat) ;
PRINT FBAR;
sigma=(Fstat[1]-FBAR)**2;
DO i=2 TO NoSimule ;
  sigma=sigma+(Fstat[i]-FBAR)**2 ;
  END;
sigma=sqrt(sigma/(NoSimule-1)) ;
PRINT sigma ;

Nbclasse=20 ;
Fmin=0;
Fmax=FBAR+(2.5*sigma) ;
etendue=Fmax-Fmin;
largeur=etendue/Nbclasse ;
compteur=J(Nbclasse+1,1) ;
DO i=1 TO Nbclasse+1 ;
  compteur[i]=0 ;
  END;
DO i=1 TO NoSimule ;
  Noclasse=FLOOR((Fstat[i]-Fmin)/etendue*Nbclasse)+1 ;
  IF Noclasse>Nbclasse THEN Noclasse=Nbclasse+1 ;
  compteur[Noclasse]=compteur[Noclasse]+1 ;
  END;
BINF=J(Nbclasse,1);
BSUP=J(Nbclasse,1);
DO i= 1 TO Nbclasse;
  BINF[i]=Fmin+(i-1)*largeur;
  BSUP[i]=Fmin+i*largeur;
END;

THEO=J(Nbclasse,1);
DO i= 1 TO Nbclasse ;
  BINF[i]=Fmin+(i-1)*largeur ;
  BSUP[i]=Fmin+i*largeur ;

```

```

THEO[i]=NoSimule*(PROBF(BSUP[i],p+1,n-d-m-p-h)-
PROBF(BINF[i],p+1,n-d-m-p-h));
END;
borninf=J(Nbclasse+1,1);
bornsup=J(Nbclasse+1,1);
lastBINF=Fmin+Nbclasse*largeur;
borninf=BINF//lastBINF;
bornsup=BSUP;
lastTHEO=NoSimule*(1-PROBF(lastBINF,p+1,n-d-m-p-h));
theoric=THEO//lastTHEO;
efftheo=theoric;

Chi2=(compteur[1]-THEO[1])**2/THEO[1];
DO i=2 TO Nbclasse ;
  Chi2=Chi2+((compteur[i]-THEO[i])**2/THEO[i]);
END;
print Chi2;
CREATE ssd.donnee FROM simser (lcolname={ Y }l) ;
APPEND FROM simser ;
CREATE ssd.element FROM test (lcolname={Fstat}l) ;
APPEND FROM test ;
CREATE ssd.obstheo FROM efftheo (lcolname={theoric}l) ;
APPEND FROM efftheo;
FINISH ;
RUN;

PROC TIMEPLOT data=ssd.donnee;
  TITLE3 ' laserie Y ' ;
  plot Y='*' ;
  RUN ;

PROC UNIVARIATE data=ssd.element plot normal;
  title4 ' La statistique F ' ;
RUN;

PROC UNIVARIATE data=ssd.obstheo plot normal;
  title 'les effectifs theoriques ' ;
RUN;

```

ANNEXE II

Résultats des simulations

$\alpha_1=0,2$ et $\alpha_2=0,5$	<u>statistique</u>	<u>valeur</u>	<u>valeur</u>	<u>taux de rejet</u>	<u>taux de rejet</u>
	<u>moyenne</u>	<u>théorique</u> à 10%	<u>théorique</u> à 5%	à 10 %	à 5 %
$n = 200$	0,8969	2,1157	2,6568	$7,6 \pm 1,642$	$3,6 \pm 1,155$
$n = 100$	0,9249	2,1509	2,7145	$8,2 \pm 1,699$	$4,9 \pm 1,34$

Tableau II.1: Test de l'hypothèse nulle (modèle (4.2)).

$\alpha_1=0,3$ et $\alpha_2=0,4$	<u>statistique</u>	<u>valeur</u>	<u>valeur</u>	<u>taux de rejet</u>	<u>taux de rejet</u>
	<u>moyenne</u>	<u>théorique</u> à 10%	<u>théorique</u> à 5%	à 10 %	à 5 %
$n = 200$	1,0888	2,1157	2,6568	$8,5 \pm 1,727$	$5,1 \pm 1,362$
$n = 100$	0,9971	2,1509	2,7145	$9,3 \pm 1,799$	$5,2 \pm 1,375$

Tableau II.2: Test de l'hypothèse nulle (modèle (4.3)).

$\alpha_1=0,1$ et $\alpha_2=0,5$	<u>statistique</u> <u>moyenne</u>	<u>valeur</u>	<u>valeur</u>	<u>taux de rejet</u>	<u>taux de rejet</u>
		<u>théorique</u> à 10%	<u>théorique</u> à 5%	à 10 %	à 5 %
$n = 200$	1,553	2,1157	2,6568	$14,9 \pm 2,206$	$8,1 \pm 1,691$
$n = 100$	1,0760	2,1509	2,7145	$11,2 \pm 1,954$	$5,6 \pm 1,424$

Tableau II.3: Test de l'hypothèse nulle (modèle (4.4)).

$\alpha_1=0,7$ et $\alpha_2=0,1$	<u>statistique</u> <u>moyenne</u>	<u>valeur</u>	<u>valeur</u>	<u>taux de rejet</u>	<u>taux de rejet</u>
		<u>théorique</u> à 10%	<u>théorique</u> à 5%	à 10 %	à 5 %
$n = 200$	0,8360	2,1157	2,6568	$5,5 \pm 1,412$	$2,6 \pm 0,986$
$n = 100$	0,7994	2,1509	2,7145	$4,7 \pm 1,310$	$2,0 \pm 0,868$

Tableau II.4: Test de l'hypothèse nulle (modèle (4.5)).

$\alpha_1^{(1)}=0,3$ $\alpha_2^{(1)}=0,2$ et $\alpha_1^{(2)}=0,4$ $\alpha_2^{(2)}=0,1$	<u>statistique</u>	<u>valeur</u>	<u>valeur</u>	<u>taux de rejet</u>	<u>taux de rejet</u>
	<u>moyenne</u>	<u>théorique</u>	<u>théorique</u>	à 10 %	à 5 %
		à 10%	à 5%		
$n = 200$	3,0461	2,1157	2,6568	$60 \pm 3,036$	$92 \pm 1,681$
$n = 100$	2,1761	2,1509	2,7145	$36,7 \pm 2,987$	$79,1 \pm 2,519$

Tableau II.5: Test de l'hypothèse nulle (modèle (4.6)).

Représentation graphique de la série telle que simulée

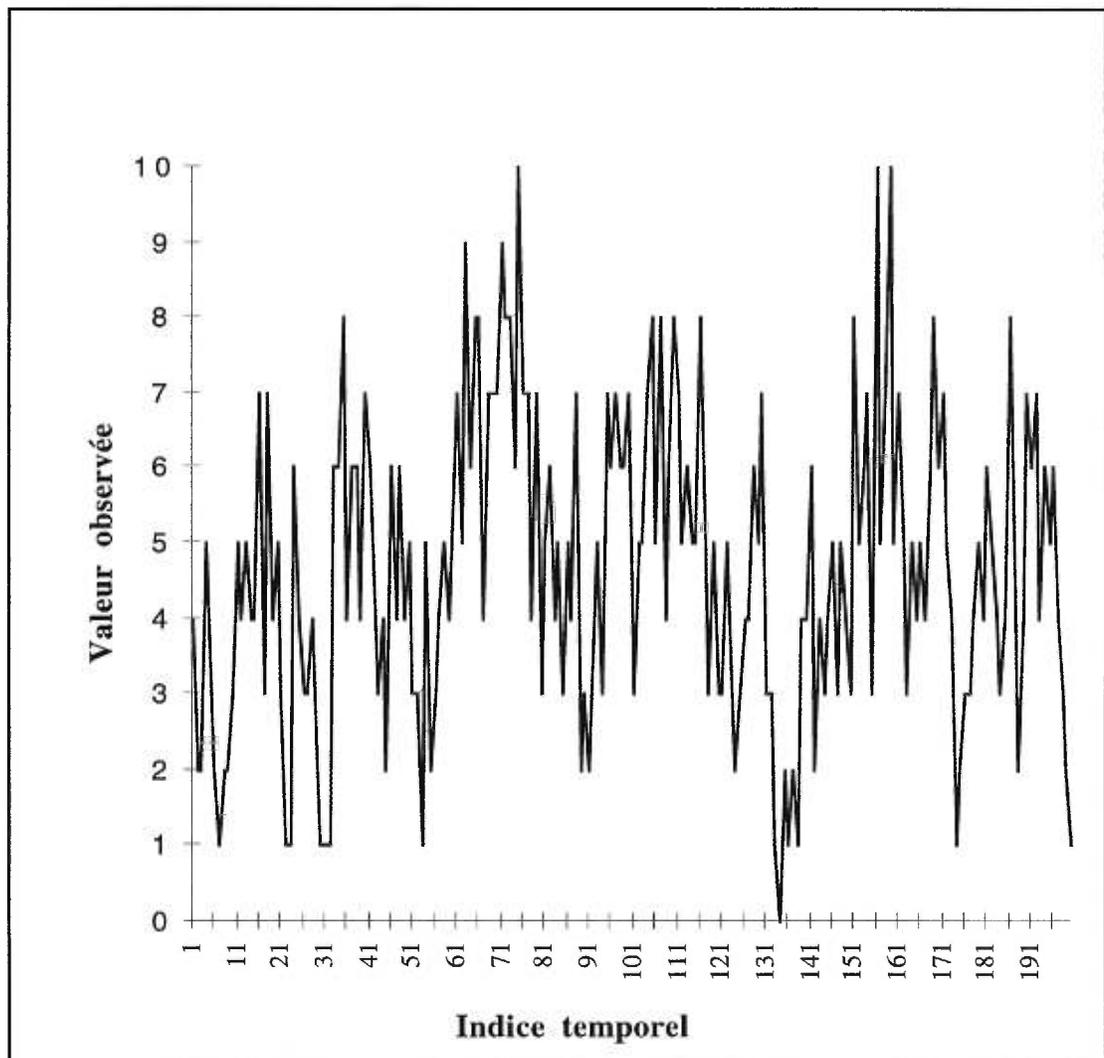


Figure II.1: Graphique de la série simulée

Résultat des simulation: Histogramme des valeurs observées de la statistique

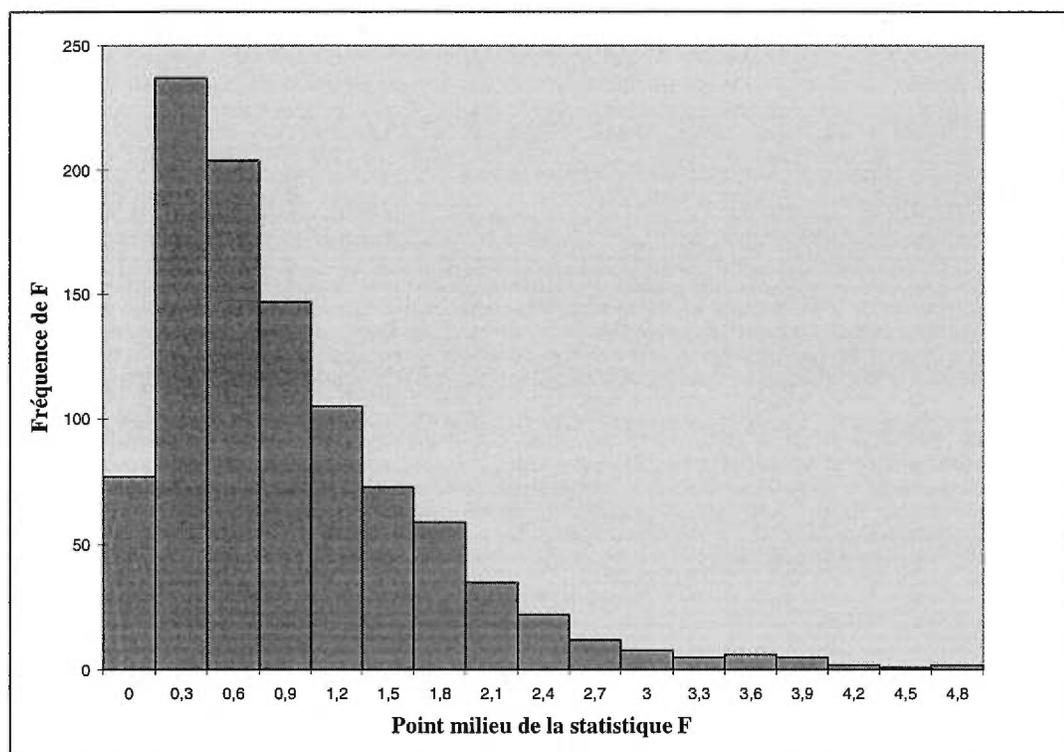


Figure II.2 : Histogramme des valeurs observées de F d'un modèle AREAX (1 ;1,2) lorsque n=200.

Résultat des simulations: Courbe des valeurs observées de la statistique

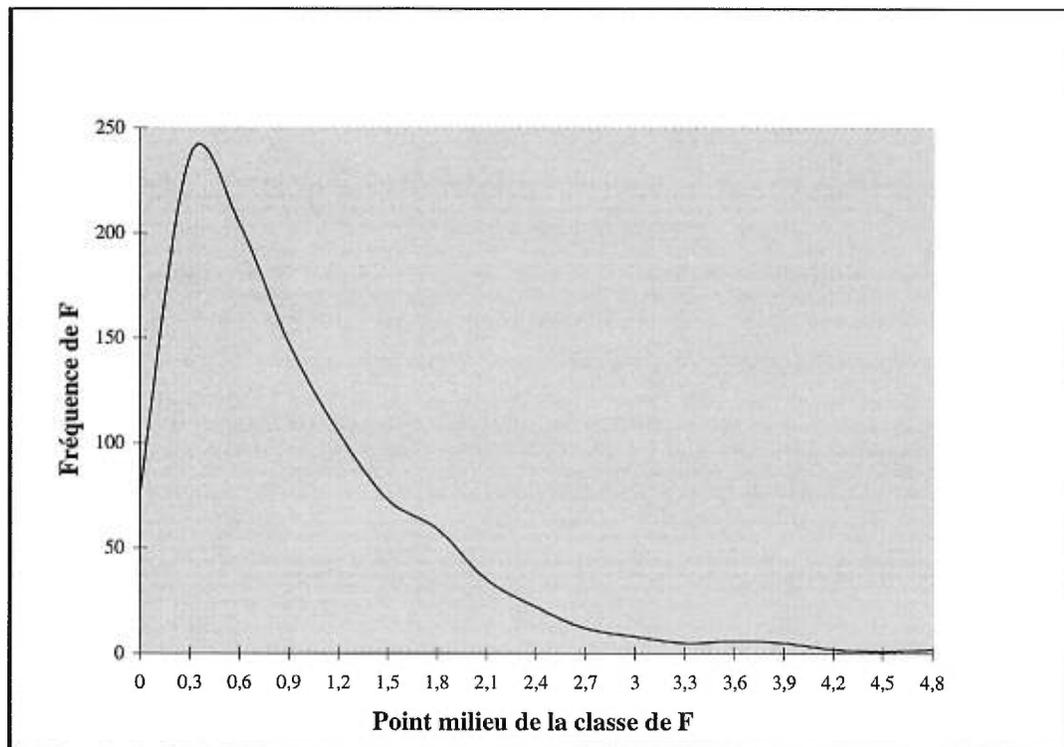


Figure II.3: Polygone de fréquence des valeurs observées de F d'un modèle AREAX(1;1,2) lorsque n=200

ANNEXE III

Programmes ayant servi à l'application

```
/* Programme ayant servi à simuler le jeu de données de l'application */;
```

```
LIBNAME ssd '~';
options linesize=95 PS=72;
proc IML;
start;
n=200;
p=2;
r=4;
d=1;
alpha_1=1/2;
alpha_2=1/6;
alpha_3=1/5 ;
alpha_4=1/2 ;
lambda_1=4/3;
lambda_2=3/2;
pr=0.5;
h=max(1,p+1-d);
Yh=J(n-d-h+1,1);
ih=J(n-d-h+1,1);
Y_pi=J(n-d-h+1,1);
pi=J(n-d-h+1,1);
Aord=J(n-d-h+1,p+2);
Rind=J(n+100,2);
X_init=J(n+100,1) ;
Y=J(n,1);
IF RANUNI(12345)<=pr THEN DO;
  X_init(11,1)=RANPOI(0,lambda_1/(1-alpha_1-alpha_2)) ;
  Rind(11,1)=1;
  Rind(11,2)=X_init(11,1);
  X_init(12,1)=RANPOI(0,lambda_1/(1-alpha_1-alpha_2)) ;
  Rind(12,1)=1;
  Rind(12,2)=X_init(12,1);
END;
ELSE DO ;
  X_init(11,1)=RANPOI(0,lambda_2/(1-alpha_3-alpha_4)) ;
  Rind(11,1)=2;
  X_init(12,1)=RANPOI(0,lambda_2/(1-alpha_3-alpha_4)) ;
  Rind(12,1)=2;
END;
```

```

DO t=3 TO n+100 ;
  IF X_init(lt-d,1l)<= r THEN DO;
    IF X_init(lt-1,1l)=0
      THEN tempo_1=0;
      ELSE tempo_1=RANBIN(0,X_init(lt-1,1l),alpha_1) ;
    IF X_init(lt-2,1l)=0
      THEN tempo_2=0;
      ELSE tempo_2=RANBIN(0,X_init(lt-2,1l),alpha_2) ;
    X_init(lt,1l)=tempo_1+tempo_2+RANPOI(0,lambda_1) ;
    Rind(lt,1l)=1;
    Rind(lt,2l)=X_init(lt,1l);
    END;
  ELSE DO;
    IF X_init(lt-1,1l)=0
      THEN tempo_1=0;
      ELSE tempo_1=RANBIN(0,X_init(lt-1,1l),alpha_3) ;
    IF X_init(lt-2,1l)=0
      THEN tempo_2=0;
      ELSE tempo_2=RANBIN(0,X_init(lt-2,1l),alpha_4) ;
    X_init(lt,1l)=tempo_1+tempo_2+RANPOI(0,lambda_2) ;
    Rind(lt,1l)=2;
    Rind(lt,2l)=X_init(lt,1l);
    END;
  END;
RRet=Rind(l101:n+100,l);
Print RRet;
Y=X_init(l101:n+100,1l);
DO i=1 TO n-d-h+1;
  ih[i]=i+h-1;
  Yh[i]=Y[ih[i]];
  END;
mat1=ih||Yh;
ind=RANK(Yh);
DO i=1 TO n-d-h+1;
  Y_pi[ind[i]]=Yh[i];
  pi[ind[i]]=ih[i];
  END;
mat2=pi||Y_pi;
DO i=1 TO n-d-h+1;
  Aord[i,1]=Y[pi[i]+d];

```

```
Aord[i,2]=1;  
  DO j=1 TO p;  
    Aord[i,j+2]=Y[pi[i]+d-j];  
  END;  
END;  
FINISH;  
RUN ;
```

```
/* Programme ayant servi à l'identification du modèle AR(2) */;
```

```
options linesize=95 PS=72;

data données;
input x @@;
cards;

    «On écrit ici les 200 observations du jeu de donnée»
;
PROC ARIMA data=données;
  identify var=x;
  estimate p =2;
RUN;
```

```
/* Programme ayant servi à la sélection du paramètre de retard */;
```

```
LIBNAME ssd '~';
options linesize=95 PS=72;
proc IML;
start;
n=200;
p=2;
s=5;
m=FLOOR((n/10)+p);
F=J(1,1);
Y=J(n,1);
      Y={ On écrit ici les 200 observations du jeu de donnée à analyser };
pvalue=1;
DO d=1 TO s;
  h=max(1,p+1-d);
  A=J(n-d-h+1,2);
  Yh=J(n-d-h+1,1);
  ih=J(n-d-h+1,1);
  mat1=J(n-d-h+1,2);
  mat2=J(n-d-h+1,2);
  Y_pi=J(n-d-h+1,1);
  pi=J(n-d-h+1,1);
  Aord=J(n-d-h+1,4);
    DO i=1 TO n-d-h+1;
      ih[i]=i+h-1;
      Yh[i]=Y[ih[i]];
    END;
mat1=ih||Yh;
ind=RANK(Yh);
  DO i=1 TO n-d-h+1;
    Y_pi[ind[i]]=Yh[i];
    pi[ind[i]]=ih[i];
  END;
mat2=pi||Y_pi;
  DO i=1 TO n-d-h+1;
    Aord[i,1]=Y[pi[i]+d];
    Aord[i,2]=1;
    DO j=1 TO p;
      Aord[i,j+2]=Y[pi[i]+d-j];
```

```

        END;
    END;
Pm=GINV(Aord(11:m,2:p+2l)*Aord(11:m,2:p+2l));
beta_m=Pm*(Aord(11:m,2:p+2l))*Aord(11:m,1l);
resid=J(NROW(Aord)-m,1);
res=J(1,1);
j=1;
DO i=m+1 TO NROW(Aord);
    D_m1=1.0+Aord(li,2:p+2l)*Pm*(Aord(li,2:p+2l));
    K_m1=(Pm*(Aord(li,2:p+2l)))/D_m1 ;
    P_m1=(I(p+1)-Pm*((Aord(li,2:p+2l))*Aord(li,2:p+2l)/D_m1))*Pm ;
    beta_m1=beta_m+K_m1*(Aord(li,1l)-Aord(li,2:p+2l)*beta_m) ;
    a_m1=Aord(li,1l)-Aord(li,2:p+2l)*beta_m ;
    res=a_m1/SQRT(D_m1);
    resid(lj,l)=res;
    j=j+1;
    Pm=P_m1;
    beta_m=beta_m1 ;
    END;
Pe=GINV((Aord(lm+1:NROW(Aord),2:p+2l))*Aord(lm+1:NROW(Aord),2:p+2l));
omega_m1=J(p+1,1);
omega_m1=Pe*(Aord(lm+1:NROW(Aord),2:p+2l))*resid ;
Sepsilca=J(1,1);
epsilon=J(NROW(Aord)-m,1);
epsilon=resid(11:NROW(resid),1l)-Aord(lm+1:NROW(Aord),2:p+2l)*omega_m1 ;
Sepsilca=(epsilon(11:NROW(epsilon),1l))*epsilon(11:NROW(epsilon),1l) ;
/* Sepsilca=(resid(11:NROW(resid),1l)-
Aord(11:NROW(Aord),2:p+2l)*omega_m1)*
(resid(11:NROW(resid),1l)-Aord(11:NROW(Aord),2:p+2l)*omega_m1) */;
Sresidca=(resid(11:NROW(resid),1l))*resid(11:NROW(resid),1l);
F=((Sresidca-Sepsilca)/(p+1))/(Sepsilca/(n-d-m-p-h));
pr=(1-probF(F,p,n-d-m-p-h));
    IF pr<=pvalue THEN DO;
        pvalue=pr ;
        dp=d;
    END;
END;
print dp ; print F;
FINISH ;
RUN;

```

/* Programme ayant servi au calcul des rapports standardisés */;

```

LIBNAME ssd '~';
options linesize=95 PS=72;
proc IML;
start;

n=200;
p=2;
d=1;
dp=J(1,1);
m=FLOOR((n/10)+p);
print n; print s; print p; /* print d*/; print m;
F=J(1,1);
Y=J(n,1);
Y={On écrit ici les données à analyser};
h=max(1,p+1-dA=J(n-d-h+1,2);
Yh=J(n-d-h+1,1);
ih=J(n-d-h+1,1);
mat1=J(n-d-h+1,2);
mat2=J(n-d-h+1,2);
Y_pi=J(n-d-h+1,1);
pi=J(n-d-h+1,1);
Aord=J(n-d-h+1,p+2);
DO i=1 TO n-d-h+1;
  ih[i]=i+h-1;
  Yh[i]=Y[ih[i]];
  END;
mat1=ih||Yh;
ind=RANK(Yh);
DO i=1 TO n-d-h+1;
  Y_pi[ind[i]]=Yh[i];
  pi[ind[i]]=ih[i];
  END;
mat2=pi||Y_pi;
DO i=1 TO n-d-h+1;
  Aord[i,1]=Y[pi[i]+d];
  Aord[i,2]=1;
  DO j=1 TO p;
    Aord[i,j+2]=Y[pi[i]+d-j];
  
```

```

        END;
    END;
    Pm=GINV(Aord(11:m,2:p+2l)*Aord(11:m,2:p+2l));
    beta_m=Pm*(Aord(11:m,2:p+2l))*Aord(11:m,1l);
    /*print beta_m*/;
    resid_m=J(m,1);
    resid_m[1:m,1]=Aord[1:m,1]-(Aord[1:m,2:p+2l]*beta_m);
    Sresca_m=(resid_m(11:NROW(resid_m),1l))*resid_m(11:NROW(resid_m),1l);
    SSCM=Sresca_m;
    Scarre_m=SSCM/(m-p-1);
    Pe=GINV((Aord(lm+1:NROW(Aord),2:p+2l))*Aord(lm+1:NROW(Aord),2:p+2l));
    resid=J(NROW(Aord)-m,1);
    Scarrem1=J(NROW(Aord)-m,1);
    Ytd=Y_pi[m+1:NROW(Y_pi)];
    T1=J(NROW(Ytd),1);
    T2=J(NROW(Ytd),1);
    T3=J(NROW(Ytd),1);
    j=1;
    DO i=m+1 TO NROW(Aord);
        D_m1=1.0+Aord(li,2:p+2l)*Pm*(Aord(li,2:p+2l));
        K_m1=(Pm*(Aord(li,2:p+2l)))/D_m1;
        P_m1=(I(p+1)-Pm*((Aord(li,2:p+2l))*Aord(li,2:p+2l)/D_m1))*Pm ;
        beta_m1=beta_m+K_m1*(Aord(li,1l)-Aord(li,2:p+2l)*beta_m) ;
        a_m1=Aord(li,1l)-Aord(li,2:p+2l)*beta_m ;
        res=a_m1/SQRT(D_m1);
        resid[j]=res;
        SSCM1=SSCM+res**2;
        Scarrem1(lj,l)=SSCM/(i-p-1);
        SSCM=SSCM1;
        beta_m=beta_m1 ;
        Bm1STD=beta_m1[1]/SQRT(Scarrem1[j]*Pm[1,1]);
        Bm2STD=beta_m1[2]/SQRT(Scarrem1[j]*Pm[2,2]);
        Bm3STD=beta_m1[3]/SQRT(Scarrem1[j]*Pm[3,3]);
        Pm=P_m1;
        T1[j]=Bm1STD;
        T2[j]=Bm2STD;
        T3[j]=Bm3STD;
        j=j+1;
    END;
    Pe=GINV((Aord(lm+1:NROW(Aord),2:p+2l))*Aord(lm+1:NROW(Aord),2:p+2l));

```

```

omega_m1=J(p+1,1);
omega_m1=Pe*(Aord(lm+1:NROW(Aord),2:p+2l))*resid ;
Sepsilca=J(1,1);
epsilon=J(NROW(Aord)-m,1);
epsilon=resid(l1:NROW(resid),1l)-Aord(lm+1:NROW(Aord),2:p+2l)*omega_m1 ;
Sepsilca=(epsilon(l1:NROW(epsilon),1l))*epsilon(l1:NROW(epsilon),1l);
/* Sepsilca=(resid(l1:NROW(resid),1l)-
Aord(l1:NROW(Aord),2:p+2l)*omega_m1)*
(resid(l1:NROW(resid),1l)-Aord(l1:NROW(Aord),2:p+2l)*omega_m1) */;
Sresidca=(resid(l1:NROW(resid),1l))*resid(l1:NROW(resid),1l);
F=((Sresidca-Sepsilca)/(p+1))/(Sepsilca/(n-d-m-p-h));
MATG=J(NROW(Ytd),4);
MATG=T1||T2||T3||Ytd;
T1Ytd=J(NROW(Ytd),2);
T2Ytd=J(NROW(Ytd),2);
T3Ytd=J(NROW(Ytd),2);
T1Ytd=Ytd||T1;
T2Ytd=Ytd||T2;
T3Ytd=Ytd||T3;
call pgraf(T1Ytd,'*', 'Ytd', 'T1', ' Tratios1  $\Psi_{t-d}$ ');
call pgraf(T2Ytd,'*', 'Ytd', 'T2', ' Tratios1  $\Psi_{t-d}$ ');
call pgraf(T3Ytd,'*', 'Ytd', 'T3', ' Tratios1  $\Psi_{t-d}$ ');
FINISH;
RUN;

```

/*Programme d'ajustement de la première régression linéaire */;

```
option pagesize=54 linesize=80 nodate nonumber;
data RÉGIME1;
input temps obs Xt X1 X2 @@;

list;cards;

      Y= « On écrit ici les données associées au premier régime»
;
PROC REG data=REGIME1 ;
      model Xt= X1 X2 /p r vif influence collin ;
      output out =reg1 p=predit r=resid student=rs;
run;

PROC PLOT data=reg1;
      plot rs*(X1 X2)='*/vref=0;
      plot resid*(X1 X2)='+/vref=0;
      plot resid*predit='o'/vref=0;
run;

PROC UNIVARIATE plot normal;
      var Xt;
run;
```

/*Programme d'ajustement de la deuxième régression linéaire */;

```
option pagesize=54 linesize=80 nodate nonumber;
data RÉGIME2;
input temps obs Xt X1 X2 @@;

list;cards;

      Y= « On écrit ici les données associées au deuxième régime»
;
PROC REG data=REGIME2 ;
      model Xt= X1 X2 /p r vif influence collin ;
      output out =reg2 p=predit r=resid student=rs;
run;

PROC PLOT data=reg2;
      plot rs*(X1 X2)='*/vref=0;
      plot resid*(X1 X2)='+/vref=0;
      plot resid*predit='o'/vref=0;
run;

PROC UNIVARIATE plot normal;
      var Xt;
run;
```

ANNEXE IV

Résultats de l'application

Liste des autocorrelations partielles

Lag	Correlation	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1	
1	0.41255											.	*****										
2	0.32805											.	*****										
3	0.03584											.	*										
4	-0.07738											.	**										
5	0.11518											.	**										
6	-0.02690											.	*										
7	-0.01945											.	.										
8	0.00812											.	.										
9	0.04024											.	*										
10	0.03055											.	*										
11	0.03350											.	*										
12	-0.07484											.	*										
13	-0.01041											.	.										
14	-0.08676											.	**										
15	-0.04982											.	*										
16	-0.06215											.	*										
17	0.00733											.	.										
18	-0.12100											.	**										
19	-0.00887											.	.										
20	0.01877											.	.										
21	0.08684											.	**										
22	-0.01125											.	.										
23	0.09889											.	**										
24	-0.02292											.	.										

Tableau IV.1. Identification du modèle simulé.

"." erreur standard

Représentation graphique des rapports standardisés en fonction des valeurs de la variable seuil

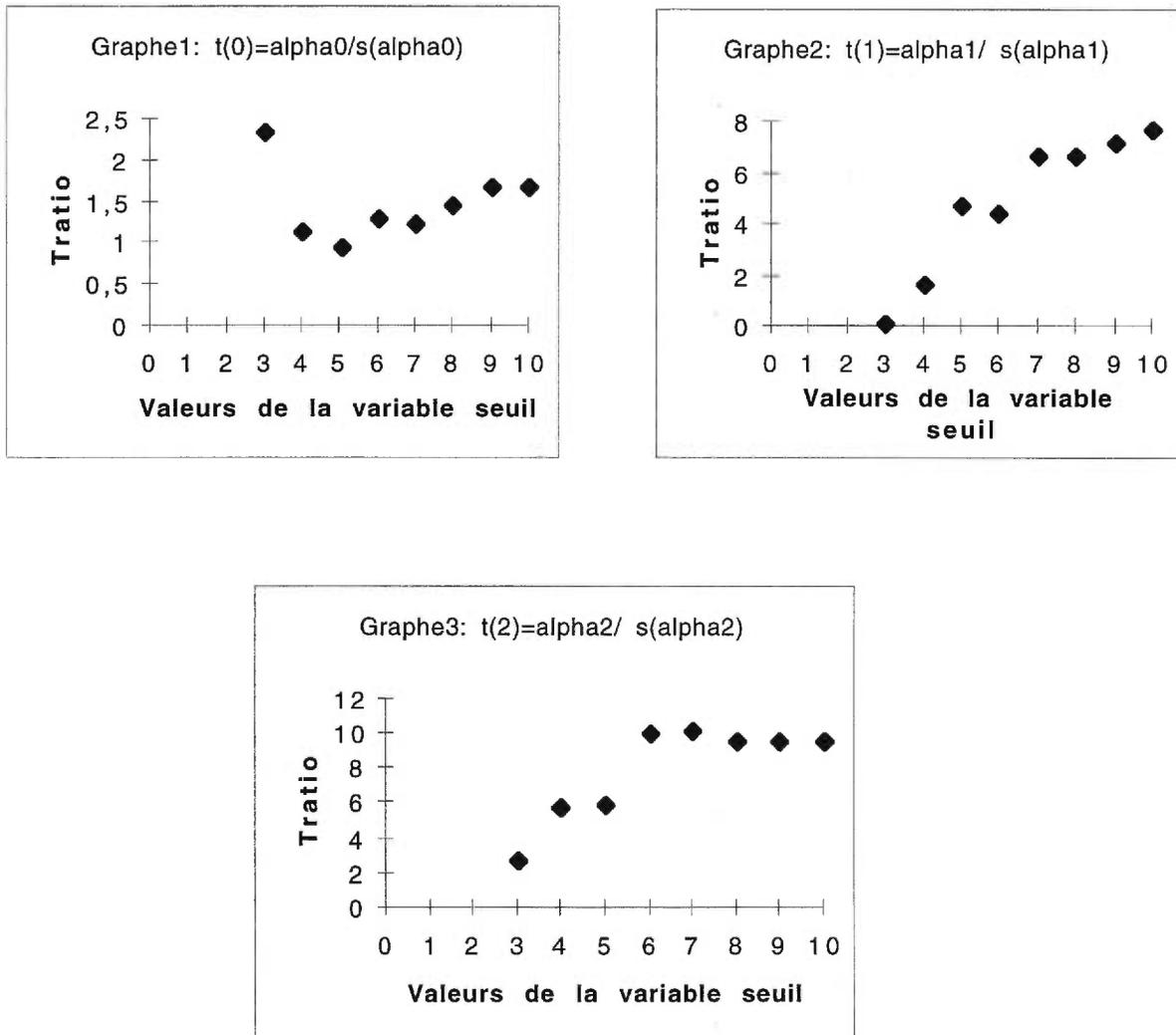


Figure IV.2: Détermination de la valeur du seuil.

Représentation graphique des rapports standardisés en fonction de l'indice des observations ordonnées

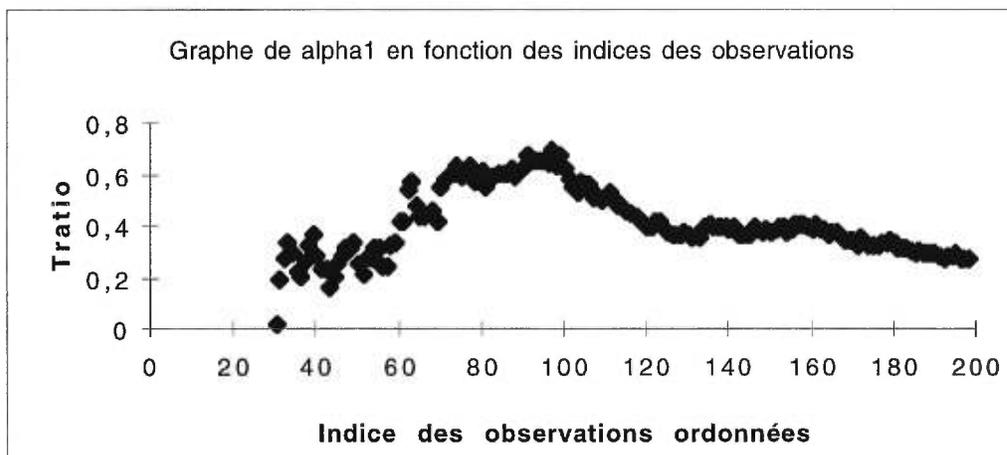
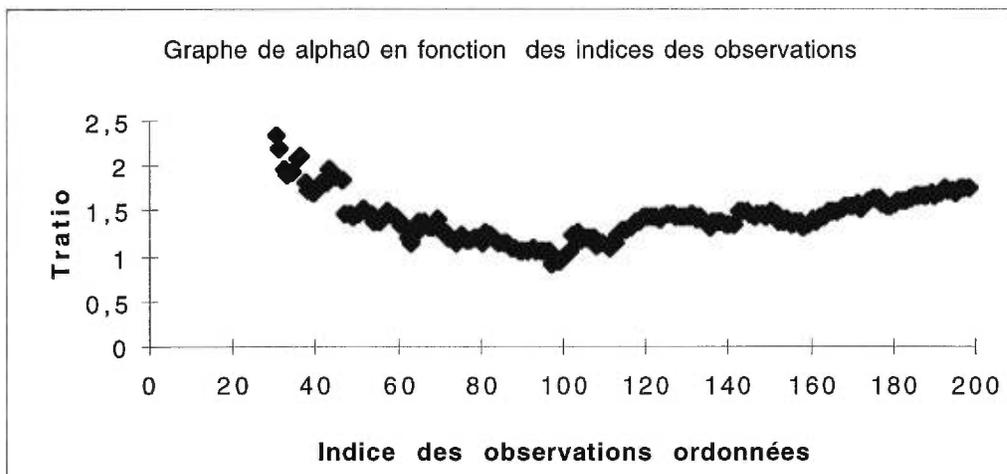


Figure IV.3: Détermination de la valeur du seuil en fonction des indices des observations ordonnées.

**Représentation graphique des estimations
en fonction des valeurs de la variable seuil**

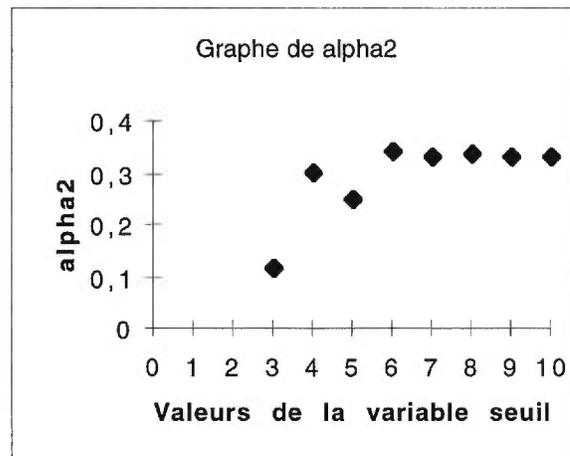
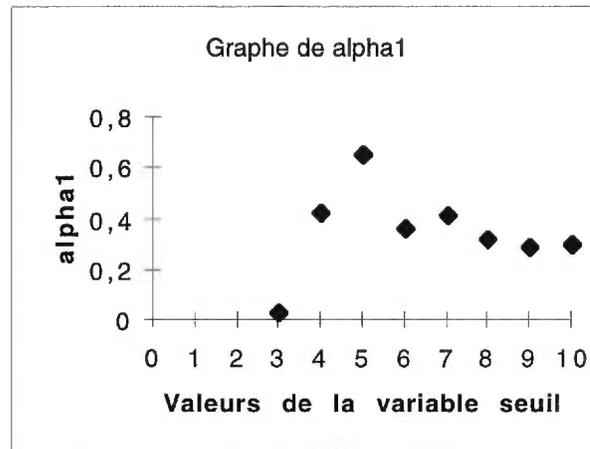


Figure IV.4: Estimations des paramètres en fonction des valeurs de la variable seuil.

Liste des résidus de la première régression

Obs	-2-1-0 1 2	Cook's D	Rstudent	Hat Diag H	Cov Ratio	Dffits
1		0.005	0.3863	0.0932	1.1336	0.1238
2	**	0.017	1.0099	0.0465	1.0481	0.2230
3	****	0.100	2.3683	0.0531	0.9134	0.5606
4	*	0.006	-0.5695	0.0531	1.0794	-0.1348
5		0.004	-0.4290	0.0556	1.0872	-0.1041
6	**	0.034	1.4490	0.0473	1.0133	0.3230
7	*	0.008	-0.7124	0.0465	1.0656	-0.1573
8	**	0.034	-1.4425	0.0473	1.0139	-0.3216
9	*	0.012	-0.8581	0.0473	1.0587	-0.1913
10	****	0.100	2.3683	0.0531	0.9134	0.5606
11		0.000	-0.1388	0.0465	1.0827	-0.0307
12		0.000	-0.1388	0.0465	1.0827	-0.0307
13	**	0.009	-1.2180	0.0189	1.0036	-0.1692
14	**	0.016	1.4864	0.0209	0.9825	0.2173
15		0.001	0.2042	0.0371	1.0713	0.0401
16		0.001	-0.3674	0.0303	1.0606	-0.0650
17	**	0.009	-1.2180	0.0189	1.0036	-0.1692
18	*	0.011	-0.6682	0.0712	1.0962	-0.1851
19		0.002	0.4828	0.0189	1.0449	0.0671
20	*	0.004	-0.7901	0.0209	1.0339	-0.1155
21		0.000	0.0606	0.0243	1.0586	0.0096
22	*	0.013	-0.8014	0.0573	1.0731	-0.1975
23	**	0.013	-1.3651	0.0209	0.9934	-0.1996
24	**	0.012	1.2000	0.0243	1.0105	0.1895
25		0.000	-0.0813	0.0189	1.0527	-0.0113
26	*	0.011	-0.9378	0.0371	1.0426	-0.1841
27		0.000	0.0575	0.0471	1.0839	0.0128
28		0.002	-0.3649	0.0371	1.0681	-0.0716
29		0.000	0.0575	0.0471	1.0839	0.0128
30	*	0.002	-0.7313	0.0136	1.0292	-0.0859
31		0.001	0.2589	0.0263	1.0586	0.0426
32		0.000	0.1136	0.0421	1.0779	0.0238
33		0.000	-0.0257	0.0105	1.0439	-0.0026
34	*	0.001	-0.5879	0.0105	1.0322	-0.0605
35	**	0.008	-1.3030	0.0136	0.9913	-0.1531
36	*	0.002	-0.7313	0.0136	1.0292	-0.0859
37	**	0.023	1.2650	0.0421	1.0240	0.2651
38		0.001	0.3951	0.0136	1.0419	0.0464
39		0.000	0.1158	0.0147	1.0479	0.0141
40	*	0.007	-0.8761	0.0263	1.0348	-0.1441
41	**	0.008	-1.3030	0.0136	0.9913	-0.1531
42		0.001	-0.3070	0.0263	1.0576	-0.0505
43	***	0.016	-1.8887	0.0136	0.9343	-0.2220
44		0.001	0.3951	0.0136	1.0419	0.0464
45		0.001	0.2589	0.0263	1.0586	0.0426
46		0.000	0.1158	0.0147	1.0479	0.0141
47	*****	0.130	3.1176	0.0421	0.7978	0.6534
48		0.000	-0.0257	0.0105	1.0439	-0.0026
49		0.001	0.3951	0.0136	1.0419	0.0464
50	***	0.012	-1.5926	0.0147	0.9662	-0.1946
51	**	0.005	-1.1562	0.0105	0.9997	-0.1189
52	*	0.002	0.6798	0.0147	1.0327	0.0831
53		0.000	0.1136	0.0421	1.0779	0.0238
54	***	0.011	1.5382	0.0136	0.9704	0.1808
55		0.000	-0.1672	0.0136	1.0463	-0.0196
56	***	0.012	-1.5926	0.0147	0.9662	-0.1946
57	**	0.016	-1.0329	0.0421	1.0417	-0.2165
58	*	0.006	0.8275	0.0263	1.0376	0.1361
59		0.001	0.3951	0.0136	1.0419	0.0464
60	****	0.018	2.2805	0.0105	0.8851	0.2346
61		0.001	-0.2522	0.0347	1.0679	-0.0478
62	**	0.014	1.3142	0.0242	1.0011	0.2070
63		0.001	0.3161	0.0347	1.0666	0.0599

64	**		0.023	-1.4031	0.0347	1.0042	-0.2660
65	**		0.014	-1.2538	0.0254	1.0074	-0.2025
66			0.000	-0.1125	0.0254	1.0595	-0.0182
67			0.001	0.1711	0.0562	1.0934	0.0418
68			0.000	0.0298	0.0211	1.0552	0.0044
69	**		0.014	-1.2538	0.0254	1.0074	-0.2025
70		****	0.049	2.0665	0.0347	0.9339	0.3917
71		*	0.003	0.5950	0.0211	1.0432	0.0874
72			0.002	0.4534	0.0254	1.0528	0.0732
73			0.001	0.3127	0.0371	1.0694	0.0614
74			0.004	0.4637	0.0525	1.0826	0.1092
75	**		0.009	-1.1058	0.0211	1.0143	-0.1625
76		*	0.003	0.5950	0.0211	1.0432	0.0874
77			0.001	0.3161	0.0347	1.0666	0.0599
78	**		0.023	-1.4031	0.0347	1.0042	-0.2660
79			0.000	0.0298	0.0211	1.0552	0.0044
80		**	0.010	1.1666	0.0211	1.0098	0.1714
81	****		0.036	-2.2893	0.0211	0.8937	-0.3363
82		*	0.005	0.7392	0.0242	1.0400	0.1164
83			0.001	0.3161	0.0347	1.0666	0.0599
84		*	0.010	0.8850	0.0371	1.0458	0.1738
85			0.001	-0.2564	0.0371	1.0705	-0.0503
86			0.001	0.1711	0.0562	1.0934	0.0418
87		**	0.009	1.0241	0.0254	1.0245	0.1654
88	*		0.009	-0.6854	0.0525	1.0737	-0.1614
89		*	0.005	0.7392	0.0242	1.0400	0.1164
90			0.000	0.0346	0.0778	1.1200	0.0100
91		***	0.048	1.6332	0.0525	1.0007	0.3846
92	**		0.009	-1.1058	0.0211	1.0143	-0.1625
93		*	0.011	0.7478	0.0562	1.0748	0.1825
94			0.000	0.0298	0.0211	1.0552	0.0044
95			0.000	-0.1125	0.0254	1.0595	-0.0182
96	*		0.002	-0.5352	0.0211	1.0455	-0.0786

Tableau IV.5. Résidus de la première régression

Liste des résidus de la deuxième régression

Obs	-2-1-0 1 2	Cook's D	Rstudent	Hat Diag H	Cov Ratio	Dffits
1	***	0.040	1.9946	0.0303	0.9435	0.3528
2	*	0.005	-0.6986	0.0301	1.0472	-0.1231
3	**	0.015	1.4245	0.0224	0.9916	0.2155
4	*	0.006	-0.9038	0.0224	1.0286	-0.1367
5	*	0.005	-0.6986	0.0301	1.0472	-0.1231
6	***	0.046	-1.7700	0.0434	0.9806	-0.3770
7		0.001	-0.3227	0.0224	1.0512	-0.0488
8	***	0.022	1.8088	0.0198	0.9531	0.2572
9		0.000	-0.1568	0.0225	1.0538	-0.0238
10		0.000	0.2158	0.0303	1.0617	0.0382
11	*	0.017	-0.8810	0.0613	1.0725	-0.2251
12	**	0.021	1.1851	0.0434	1.0327	0.2524
13		0.001	-0.3227	0.0224	1.0512	-0.0488
14	**	0.010	1.2115	0.0198	1.0059	0.1723
15		0.002	0.4632	0.0301	1.0560	0.0816
16	*	0.006	-0.9038	0.0224	1.0286	-0.1367
17	***	0.024	-1.5423	0.0303	0.9894	-0.2728
18	*	0.004	-0.7363	0.0225	1.0373	-0.1116
19		0.000	-0.1171	0.0301	1.0625	-0.0206
20		0.001	-0.3227	0.0224	1.0512	-0.0488
21	*	0.009	-0.9483	0.0303	1.0344	-0.1677
22		0.000	-0.1171	0.0301	1.0625	-0.0206
23	**	0.013	-1.3235	0.0225	1.0000	-0.2007
24		0.001	-0.3227	0.0224	1.0512	-0.0488
25		0.000	-0.1171	0.0301	1.0625	-0.0206
26	**	0.011	1.0484	0.0301	1.0280	0.1848
27		0.001	0.1806	0.0852	1.1258	0.0551
28	*	0.006	-0.9038	0.0224	1.0286	-0.1367
29	*	0.006	-0.9038	0.0224	1.0286	-0.1367
30		0.004	-0.4941	0.0431	1.0693	-0.1049
31		0.001	0.1806	0.0852	1.1258	0.0551
32		0.000	0.0091	0.0434	1.0777	0.0019
33		0.001	0.4211	0.0225	1.0489	0.0639
34		0.000	-0.1171	0.0301	1.0625	-0.0206
35	*	0.006	-0.9038	0.0224	1.0286	-0.1367
36	**	0.008	1.0034	0.0225	1.0228	0.1521
37		0.000	-0.0622	0.0104	1.0417	-0.0064
38	***	0.017	1.9022	0.0142	0.9380	0.2280
39	*	0.003	0.8853	0.0119	1.0187	0.0971
40	*	0.001	-0.6375	0.0104	1.0289	-0.0654
41		0.000	0.1030	0.0186	1.0501	0.0142
42		0.001	0.1868	0.0566	1.0917	0.0458
43	*	0.003	0.6809	0.0186	1.0356	0.0936
44		0.001	-0.4328	0.0142	1.0398	-0.0519
45	*	0.003	0.6809	0.0186	1.0356	0.0936
46	****	0.062	-2.4988	0.0304	0.8834	-0.4427
47		0.000	-0.0226	0.0373	1.0708	-0.0044
48		0.000	0.1427	0.0142	1.0451	0.0171
49		0.002	0.4773	0.0304	1.0559	0.0846
50	***	0.017	1.9022	0.0142	0.9380	0.2280
51	*	0.001	-0.6375	0.0104	1.0289	-0.0654
52		0.001	-0.4740	0.0186	1.0432	-0.0652
53	**	0.004	1.0924	0.0104	1.0046	0.1121
54	*	0.003	-0.8445	0.0119	1.0209	-0.0926
55	*	0.002	0.7197	0.0142	1.0293	0.0863
56	****	0.052	2.2742	0.0304	0.9113	0.4029
57		0.000	0.1427	0.0142	1.0451	0.0171
58	***	0.012	-1.6032	0.0142	0.9676	-0.1922
59	**	0.009	1.4723	0.0119	0.9770	0.1614
60	**	0.038	1.3757	0.0566	1.0319	0.3371
61		0.000	0.1427	0.0142	1.0451	0.0171
62	*	0.012	0.8607	0.0475	1.0582	0.1922
63		0.000	-0.1735	0.0134	1.0440	-0.0202

64	*		0.007	-0.7976	0.0299	1.0422	-0.1399
65		***	0.011	1.5706	0.0134	0.9699	0.1833
66	***		0.011	-1.5470	0.0137	0.9723	-0.1824
67			0.000	0.1965	0.0137	1.0440	0.0232
68	*		0.004	-0.9579	0.0137	1.0164	-0.1129
69		*	0.002	0.6090	0.0184	1.0384	0.0833
70		*	0.007	0.8204	0.0285	1.0396	0.1406
71	***		0.018	-1.7229	0.0184	0.9603	-0.2357
72			0.001	-0.3788	0.0137	1.0407	-0.0446
73			0.000	-0.1735	0.0134	1.0440	-0.0202
74		*	0.002	0.6090	0.0184	1.0384	0.0833
75	**		0.008	-1.1284	0.0184	1.0103	-0.1544
76		*****	0.033	2.8075	0.0134	0.8285	0.3276
77	***		0.011	-1.5470	0.0137	0.9723	-0.1824
78	**		0.009	-1.1697	0.0192	1.0083	-0.1636
79			0.000	-0.0498	0.0669	1.1047	-0.0133
80			0.001	-0.3413	0.0285	1.0574	-0.0585
81		***	0.020	1.7478	0.0192	0.9586	0.2444
82	*		0.002	0.5690	0.0192	1.0407	0.0796
83		**	0.009	1.1928	0.0184	1.0058	0.1632
84	**		0.008	-1.3352	0.0134	0.9899	-0.1558
85	**		0.008	-1.1284	0.0184	1.0103	-0.1544
86			0.001	-0.2870	0.0289	1.0589	-0.0495
87		*	0.004	0.6666	0.0279	1.0463	0.1130
88	*		0.008	-0.8697	0.0289	1.0374	-0.1499
89	*		0.006	-0.7030	0.0322	1.0493	-0.1283
90		*	0.003	0.5013	0.0350	1.0601	0.0955
91			0.000	0.0859	0.0279	1.0603	0.0146
92			0.000	0.1274	0.0463	1.0805	0.0281
93			0.002	-0.3296	0.0417	1.0722	-0.0688
94		*	0.008	0.6355	0.0564	1.0792	0.1554
95	**		0.011	-1.0788	0.0279	1.0237	-0.1829
96	***		0.033	-1.5134	0.0417	1.0038	-0.3157
97		**	0.015	1.2542	0.0279	1.0111	0.2126
98			0.001	0.1831	0.0567	1.0918	0.0449
99		*	0.018	0.9445	0.0577	1.0647	0.2337
100	*		0.033	-0.9622	0.0955	1.1081	-0.3127
101			0.000	-0.1028	0.1192	1.1701	-0.0378
102			0.007	0.4572	0.0937	1.1302	0.1470

Tableau IV.6. Résidus de la deuxième régression.

BIBLIOGRAPHIE

AKAIKE, H., (1994). "A New Look at Statistical Model Identification", IEEE Transactions on Automatic Control, 19, 716-722.

AL-OSH, M. A. and ALZAID, A. A., (1987). "First order integer-valued autoregressive ($INAR(1)$) process", Journal of Time Series Analysis, 8, 261-275.

ALZAID, A. A., and AL-OSH, M. A., (1990). "An integer-valued p^{th} order autoregressive structure ($INAR(p)$) process", Journal of Applied Probability, 27, 314-324.

BILLINGSLEY, P., (1968). "Convergence of Probability Measures", Wiley, New York.

BOX, G.E.P., and JENKINS, G. M., (1976). "Time Series Analysis: Forecasting and Control", Holden-Day, San Francisco.

BROCKWELL, P. J. and DAVIS, R. A., (1991). "Time Series: Theory and Methods", (second edition), Springer-Verlag, New York.

BRILLINGER, D. R., (1969). "The calculation of cumulants via conditioning", Ann. Inst. Statistic. Math., 21.

CHAN, N. H., (1988). "On the existence of the stationary and ergodic $NEAR(p)$ model", J. T. S. A., 9, 319-328.

CHAN, K. S., and TONG, H., (1986). "On estimating thresholds in autoregressive models", Journal of Time Series Analysis, 7, 179-190.

DIEBOLT, J., (1990). "Testing the functions defining a nonlinear autoregressive time series", Stoch. Proc. and their Appl., 36, 85-106.

DIEBOLT, J., and LAIB, N., (1991). "Un principe d'invariance faible pour l'étude d'un test non paramétrique relatif à un modèle de régression non-linéaire", Rapport Technique L. S. T. A. Paris VI.

- DAVIES, N., and PETRUCELLI, J. D.,** (1986). "*Detecting nonlinearity in time series*", *The Statistician* 35, 271-280.
- DION, J. P., GAUTHIER G. and LATOUR, A.,** (1993). "*Branching processes with immigration and integer-valued time series*", *Serdica Math. J.* 21, 123-136.
- DU J. G. and LI Y.,** (1991). "*The integer-valued autoregressive (INAR(p)) model* ", *J. Times Series. Anal.* 12, 129-142.
- ERTEL, J. E., and FOWLKES, E. B.,** (1976). "*Some Algorithms for Linear Spline and Piecewise Multiple Linear Regression*", *Journal of the American Statistical Association*, 71. 640-648.
- GAUTHIER, G., and LATOUR, A.,** (1994). "*Convergence forte des estimateurs des paramètres d'un processus GNAR(p)*", *Ann. sc.math, Québec*, 18, 37-59.
- GOODWIN, G. C., and PAYNE, R. L.,** (1977). "*Dynamic System Identification: Experiment Design and Data Analysis*", Academic Press, New York.
- GRIMMETT, G. R. and STIRZAKER, D. R.,** (1983). "*Probability and random process*", Clarendon Press, Oxford.
- GUÉGAN, D.,** (1994). "*Séries chronologiques non-Linéaires à temps discret*", Economica, Paris.
- HALL, P. and HEYDE, C. C.,** (1980). "*Martingale limit theory and its application*", Academic Press, New York.
- HILI, O.,** (1992). "*Sur les modèles autorégressifs à seuil*", *C. R. A. S. Série I.*, 573-576.
- KEENAN, D. M.,** (1985). "*A Tukey Nonadditivity-Type Test for Time series Nonlinearity*", *Biometrika*, 72, 39-44.
- KLIMKO, L. A., and NELSON, P. J.,** (1978). "*On conditional least squares estimation for stochastic process*", *Ann. Statistic*, 6.

- LAI, T. L., and WEI, C. Z.,** (1982). "*Least Squares Estimates in Stochastic Regression Models With Applications to Identification and Control of Dynamic Systems*", *Annals of Statistics*, 10, 154–166.
- LATOURE, A.,** (1997). "*The Multivariate GNAR(p) Process*", *Adv. Appl. Prob.* 29, 228–248.
- LORD, Y.,** (1990). "*Application de la méthode de bootstrap à la régression linéaire multiple*", Mémoire de maîtrise ès sciences, Université de Montréal.
- MCLEOD, A. J., and LI W. K.,** (1983). "*Diagnostic checking ARMA time series models using squared residual autocorrelations*", *J. T. S. A.* 4, 269–273.
- OZAKI, T.,** (1981). "*Nonlinear threshold autoregressive models for nonlinear random vibrations*", *J. Appl. Prob.* 18, 443–451.
- PETRUCCELLI, J., and DAVIS, N.,** (1986). "*A Portmanteau Test for Self-Exciting Threshold Autoregressive-Type Nonlinearity in Time Series*", *Biometrika*, 73, 687–694.
- PETRUCCELLI, J., and WOOLFORD, S. W.,** (1984). "*A threshold AR(1) model*", *Journal of Applied Probability*, 21, 270–286.
- PHAM, T., CHAN, H and TONG.,** (1991). "*Least squares estimates of threshold parameters*", *Statistica Sinica* 1, 361–371.
- POLE, A. M., and SMITH, A. F. M.,** (1985). "*A Bayesian Analysis of Some Threshold Switching Models*", *Journal of Econometrics*, 29, 97–119.
- QUANDT, R. E.,** (1960). "*Tests of the Hypothesis That a Linear Regression System Obeyes Two Separate Regimes*", *Journal of the American Statistical Association*, 55, 324–330.
- SHABAN, S. A.,** (1980). "*Change-Point Problem and Two-Phase Regression: An Annotated Bibliography*", *International Statistical Review*, 48, 83–93.
- SIEGMUND, D.,** (1988). "*Confidence Sets in Change-Point Problems*", *International Statistical Review*, 56, 31–48.

STEUTEL F. W., and VAN HARN, K., (1979). "*Discrete analogues of self-decomposability and stability*", *Ann. of Prob.*, 7, 893–899.

TONG, H., (1978). "*On a Threshold Model in Pattern Recognition and Signal Processing*", ed. C. H. Chen, Amsterdam: Sijhoff & Noordhoff.

TONG, H., (1983). "*Thresold Models in Non-Linear Time Series Analysis*", (Lecture Notes in Statistics, No 21), Springer-Verlag, New York.

TONG, H. and LIM, K. S., (1980). "*Thresold autoregression, limit cycles and cyclical data*", *Journal of the Royal Statistical Society, Series B*, 42, 245–292.

TSAY, R. S., (1986). "*Nonlinearity Tests for Time Series*", *Biometrika*, 73, 461–466.

TSAY, R. S., (1989). "*Testing and Modeling Threshold Autoregressive Process*", *Journal of the American Statistical Association*, 84, 231–240.