

REPUBLIQUE ALGERIENNE DEMOCRATIQUE et POPULAIRE.  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique.



UNIVERSITE MOULOU D MAMMERI, TIZI-OUZOU  
Faculté des Sciences Département de Mathématiques

\*\*\*

# MEMOIRE DE MASTER en Mathématiques

Spécialité : Mathématiques appliquées à la gestion

## Modélisation des dépenses de santé en Algérie : Approche ARDL

présenté par

**INGRACHEN Tinhinane**

Devant le jury composé de

\*\*\*

Fazia BEDOUHENE	Professeur	U.M.M.T.O	Présidente
Abdelghani HAMAZ	MCA	U.M.M.T.O	Rapporteur
Farida ACHEMINE	MCA	U.M.M.T.O	Examinatrice
Mouloud GOUBI	MCB	U.M.M.T.O	Examineur

\*\*\*

soutenu le 08 / 10 / 2020

## *Remerciements*

La réalisation de ce mémoire a été possible grâce au concours de plusieurs personnes auxquelles je voudrais témoigner toute ma reconnaissance.

Tout d'abord je remercie le bon Dieu tout puissant de m'avoir donné le courage, la patience et la volonté d'aller vers l'avant malgré les circonstances.

Ma grande gratitude va aussi à mon enseignant et promoteur Mr HAMAZ Abdelghani, ses conseils judicieux et ses encouragements ont permis à ce travail d'aboutir, ses capacités scientifiques et ses compétences étaient mon plus grand support. Travailler sous sa direction était pour moi un grand honneur. Je tiens à le remercier pour sa patience et sa disponibilité.

Je tiens aussi à remercier les membres du jury, pour leurs présente lecture attentive de mon mémoire ainsi que pour les remarques qu'ils m'adresseront lors de cette soutenance afin d'améliorer mon travail.

Je souhaite aussi remercier la directrice du Laboratoire de Mathématiques Pures et Appliquées (*LMPA*), pour sa générosité et son accueil.

Je désire aussi remercier les enseignants du département Mathématiques de l'université Mouloud Mammeri, leurs générosité et leurs soutien m'obligent de leurs témoigner mon profond respect et ma loyale considération.

Mes remerciements vont aussi à l'encontre de Mr KOURAT Hocine et toute l'équipe administrative du département Mathématiques de l'université Mouloud Mammeri, pour leurs disponibilité et générosité.

Je remercie aussi tout(e)s mes ami(e)s pour le soutien que vous m'avez apporté et votre sincère amitié : Syla, Karima, Amina, Said, Farid, Naim et Mohamed.

Sans oublier des personnes que j'ai connu durant ma formation de master et qui sont devenus de très bons ami(e)s et une belle rencontre : Louiza, Lynda, Djafar et Rachid. Je vous remercie pour votre soutien, sincérité, disponibilité, conseils et surtout générosité.

Enfin, un très grand merci à toute ma famille qui m'a gratifié son amour et fourni les motivations. Je leurs adresse toute ma gratitude du fond du coeur.

## *Dédicaces*

Que ce travail témoigne de mes respects :

A mes très chers parents, mon exemple éternel et ma source de motivation, votre présence, soutien et sacrifices m'ont permis d'être celle d'aujourd'hui, je suis redevable d'une éducation dont je suis fière. Vous êtes gravés en lettres d'or dans mon cœur. Je porte et je porterai toujours dans mon cœur la confiance que vous m'avez accordé.

A mes très chères soeurs : Taous, Melissa et Letissia, ma source de bonheur et de confiance, pour votre présence, motivations, soutien, encouragements et remarques ; je vous souhaite un radieux avenir, plein de bonheur et de succès.

A mon très cher petit frère Koceila, ma source de bonheur, pour ta présence, soutien et notamment ton côté protecteur, je te souhaite un radieux avenir, que dieu te garde pour moi.

A mon frère Farid, un grand frère que la vie m'a offert, pour tes conseils, tes encouragements et ton soutien moral, je te souhaite tout le bonheur et la réussite, je suis reconnaissante à tout jamais.

A tous ceux qui sont chers de près et de loin, ils vont trouver ici l'expression de mes sentiments de respect et de reconnaissance pour leurs soutien qu'ils n'ont jamais cessé de me porter.

Enfin, je voudrai dédier ce modeste travail à l'école de la vie, l'école qui m'a beaucoup servi à découvrir et apprendre beaucoup de choses, persévérer, ne pas lâcher, estimer soi même et surtout faire face aux difficultés avec un sourire étampé sur le visage.

## *Liste des abréviations*

- ACF : Autocorrélation simple.
- ADF : Dickey-Fuller Augmenté.
- AIC : Critère d'information d'Akaike.
- AR : Autoregressif.
- ARDL : Autoregressif distributed lag.
- ARMA : Autoregressif moving average.
- DF : Dickey-Fuller simple.
- DL : Distributed lag.
- DS : Difference stationary.
- EMV : Estimateur du maximum de vraisemblance.
- i.i.d : indépendantes et identiquement distribuées.
- MA : Moyenne mobile.
- MCO : Moindre carrés ordinaires.
- ONS : Office National de santé.
- PACF : Autocorrélation partielle.
- PP : test de Phillips-Perron.
- SIC : critère d'information de Schwarz.
- TS : trend stationary.
- VAR : vector autoregressif.
- VECM : vector error correction model.

# Table des figures

3.1	Evolution des dépenses de santé per capita (en log) . . . . .	51
3.2	Evolution du PIB par habitant . . . . .	51
3.3	Evolution de la densité médicale . . . . .	52
3.4	Evolution de la densité de lits d'hôpitaux . . . . .	52
3.5	Evolution de la prise en charge publique . . . . .	53
3.6	Evolution de la part des 60 ans et plus dans la population totale . . . . .	54
3.7	Evolution de la population totale . . . . .	54
3.8	Le graphique du critère d'information Schwarz (SIC) . . . . .	67
3.9	Test de normalité des résidus . . . . .	71
3.10	Courbe de la somme cumulée des résidus (CUSUM) . . . . .	72

# Liste des tableaux

3.1	Application du test ADF sur la série log (DS) . . . . .	56
3.2	Application du test ADF sur la série $\Delta$ log (DS) . . . . .	57
3.3	Application du test ADF sur la série log (DM) . . . . .	58
3.4	Application du test ADF sur la série $\Delta$ log (DM) . . . . .	59
3.5	Application du test ADF sur la série log (POP) . . . . .	60
3.6	Application du test ADF sur la série log (PCP) . . . . .	61
3.7	Application du test ADF sur la série log (PIBH) . . . . .	62
3.8	Application du test ADF sur la série $\Delta$ log (PIBH) . . . . .	63
3.9	Application du test ADF sur la série log (POP60) . . . . .	64
3.10	Application du test ADF sur la série log (POP60) . . . . .	65
3.11	Application du test ADF sur la série log (DLH) . . . . .	65
3.12	Résultats du test ADF de racine unitaire . . . . .	66
3.13	Estimation du modèle ARDL(1.0.0.1.1.1.1) . . . . .	67
3.14	Bounds-test . . . . .	68
3.15	Résultats de l'estimation de la relation à long terme . . . . .	68
3.16	Résultats de l'estimation de la relation à court terme . . . . .	69
3.17	Test d'autocorrélation des erreurs . . . . .	71

## *Résumé*

Vouloir identifier la relation entre les variables nous mène au cadre multivarié. Il est donc nécessaire de définir les outils statistiques qui nous permettront de bien mener notre étude et application. Au cours de notre travail, nous allons s'appuyer sur l'approche ARDL, les modèles VAR et VECM comme facteurs de rôle très important pour l'atteinte de notre objectif : l'analyse de l'existence d'une relation de long terme ou de court terme entre les variables.

Enfin, un exemple d'application sur des données macroéconomiques réelles collectées de différentes sources connues sera présenté.

**mots clés :** Série chronologique, Stationnarité, ARDL, VAR , Causalité, Cointègation, Estimation, Test ADF.

## *Abstract*

Want to identify the relationship between variables leads us to the multivariate framework. It is therefore necessary to define the statistical tools that will allow us to carry out our study and application. In the course of our work we will rely on the ARDL approach, the VAR and VECM models such as very important factors for the achievement of our objective : to analyze the existence of a long-term or short-term relationship between the variables.

Finally, an example of application on macroeconomic data actual collected from various known sources will be presented.

**Key words :** Time series, stationarity, ARDL, VAR, Causality, Cointegration, Forecasting, ADF Test.

# Table des matières

<b>Introduction</b>	<b>8</b>
<b>1 Notions Préliminaires</b>	<b>11</b>
1.1 Introduction . . . . .	11
1.2 Concepts de base des séries chronologiques . . . . .	11
1.2.1 Processus aléatoire . . . . .	11
1.2.2 Définition d'une série chronologique . . . . .	12
1.2.3 Composantes d'une série chronologique . . . . .	12
1.2.4 Domaine d'application et Objectifs de l'analyse d'une série chronologique . . . . .	13
1.2.5 La Stationnarité . . . . .	13
1.2.6 Processus Bruit Blanc . . . . .	14
1.2.7 Caractéristiques d'une série chronologique . . . . .	14
1.2.8 Les opérateurs $B$ et $\Delta$ . . . . .	18
1.3 Processus aléatoires stationnaires . . . . .	19
1.3.1 Processus MA ( $q$ ) . . . . .	20
1.3.2 Les caractéristiques d'un MA( $q$ ) . . . . .	20
1.3.3 Prévission à l'aide d'un MA( $q$ ) . . . . .	21
1.3.4 Processus AR ( $p$ ) . . . . .	22
1.3.5 Existence des processus AR(1) . . . . .	22
1.3.6 Les caractéristiques d'un AR( $p$ ) . . . . .	23
1.3.7 Estimation dans le modèle autorégressif AR( $p$ ) . . . . .	24
1.3.8 Prévission à l'aide d'un AR( $p$ ) . . . . .	26
1.3.9 Processus ARMA ( $p,q$ ) . . . . .	26
1.3.10 Caractéristiques du modèle ARMA ( $p,q$ ) . . . . .	28
1.3.11 Estimation dans modèle ARMA( $p,q$ ) . . . . .	29
1.3.12 Prévission à l'aide d'un modèle ARMA( $p,q$ ) . . . . .	30
1.3.13 Méthodologie de Box-Jenkins . . . . .	30
1.4 Processus aléatoires non-stationnaires . . . . .	31
1.4.1 Processus TS(Trend Stationary) . . . . .	31
1.4.2 Processus DS( Difference Stationary) . . . . .	31
1.5 Tests de racines unitaires . . . . .	32
1.5.1 Test de Dickey-Fuller . . . . .	33

1.5.2	Test de Phillips et Perron (1988) . . . . .	36
<b>2</b>	<b>Modélisation VAR et ARDL</b>	<b>37</b>
2.1	Introduction . . . . .	37
2.2	Le modèle ARDL . . . . .	37
2.2.1	Caractéristiques du modèle ARDL . . . . .	38
2.2.2	La méthodologie du modèle ARDL . . . . .	39
2.3	Modélisation VAR . . . . .	41
2.3.1	Représentation générale d'un modèle VAR . . . . .	41
2.3.2	La fonction d'autocovariance d'un VAR(p) . . . . .	43
2.3.3	Estimation d'un modèle VAR . . . . .	44
2.3.4	Détermination du nombre de retards . . . . .	45
2.3.5	Prévision dans un VAR stationnaire . . . . .	45
2.3.6	Causalité . . . . .	46
2.4	Le concept de cointégration . . . . .	48
2.4.1	Les propriétés de l'ordre d'intégration d'une série chronologique . . . . .	48
2.4.2	Représentation du modèle à correction d'erreur(VECM) . . . . .	49
<b>3</b>	<b>La modélisation des dépenses sanitaires en Algérie</b>	<b>50</b>
3.1	Introduction . . . . .	50
3.2	Choix des variables et données utilisées . . . . .	50
3.2.1	Spécification du modèle . . . . .	54
3.3	Analyse Empirique . . . . .	55
3.3.1	Etude de la stationnarité . . . . .	56
3.3.2	Test de conitègration et estimation du modèle ARDL . . . . .	66
3.3.3	L'estimations de long et de court terme selon le modèle ARDL . . . . .	68
3.3.4	Validation du modèle . . . . .	70
	<b>Conclusion</b>	<b>72</b>
	<b>Bibliographie</b>	<b>74</b>

# Introduction Générale

Partout dans le monde, la part des dépenses publiques de santé dans le PIB et dans le total des dépenses publiques a beaucoup augmenté au cours de ces trente dernières années. Jusqu'ici, les efforts de réforme ont permis de ralentir cette tendance, mais n'ont pas été suffisants pour supprimer toutes les pressions sous-jacentes qui contribuent à la croissance des dépenses. A moins d'un changement majeur de la politique des pouvoirs publics en matière de santé, les ponctions des dépenses de santé sur les budgets publics continueront sans doute d'augmenter dans les années à venir, en même temps que les dépenses au titre des pensions du régime public. Par conséquent, la réforme du secteur de la santé représente dans de nombreux pays membres un problème politique majeur. Faute de réformes plus fondamentales que celles appliquées dans les années 80, il apparaît inévitable d'augmenter les impôts ou de réduire les prestations.

Dans le cas d'Algérie, les pouvoirs publics sont interpellés sur la question des inégalités d'accès aux soins qui se creusent en Algérie, avec le développement d'un système à double vitesse, leur prise en charge est indispensable. De plus, la crise économique actuelle que traverse le pays ne peut que renforcer les difficultés, déjà présentes à assurer la pérennité aussi bien du système de santé que celui de la sécurité sociale.

Il est de la responsabilité des pouvoirs publics de prendre des décisions rationnelles permettant de faire face à une croissance insoutenable des dépenses de santé. A priori, cette tendance à la hausse des dépenses amène à s'interroger sur les facteurs qui y sont à l'origine et qui représentent, à notre sens, la clé de voûte d'une politique avérée de régulation des dépenses de santé. Manifestement, plusieurs déterminants exercent une influence plus ou moins importante sur le dynamisme des dépenses de santé. Les modèles explicatifs de l'évolution de la dépense de santé mettent généralement en avant trois facteurs principaux : des facteurs de demande (niveau de vie, vieillissement de la population et degrés de couverture), des facteurs d'offre (progrès technique et densité médicale) et enfin des facteurs institutionnels (organisation et financement du système de soins).

Afin de bien comprendre la relation entre les dépenses sanitaires et ces facteurs, une étude économétrique sera mise en oeuvre en se basant sur des fondements théoriques des séries chronologiques et de modèles développés permettant de déterminer le type et le sens de la relation existante entre les dépenses sanitaires et ces facteurs spécialement en Algérie.

Pour se faire, on fait recours à l'approche ARDL (Autoregressif Distributed Lag) ou modèle à retard échelonné. Une approche très utilisée car elle est considérée par beaucoup d'économistes comme un des concepts nouveaux et relativement simple. Cette dernière est fondée sur deux concepts primordiales : le concept de cointégration spécifiée par son pouvoir d'identifier et de définir d'une manière rigoureuse et claire la véritable relation existante entre deux variables en faisant recours au vecteur de cointégration qu'on peut définir à partir des modèles utilisés dans le cas multivarié, on prend comme principal exemple : les modèles VAR (Vector autoregressif) qui sont des modèles très efficaces, en s'appuyant aussi sur le Bound test développé par *Pesaran et al* et un deuxième concept important c'est la causalité, proposée par Granger 1969 et Sims 1980.

Nous avons structuré notre mémoire de recherche en deux parties : théorique qui comporte deux chapitres :

Dans le premier chapitre de ce mémoire, nous présentons les fondements théoriques des séries chronologiques (définition, composantes, objectifs et domaines d'application, stationnarité , processus stationnaires et non stationnaires, Box-jenkins).

le deuxième chapitre porte sur les modèles ARDL, VAR et VECM, dont les utilisations sont fréquentes pour modéliser des séries temporelles multivariées. Après avoir donné la représentation générale, on s'intéresse aux problèmes d'estimation et de prévision liés à ces processus. Ce chapitre abordera aussi des concepts très importants dans la modélisation : le concept de cointégration et de causalité.

La partie pratique est présentée au troisième chapitre. Il s'agit de l'analyse économétrique des dépenses sanitaires en Algérie de 1990 à 2016, avec le modèle VAR ou VECM et ARDL, à l'aide du logiciel Eviews 10 qui est un logiciel très pratique, très adopté par les entreprises, il aide à étudier des phénomènes économiques et tester l'existence d'une relation à long ou à court terme entre les variables.

Ce mémoire sera achevé par une conclusion portant des perspectives.

# Chapitre 1

## Notions Préliminaires

### 1.1 Introduction

Les séries chronologiques sont considérées à tort comme étant une branche exclusive de l'économie, leur célébrité dans de variés domaines est preuve de leur grande utilité et importance dans l'étude de l'évolution d'un phénomène en fonction du temps telque les ventes en finance, la production en industrie et autres .

L'objectif est de décrire, expliquer et prévoir ce qui signifie effectuer des prévisions en partant de l'hypothèse que le passé pourrait être garant du futur.

Ce chapitre a un but introductif, au cours du quel l'attention sera accentuée sur certains concepts fondamentaux des séries chronologiques et modèles de leurs représentation, donc on va faire en sorte de présenter un bref aperçu sur des généralités liées à notre prochaine étude.

### 1.2 Concepts de base des séries chronologiques

#### 1.2.1 Processus aléatoire

Le processus aléatoire, appelé aussi processus stochastique, est une suite de variables aléatoires réelles indexées par le temps  $t$  noté  $\{X_t; t \in \mathbb{I}\}$  avec  $\mathbb{I} \subset \mathbb{Z}$ . L'ensemble  $\mathbb{I}$  est appelé espace de temps. Ce dernier peut être :

- **Discret** : ( nombre de voyageurs d'un vol d'avion, température maximale, ... ), les dates d'observations sont souvent équidistantes : par exemple relevés mensuels, trimestriels.
- **Continu** : ( signal radio, résultat d'un électrocardiogramme, ...), l'indice de temps est à valeurs dans un intervalle de  $\mathbb{R}$  et on dispose d'une infinité d'observations issues d'un processus où  $\mathbb{I}$  est un intervalle de  $\mathbb{R}$  . Un tel processus est dit à temps continu.

Le processus aléatoire est aussi défini comme une famille de variables aléatoires  $X$  as-

soit à chaque  $\omega \in \Omega$  une réalisation  $X(\omega)$ , un processus stochastique  $\{X_t; t \in \mathbb{Z}\}$  associé à chaque  $\omega$  une fonction (ou trajectoire) qu'on note  $X_t(\omega)$  avec  $t \in \mathbb{Z}$

$$\begin{aligned} \mathbb{Z} &\longrightarrow \mathbb{E} \\ t &\longrightarrow X_t(\omega) \end{aligned}$$

où

$\Omega$  : l'ensemble fondamental.

$\mathbb{E}$  : l'espace d'arrivé des variables aléatoires  $X$ .

### 1.2.2 Définition d'une série chronologique

Série chronologique, chronique ou encore temporelle est une suite  $(x_1, x_2, \dots, x_n)$  d'observations enregistrées à un temps spécifique  $t$ . Le nombre  $n$  est appelé la longueur de la série.

Cette dernière peut être journalière, mensuelle, trimestrielle ou annuelle.

Mathématiquement, une série chronologique est définie comme étant un processus aléatoire noté  $\{X_t; t \in \mathbb{I}\}$  avec  $t \in \mathbb{I}$  où  $\mathbb{I} \subset \mathbb{Z}$  espace des temps.

### 1.2.3 Composantes d'une série chronologique

Une série chronologique  $X_t$  est considérée comme la résultante de différentes composantes fondamentales :

- **La tendance** : ou trend notée  $f_t$ , elle représente l'évolution à long terme de la série étudiée et traduit son comportement "moyen".
- **La composante saisonnière** : ou saisonnalité notée  $S_t$ , correspond à un phénomène qui se répète à intervalles de temps réguliers (périodique).  
En général, c'est un phénomène saisonnier d'où le terme de variations saisonnières .
- **La composante résiduelle** : bruit ou résidu notée  $\varepsilon_t$ , correspond à des fluctuations irrégulières. En général de faible intensité mais de nature aléatoire. On parle aussi d'aléas.
- **La composante cyclique** : notée  $C_t$ , correspond à un phénomène qui se répète sur des durées qui ne sont pas fixes, généralement plus longues. Il est généralement très difficile de dissocier tendance et cycle.

- **Des phénomènes accidentels** : (grèves, conditions météorologiques exceptionnelles, crash financier) peuvent notamment intervenir.

### 1.2.4 Domaine d’application et Objectifs de l’analyse d’une série chronologique

L’étude d’une série chronologique est une démarche très importante et utile connaissant une vaste utilisation dans de nombreux domaines d’application comme l’économie, la médecine (étude du rythme cardiaque), la finance (ventes, bourses), l’industrie (production, consommation), sociologie (chômage, grèves) et autres.

Cette étude permet d’analyser, décrire, expliquer et prévoir un phénomène évoluant dans le temps et notamment d’en tirer des conséquences et conclusions qui serviront dans la prise de décisions à partir des séries observées.

L’objectif principal de cette étude est la prévision qui consiste à prévoir des valeurs futures  $X_{t+h}$  de la série à partir des valeurs déjà observées  $(X_1, X_2, \dots, X_t)$  en suivant les étapes suivantes :

1. **Proposer un modèle probabiliste afin de représenter les données.**
2. **Estimation et Vérification** : Il s’agit d’estimer les paramètres du modèle choisi et vérifier la qualité d’ajustement aux données.
3. **Application du modèle** : C’est la prévision, elle peut être ponctuelle ou sous forme d’un intervalle de prévision.

### 1.2.5 La Stationnarité

Une série chronologique est dite non stationnaire si elle se trouve modifiée dans le temps et stationnaire si elle reste elle même .

#### Stationnarité Stricte

Un processus aléatoire  $\{X_t; t \in \mathbb{Z}\}$  est dit strictement stationnaire si pour tout  $(t_1, t_2, \dots, t_n)$  avec  $t_i \in \mathbb{Z}$  et  $i = 1, \dots, n$  et pour tout  $h \in \mathbb{Z}$  avec  $t_{i+h} \in \mathbb{Z}$  on a la loi de probabilité de la suite  $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})$  est identique à celle de la suite  $(X_{t_1}, \dots, X_{t_n})$ .

#### Stationnarité au second ordre

Un processus aléatoire  $\{X_t; t \in \mathbb{Z}\}$  est dit stationnaire au second ordre ou faiblement stationnaire si :

1.  $E(X_t^2) < \infty, \forall t \in \mathbb{Z}$

2.  $E(X_t) = m$ , indépendant de  $t \forall t \in \mathbb{Z}$
3.  $cov(X_t, X_{t+h}) = E[(X_{t+h} - m)(X_t - m)] = \gamma(h)$  indépendante de  $t \forall (t, h) \in \mathbb{Z}^2$

### 1.2.6 Processus Bruit Blanc

On dit que le processus  $\{\varepsilon_t; t \in \mathbb{Z}\}$  est un bruit blanc faible ( processus hasard pure ) s'il vérifie :

1.  $E(\varepsilon_t) = 0, \forall t \in \mathbb{Z}$
2.  $E((\varepsilon_t)^2) = \sigma^2, \forall t \in \mathbb{Z}$
3.  $cov(\varepsilon_t, \varepsilon_s) = E(\varepsilon_t \varepsilon_s) = 0, \forall t \neq s \in \mathbb{Z}$ .

Si de plus on a les variables  $\varepsilon_t$  et  $\varepsilon_s$  indépendantes on parle donc d'un bruit blanc fort.

### 1.2.7 Caractéristiques d'une série chronologique

#### La moyenne

la moyenne d'un processus est estimée par la moyenne empirique définie par :

$$\bar{X} = \frac{1}{T} \sum_{i=1}^T X_i \quad (1.1)$$

#### Propriétés :

1.  $\bar{X}$  est un estimateur sans biais de  $\mu$
2.  $E(\bar{X}) = \mu$
3.  $var(\bar{X}) = \frac{1}{T} \sum_{|h| < T} (1 - \frac{|h|}{T}) \gamma(h)$

#### La variance

La variance d'une série chronologique permet d'évaluer la dispersion autour de la moyenne. Dans le cas d'une population totalement connue la variance est donnée par :

$$Var(X) = \sum_{t=1}^T \frac{(X_t - \bar{X})^2}{T} \quad (1.2)$$

La racine carrée de la variance est appelée l'écart type noté  $\sigma(x)$  défini par :

$$\sigma(x) = \sqrt{var(x_t)} \quad (1.3)$$

## La fonction d'autocovariance

La fonction d'autocovariance mesure la covariance entre les termes d'une série chronologique. De façon général la covariance entre deux variables  $X_1$  et  $X_2$  est donnée par :

$$\text{cov}(X_1, X_2) = E[(X_1 - \mu_1)(X_2 - \mu_2)] \quad (1.4)$$

où :  $\mu_1 = E(X_1)$  et  $\mu_2 = E(X_2)$

Spécifiquement, on appelle fonction d'autocovariance d'un processus stationnaire  $\{X_t; t \in \mathbb{Z}\}$  la fonction  $\gamma$  définie de  $\mathbb{Z}$  dans  $\mathbb{R}$  par :

$$\begin{aligned} \forall h, t \in \mathbb{Z} \\ \gamma(h) = \text{cov}(X_t, X_{t+h}) \end{aligned} \quad (1.5)$$

On estime la fonction d'autocovariance par :

$$\hat{\gamma}(h) = \frac{1}{T} \sum_{t=1}^{T-|h|} (X_t - \bar{X})(X_{t+|h|} - \bar{X}) \quad (1.6)$$

appelée la fonction d'autocovariance empirique, son graphe est appelé **variogramme**.

### Propriétés :

la fonction d'autocovariance d'un processus stationnaire vérifie :

1.  $\forall h \in \mathbb{Z}, \gamma(-h) = \gamma(h)$  elle est paire
2.  $\forall n \in \mathbb{N}, \forall (a_i) \in \mathbb{R}^{\mathbb{N}}, \sum_{i=1}^n \sum_{j=1}^n a_i a_j \gamma(t_i - t_j) > 0$  elle est positive
3.  $\gamma(0) = V(X_t)$
4.  $|\gamma(h)| \leq \gamma(0); \forall h$

**Remarque 1.1.** *La fonction d'autocovariance peut être normalisée et la nouvelle fonction obtenue est appelée fonction d'autocorrélation.*

## La fonction d'autocorrélation

La fonction d'autocorrélation mesure la corrélation entre les termes d'une série chronologique. De façon général la corrélation est une version standardisée de la covariance , spécifiquement la corrélation entre deux variables  $X_1$  et  $X_2$  est donnée par :

$$\text{cor}(X_1, X_2) = \frac{\text{cov}(X_1, X_2)}{\sigma_{X_1} \sigma_{X_2}} \quad (1.7)$$

où :  $\sigma_{X_1}$  et  $\sigma_{X_2}$  sont les écarts types de  $X_1$  et  $X_2$ .

On appelle fonction d'autocorrélation d'un processus stationnaire  $\{X_t; t \in \mathbb{Z}\}$ , la fonction  $\rho$  définie de  $\mathbb{Z}$  dans  $\mathbb{R}$  par :

$$\forall h \in \mathbb{Z}$$

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} \quad (1.8)$$

On peut estimer la fonction d'autocorrélation à partir de l'autocovariance empirique définie par :

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{T-|h|} (X_t - \bar{X})(X_{t+|h|} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2} \quad (1.9)$$

appelée l'autocorrélation empirique, son graphe est appelé **corrélogramme**.

### Propriétés :

La fonction d'autocorrélation d'un processus stationnaire vérifie :

1.  $\forall h \in \mathbb{Z}, \rho(-h) = \rho(h)$  elle est paire
2.  $\rho(0) = 1$
3.  $|\rho(h)| \leq 1, \forall h$

**Remarque 1.2.** la fonction  $\rho(h)$  est l'expression du lien linéaire entre  $X_t$  et  $X_{t-h}$  qui veut dire le lien entre le présent et le passé d'ordre  $h$ .

On remarque que les valeurs  $\rho(h), h=1,2,\dots$  ne sont souvent pas suffisantes pour expliquer le lien entre le présent et le passé car les variables du passé sont elles même reliées entre elles. Autrement dit,  $X_{t-h}$  dépend en général des variables  $X_{t-h+1}, \dots, X_{t-1}$  donc on fait recours à une nouvelle fonction appelée la fonction d'autocorrélation partielle.

### La fonction d'autocorrélation partielle

L'autocorrélation partielle d'ordre  $h$  désigne la corrélation entre  $X_t$  et  $X_{t-h}$  obtenue lorsque l'influence des variables  $X_{t-h-i}$  avec  $i < h$ , a été retirée.

L'autocorrélation partielle d'ordre  $h$  d'un processus  $\{X_t; t \in \mathbb{Z}\}$  de moyenne  $E(X_t) = m$ , notée  $\rho_h$ , est définie par le dernier coefficient de la projection linéaire de  $X_{t+1}$  sur ces  $h$  plus récentes valeurs .  $\forall h \in \mathbb{Z}$  :

$$X_{t+1} - m = c_1(X_t - m) + c_2(X_{t-1} - m) + \dots + c_{h-1}(X_{t-h+1} - m) + \rho_h(X_{t-h+1} - m) \quad (1.10)$$

En général, la fonction d'autocorrélation partielle d'un processus stationnaire  $\{X_t; t \in \mathbb{Z}\}$  satisfait la relation

$$\rho_h = \frac{|\rho_h^*|}{|\rho_h|} \quad (1.11)$$

$$\forall h \in \mathbb{N}$$

avec

$$P_h = \begin{pmatrix} 1 & \rho_1 & \dots & \dots & \rho_{h-1} \\ \cdot & 1 & & & \\ \cdot & & & & \\ \rho_{h-1} & & & & 1 \end{pmatrix}$$

et

$$P_h^* = \begin{pmatrix} 1 & \rho_1 & \dots & \dots & \rho_1 \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ \rho_{h-1} & & & & \rho_h \end{pmatrix}$$

où :  $|\rho_h^*|$  est le déterminant de la matrice  $P_h$ .

Pour calculer les  $\rho_{hh}$  on utilise l'écriture récurrente de  $\rho_{ii}$  :

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

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

**Remarque 1.3.** Pour obtenir les estimateurs  $\hat{\rho}_h$  on utilise les estimateurs des autocorrélations de la façon suivante :

$$\hat{\rho}_h = \frac{|\hat{\rho}_h^*|}{|\hat{\rho}_h|} \quad (1.12)$$

$$\forall h \in \mathbb{N}$$

avec

$$P_h = \begin{pmatrix} 1 & \hat{\rho}_1 & \dots & \dots & \hat{\rho}_{h-1} \\ \cdot & 1 & & & \\ \cdot & & & & \\ \hat{\rho}_{h-1} & & & & 1 \end{pmatrix}$$

et

$$P_h^* = \begin{pmatrix} 1 & \hat{\rho}_1 & \dots & \dots & \hat{\rho}_1 \\ \cdot & & & & \\ \cdot & & & & \\ \hat{\rho}_{h-1} & & & & \hat{\rho}_h \end{pmatrix}$$

### 1.2.8 Les opérateurs $B$ et $\Delta$

#### L'opérateur de retard $B$

On appelle opérateur  $B$  ( back ward ) noté aussi pour L log) qui à tout processus  $\{X_t; t \in \mathbb{Z}\}$  associe le processus  $\{Y_t; t \in \mathbb{Z}\}$  défini par :

$$\begin{aligned} \forall t \in \mathbb{Z} \\ Y_t = BX_t = X_{t-1} \end{aligned} \tag{1.13}$$

Il décale le processus d'une unité de temps vers le passé.

**Remarque 1.4.** L'opérateur  $B$  est linéaire et inversible. Son inverse  $B^{-1} = F$  est défini par :

$$\begin{aligned} \forall t \in \mathbb{Z} \\ FX_t = X_{t+1} \end{aligned} \tag{1.14}$$

telque  $F$  est appelé opérateur avance.

- Si on compose  $B$  avec lui même on obtient  $B^2 = B \circ B$  telque :

$$\begin{aligned} \forall t \in \mathbb{Z} \\ B^2 X_t = X_{t-2} \end{aligned} \tag{1.15}$$

on peut itérer cette application et définir par récurrence :

$$B^k X_t = X_{t-k}, \forall k \in \mathbb{N}$$

- Par convention,  $B^0$  est l'opérateur identité  $I$ .
- $Ba = a$ , l'opérateur d'une constante est une constante.

### L'opérateur de Différenciation $\Delta$

On appelle l'opérateur de différenciation l'opérateur **I-B** noté  $\Delta$ .  
Si on a une série chronologique ayant une tendance polynômiale on peut annuler celle-ci par différenciation successive du processus telque le nombre de différenciation est égal au degré du polynôme de tendance.

**Remarque 1.5.** L'opérateur  $\Delta$  permet d'éliminer la tendance de la série. Il peut être répété plusieurs fois si la tendance n'est pas linéaire, par exemple :

$$\Delta^2 X_t = (1 - B)^2 X_t = (1 - 2B + B^2) X_t \quad (1.16)$$

Le nombre de fois où on applique  $\Delta$  est appelé **ordre de différenciation**.

### L'opérateur de Désaisonnalisation $\Delta_s$

Il est défini par :

$$\Delta_s X_t = X_t - X_{t-s} \quad (1.17)$$

où :

$s$  est la période de saisonnalité.

on peut aussi écrire :

$$\Delta_s = (1 - B^s) \quad (1.18)$$

**Remarque 1.6.** l'opérateur  $\Delta_s$  permet d'éliminer la saisonnalité de période  $s$ .  
On peut l'appliquer plusieurs fois :

$$\Delta_s = \Delta_s(X_t - X_{t-s}) \quad (1.19)$$

Le nombre de fois où on applique  $\Delta_s$  est appelé **ordre de désaisonnalisation**.

## 1.3 Processus aléatoires stationnaires

Dans cette partie, on entame l'étude détaillée de la partie aléatoire d'une série chronologique. Les processus linéaires constituent le modèle le plus simple pour d'écrire cette

composante. On va d'abord présenter comme exemples introductifs les processus autorégressifs et les processus moyennes mobiles qui forment une classe flexible de modèles pour de nombreux phénomènes observés et construits à partir de l'idée que l'observation au temps  $t$  s'explique linéairement par les observations précédentes. Puis on va définir les processus ARMA qui sont la combinaison des processus autorégressifs et moyennes mobiles. Cette classe de processus ARMA est encore un cas particulier de processus linéaires et jouera un rôle dans la modélisation concrète des processus stationnaires. Elle présente l'avantage d'être plus souple à l'utilisation et de fournir généralement de bonnes approximations des séries réelles avec moins de paramètres que les modèles purs. En particulier, on exposera brièvement la méthode de Box et Jenkins qui est une des méthodes de prévision la plus couramment utilisée (en particulier sous R) en raison de sa simplicité, de l'économie du temps qu'elle permet et de la fiabilité de ses résultats.

### 1.3.1 Processus MA (q)

Un processus  $\{X_t; t \in \mathbb{Z}\}$  est dit à moyenne ajustée ( Moving average ) pour un entier  $q \geq 0$  noté MA(q) lorsqu'il est du second ordre stationnaire et s'il est solution de l'équation de récurrence :

$$X_t = Z_t - \theta_1 Z_{t-1} \dots - \theta_q Z_{t-q} \quad (1.20)$$

$$\forall t \in \mathbb{Z}$$

où

$\theta_1, \theta_2, \dots, \theta_q$  des réels fixes.

$\{Z_t; t \in \mathbb{Z}\}$  est un  $BB(0, \sigma^2)$ .

### 1.3.2 Les caractéristiques d'un MA(q)

#### La fonction d'auto-covariance

Soit un processus MA(q) centré

$$X_t = \theta_q(B)\varepsilon_t \quad (1.21)$$

ou encore

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (1.22)$$

on a

$$\gamma_h = E[X_t X_{t-h}] \quad (1.23)$$

Or

$$\begin{cases} E[\varepsilon_t \varepsilon_{t-h}] = 0 & h > q \\ E[\varepsilon_t \varepsilon_{t-h}] = \sigma_\varepsilon^2 & h \leq q \end{cases}$$

On obtient

$$\gamma(h) = \begin{cases} \sigma_\varepsilon^2(1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) & h = 0 \\ \sigma_\varepsilon^2(-\theta_h + \theta_1\theta_{h+1} + \dots + \theta_q\theta_{q+h}) & 1 \leq h \leq q \\ 0 & h > q \end{cases}$$

### La fonction d'autocorrélation

On sait que

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} \quad (1.24)$$

Donc on déduit que

$$\rho(h) = \begin{cases} 1 & \\ \frac{-\theta_h + \theta_1\theta_{h+1} + \dots + \theta_q\theta_{q+h}}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & 0 < h \leq q \\ 0 & \end{cases}$$

**Remarque 1.7.** La fonction d'autocorrélation d'un MA(q) s'annule à l'ordre  $q+1$ .

### La fonction d'autocorrélation partielle

Elle est définie par

$$\rho(h) = \begin{cases} \frac{-\theta_1}{1 + \theta_1^2} & h = 1 \\ 0 & h \geq 2 \end{cases}$$

### 1.3.3 Prévision à l'aide d'un MA(q)

On supposera que  $X_t$  est un processus stationnaire suivant un MA(q), donc il vérifie l'équation suivante

$$X_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \dots + \theta_q\varepsilon_{t-q} = \Theta(B)\varepsilon_t \quad (1.25)$$

La prévision à la date T+1 est donnée par

$$\hat{X}_{T+1} = E(X_{T+1}/X_T, X_{T-1}, \dots, X_T) \quad (1.26)$$

alors

$$\hat{X}_{T+1} = 0 + \theta_1\varepsilon_T + \dots + \theta_q\varepsilon_{T-q+1} \quad (1.27)$$

Donc

$$\hat{X}_{T+h} = \begin{cases} \theta_h \hat{X}_{T+1} + \dots + \theta_q X_{T+h-q} & h \leq q \\ 0 & h > q \end{cases}$$

### 1.3.4 Processus AR (p)

Un processus  $\{X_t; t \in \mathbb{Z}\}$  est dit autoregressif pour un entier  $p \geq 0$  noté AR(p) lorsqu'il est du second ordre stationnaire et s'il est solution de l'équation de récurrence :

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + Z_t \quad (1.28)$$

$$\forall t \in \mathbb{Z}$$

où

$\phi_1, \phi_2, \dots, \phi_p$  des réels fixes.

$\{Z_t; t \in \mathbb{Z}\}$  est un  $BB(0, \sigma^2)$  donc on dit que p est l'ordre du processus  $\{X_t; t \in \mathbb{Z}\}$  et  $(\phi, \sigma^2)$  sont ses paramètres.

### 1.3.5 Existence des processus AR(1)

si  $|\phi_1| = 1$  alors l'équation des AR(1) n'a pas de solution stationnaire.

si  $|\phi_1| < 1$  alors la solution est donnée par le processus linéaire causal :

$$X_t = \sum_{k=0}^{\infty} \phi_1^k Z_{t-k} \quad (1.29)$$

de moyenne nulle et d'autocovariance

$$\gamma_X(h) = \frac{\sigma^2 \phi_1^{|h|}}{1 - \phi_1^2} \quad (1.30)$$

pour tout  $h \in \mathbb{Z}$  c'est un MA ( $\infty$ ) et c'est aussi l'unique solution stationnaire.

Si  $|\phi_1| > 1$  alors la solution est donnée par le processus linéaire non-causal :

$$X_t = - \sum_{k=1}^{\infty} \phi_1^{-k} Z_{t+k} \quad (1.31)$$

de moyenne nulle et d'autocovariance

$$\gamma_X(h) = \frac{\sigma^2 \phi_1^{-|h|}}{\phi_1^2 - 1} \quad (1.32)$$

pour tout  $h \in \mathbb{Z}$  c'est aussi l'unique solution stationnaire.

### 1.3.6 Les caractéristiques d'un AR(p)

#### Fonction d'auto-covariance

Soit un processus AR(p) centré

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad (1.33)$$

On considère

$$\Phi(B)X_t = \varepsilon_t \quad (1.34)$$

avec  $E(X_t) = 0$

L'auto-covariance est définie par

$$\gamma(h) = \text{cov}(X_t, X_{t-h}) = E(X_t X_{t-h}) \quad (1.35)$$

pour  $h \geq 0$  On a

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad (1.36)$$

Donc

$$X_t^2 = \phi_1 X_t X_{t-1} + \dots + \phi_p X_t X_{t-p} + X_t \varepsilon_t \quad (1.37)$$

Alors

$$\gamma(0) = \phi_1 \rho(1) + \dots + \phi_p \rho(p) + E(X_t \varepsilon_t) \quad (1.38)$$

Or

$$E(X_t \varepsilon_t) = E[(\phi_1 X_{t-1} + \dots + \phi_p X_{t-p}) \varepsilon_t] + E(\varepsilon_t^2) \quad (1.39)$$

D'où

$$\gamma(0) = \phi_1 \gamma(1) + \dots + \phi_p \gamma(p) + \sigma_\varepsilon^2 \quad (1.40)$$

Si  $h > 0$ , on procède de la même manière et donc on aura

$$\gamma(h) = \phi_1 \gamma_{h-1} + \dots + \phi_p \gamma_{h-p} + E(\varepsilon_t X_{t-h}) \quad (1.41)$$

avec  $E(\varepsilon_t X_{t-h}) = 0$  puisque  $X_{t-h}$  ne dépend des  $\varepsilon_j$  que pour  $j < t - h$

#### Fonction d'autocorrélation

On a déjà vu que la fonction d'autocorrélation d'un processus est définie par

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} \quad (1.42)$$

Donc on peut déduire que

$$\rho(h) = \phi_1 \rho_{h-1} + \dots + \phi_p \rho_{h-p} \quad (1.43)$$

$\forall h \geq 0$

$$\rho_h - \phi_1 \rho_{h-1} + \dots + \phi_p \rho_{h-p} = 0 \rho_h (1 - \phi_1 B - \dots - \phi_p B^p) = 0 \phi_p(B) \rho_h = 0 \quad (1.44)$$

C'est une équation aux différences homogènes et elle a pour solutions générales :

$$\rho_h = A_1 \lambda_1^h + \dots + A_p \lambda_p^h \quad (1.45)$$

où  $\frac{1}{\lambda_1} \dots \frac{1}{\lambda_p}$  sont les racines du polynôme  $\phi_p(B) = 0$  et  $A_1 \dots A_p$  sont des constantes déterminées à partir de conditions initiales.

**Remarque 1.8.** *La fonction d'autocorrélation d'un AR(p) s'annule à l'ordre  $p+1$ .*

### Fonction d'autocorrélation partielle

Les autocorrélations partielles notées  $\rho_h$  d'un processus AR(p) sont nulles pour tout ordres supérieur à p et non nulles pour tout ordres inférieur à p. Donc on peut écrire :

$$\Gamma = \begin{cases} 1 & si & h = 0 \\ \rho_1 & si & h = 1 \\ \phi_p & si & h = p \\ 0 & si & h > p \end{cases}$$

### 1.3.7 Estimation dans le modèle autorégressif AR(p)

Après avoir identifier l'ordre du processus adéquat, il convient d'estimer les paramètres du modèle. Cette étape est essentielle dans l'étude d'une série chronologique, elle nous permet donc de valider le modèle choisi à l'aide d'un nombre de tests statistiques. Il existe plusieurs méthodes d'estimation des paramètres, parmi les plus connues on a :

• **La méthode des Moindres carrés** : d'une façon générale, cette méthode consiste à estimer un paramètre  $\theta$  en minimisant la somme des carrés des résidus, telque :

$$\varphi(\hat{\theta}(y)) = \min \sum_{i=1} (\hat{e}_i)^2$$

Cette méthode d'estimation ne nécessite pas que l'on pose l'hypothèse de normalité des résidus.

• **La méthode du Maximum de vraisemblance** : d'une façon générale, cette méthode est basée sur la vraisemblance du modèle linéaire gaussien :

$$\mathbf{L}(\theta, y) = \prod_{i=1}^n f(y_i, \theta)$$

Pour obtenir l'estimateur  $\hat{\theta}$  du maximum de vraisemblance, on maximise sa log-vraisemblance selon  $\theta$  en résolvant le système d'équations du maximum de vraisemblance :

$$\frac{\partial}{\partial \theta_j} \ln \mathbf{L}(\theta_1, \dots, \theta_k, y) = 0 \quad \text{pour } j = 1, \dots, k$$

Pour bien comprendre les deux méthodes, on va les appliquer dans le cas d'un modèle AR(p) et de la même façon concernant l'estimation des autres modèles dans ce qui suit.

### Méthode des moindres carrés

Un AR(p) peut s'écrire comme

$$X_t = \mu + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t = Z_t' \beta_t + \varepsilon_t \quad (1.46)$$

où  $Z_t' = (1, \phi_1, \phi_2, \dots, \phi_p)$  et  $\varepsilon_t$  est un bruit blanc. On estime le modèle  $X_t = Z_t' \beta_t + \varepsilon_t$  par la méthode MCO et on obtient donc

$$\hat{\beta} = (Z_t Z_t')^{-1} Z_t X_t \quad \text{et } \hat{\sigma}^2 = \frac{1}{T-p} \sum_{t=p+1}^T (X_t - Z_t' \hat{\beta})^2 \quad (1.47)$$

### Méthode du maximum de vraisemblance

Dans ce cas, pour déterminer la vraisemblance on supposera que la loi des erreurs est connue donc elles sont normalement distribuées et indépendantes et  $\varepsilon_t$  est un bruit blanc. La fonction de vraisemblance du processus est donnée par

$$\mathbf{L}(X, \phi, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{T/2}} \frac{1}{[\det \Omega]^{1/2}} \exp\left(-\frac{1}{2\sigma^2} X' \Omega^{-1} X\right) \quad (1.48)$$

La log-vraisemblance est donc donnée par

$$\ln \mathbf{L}(X, \phi, \sigma^2) = -\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln \sigma^2 - \frac{1}{2} \ln |\det \Omega| - \frac{1}{2\sigma^2} X' \Omega^{-1} X \quad (1.49)$$

avec  $X = (X_1, \dots, X_t)$

il est difficile de calculer le déterminant de la matrice et son inverse donc même la vraisemblance, pour cela on utilise l'algorithme de maximisation itératif.

### Exemple :

Dans le cas d'un modèle AR(1), de la forme

$$X_t = c + \phi X_{t-1} + \varepsilon_t \quad (1.50)$$

où  $\varepsilon_t \sim N(0, \sigma^2)$  et  $|\phi| < 1$  alors

$$X_t | X_{t-1} \sim N(c + \phi X_{t-1}, \sigma^2)$$

La loi conditionnelle de  $X_t$  est donnée par

$$f(x_t|x_{t-1}, (c, \phi, \sigma^2)) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x_t - c - \phi x_{t-1})^2\right] \quad (1.51)$$

C'est à dire que  $X_t \sim N(E(X_t), V(X_t))$  donc  $X_t \sim N(\frac{c}{1-\phi}, \frac{\sigma^2}{1-\phi^2})$

En posant  $\beta = (c, \sigma^2)$ , la vraisemblance conditionnelle du modèle est donc donnée par

$$\mathbf{L}(\beta|X_1, \dots, X_T) = \prod_{t=2}^T \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(X_t - c - \phi X_{t-1})^2\right] \quad (1.52)$$

et

$$\log \mathbf{L}(\beta|X_1, \dots, X_T) = -\frac{T-1}{2} \ln(2\pi) - \frac{T-1}{2} \ln \sigma^2 + \frac{1}{2\sigma^2} \sum_{t=2}^T (X_t - c - \phi X_{t-1})^2 \quad (1.53)$$

La log-vraisemblance est donnée par :

On peut dire que le problème de maximisation de la vraisemblance est un problème d'optimisation non-linéaire.

### 1.3.8 Prédiction à l'aide d'un AR(p)

On supposera que  $X_t$  est un processus stationnaire suivant un AR(p), donc il vérifie l'équation suivante

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad (1.54)$$

avec  $\varepsilon_t$  est un bruit blanc.

La prédiction à la date T+1 est donnée par

$$\hat{X}_{T+1} = E(X_{T+1}/X_T, X_{T-1}, \dots, X_T) \quad (1.55)$$

alors

$$\hat{X}_{T+1} = \phi_1 X_T + \dots + \phi_p X_{T-p+1} \quad (1.56)$$

De manière générale

$$\hat{X}_{T+h} = \phi_1 \hat{X}_{T+h-1} + \dots + \phi_p X_{T+h-p} \quad (1.57)$$

### 1.3.9 Processus ARMA (p,q)

Les processus ARMA sont représentatifs d'un processus généré par une combinaison des valeurs passées et des erreurs passées.

Un processus  $X_t$  est dit processus ARMA( p,q )(auto regressive moving average)s'il vérifie l'équation suivante :  $\forall t \in \mathbb{Z}$

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t - \theta_1 Z_{t-1} - \dots - \theta_q Z_{t-q} \quad (1.58)$$

avec : p et q des entiers,  $\phi_1, \dots, \phi_p$  et  $\theta_1, \dots, \theta_q$  des réels fixés et  $\{Z_t; t \in \mathbb{Z}\}$  est un bruit blanc de variance  $\sigma^2$ .

Posons :

$$\Phi(B) = I - \phi_1 B - \dots - \phi_p B^p \quad \text{et} \quad \Theta(B) = I - \theta_1 B - \dots - \theta_q B^q \quad (1.59)$$

On peut donc écrire

$$\forall t \in \mathbb{Z}$$

$$\Phi(B)X_t = \Theta(B)Z_t \quad (1.60)$$

**Remarque 1.9.** Notons que les processus AR(p) et MA(q) sont des cas particuliers du processus ARMA(p,q). En effet, ARMA(p,0)=AR(p) et ARMA(0,q)=MA(q).

### **Théorème de Brockwell and Davis (1991)**

- **ARMA (p,q) causal**

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q} \quad (1.61)$$

où  $(Z_t) \sim \mathcal{BB}(0, \sigma^2)$  et les  $\phi_j$  et les  $\theta_j$  sont des nombres réels. On pose

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \quad (1.62)$$

et

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q \quad (1.63)$$

On suppose que  $\phi(z)$  et  $\theta(z)$  n'ont pas de zéros communs. Alors l'équation admet une solution stationnaire causale au second ordre si et seulement si le polynôme  $\phi(z) \neq 0$  pour  $|z| \leq 1$ . Cette solution est unique et a pour expression :

$$X_t = \sum_{k \geq 0} \psi_k Z_{t-k} \quad (1.64)$$

- **ARMA(p,q) inversible**

Soit l'équation récurrente :

$$X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q} \quad (1.65)$$

où  $Z_t \sim \mathcal{BB}(0, \sigma^2)$  et les  $\phi_j$  et les  $\theta_j$  sont des nombres réels. On pose

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p \quad (1.66)$$

et

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_p z^p \quad (1.67)$$

On suppose que  $\phi(z)$  et  $\theta(z)$  n'ont pas de zéros communs. Alors l'équation admet une solution stationnaire inversible au second ordre si et seulement si le polynôme  $\theta(z) \neq 0$  pour  $|z| \leq 1$ . Cette solution est unique et a pour expression :

$$Z_t = \sum_{k \geq 0} \pi_k X_{t-k} \quad (1.68)$$

### 1.3.10 Caractéristiques du modèle ARMA (p,q)

#### Fonction d'auto-covariance

Soit un processus ARMA(p,q) centré

$$\phi_p(B)X_t = \theta_q(B)\varepsilon_t \quad (1.69)$$

ou encore

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (1.70)$$

On a pour  $h > 0$

$$\gamma_h = E[X_t X_{t-h}] = \phi_1 \gamma_{h-1} + \dots + \phi_p \gamma_{h-p} + E[\varepsilon_t X_{t-h}] - \theta_1 E[\varepsilon_{t-1} X_{t-h}] - \dots - \theta_q E[\varepsilon_{t-q} X_{t-h}] \quad (1.71)$$

par définition,  $X_{t-h}$  ne dépend que des  $\varepsilon_{t-h}, \dots, \varepsilon_{t-h-q}$

On peut écrire :

$$\begin{cases} si & h > 0 & E[\varepsilon_t X_{t-h}] = 0 \\ si & h \leq 0 & E[\varepsilon_t X_{t-h}] \neq 0 \end{cases}$$

Pour  $h \geq q + 1$ , avec  $E[\varepsilon_t X_{t-h-q}] = 0$

on aura

$$\gamma_h = \phi_1 \gamma_{h-1} + \dots + \phi_q \gamma_{h-q} \quad (1.72)$$

#### Fonction d'autocorrélation

Par définition on a

$$\rho_h = \frac{\gamma(h)}{\gamma(0)} \text{ donc } \rho_h = \phi_1 \rho_{h-1} + \dots + \phi_p \rho_{h-p} \quad (1.73)$$

#### Fonction d'autocorrélation partielle

Concernant ça, la partie AR ne se fait sentir que jusqu'à l'ordre p.

Pour  $h \leq p$ , il n'y a pas de schéma bien précis.

Pour  $h > p$ , on retrouve le même comportement que pour un MA(q).

### 1.3.11 Estimation dans modèle ARMA(p,q)

Soit un processus ARMA(p,q) centré

$$\phi_p(B)X_t = \theta_q(B)\varepsilon_t \quad (1.74)$$

alors

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (1.75)$$

On supposera que les  $\varepsilon_t$  obéissent à un bruit blanc pour pouvoir déterminer la vraisemblance d'un processus ARMA et écrire :

$$\frac{1}{(2\pi\sigma_\varepsilon^2)^{\frac{n}{2}}} \frac{1}{(\det\Omega)^{\frac{1}{2}}} \exp\left[-\frac{1}{2\sigma_\varepsilon^2} X\Omega^{-1}X\right] \quad (1.76)$$

avec  $\sigma_\varepsilon^2\Omega$  est la matrice des variances covariances du vecteur aléatoire  $X = (X_1, \dots, X_n)$ . Elle dépend des  $\phi_i$  et des  $\theta_j$  du processus ARMA. Pour n grand, il est difficile de calculer le déterminant de la matrice et son inverse aussi donc même la vraisemblance, pour cela on utilise d'autres expressions et en particulier celle de Newbold.

$$\frac{1}{(2\pi\sigma_\varepsilon^2)^{\frac{n}{2}}} \frac{1}{(\det\dot{Z}Z)^{\frac{1}{2}}} \exp\left[-\frac{s}{2\sigma_\varepsilon^2}\right] \quad (1.77)$$

avec Z matrice(p+q+n,p+q) qui dépend des  $\phi_i$  et des  $\theta_j$  du processus ARMA et

$$s = \sum_{t=1}^n [\varepsilon_t^2]$$

$$t = \begin{cases} -\infty \\ 1 \end{cases}$$

La vraisemblance pour la chronique échantillon en logarithme est donnée par :

$$\mathbf{L} = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma_\varepsilon^2 - \frac{1}{2} \log[\det[\dot{Z}Z]] + \log J - \frac{s}{2\sigma_\varepsilon^2} \quad (1.78)$$

où  $J = \prod_{t=1}^n X_t^{\lambda-1}$  est le jacobien de la transformation Box-Cox si celle-ci est utilisée.

Il faut maximiser  $\mathbf{L}$  connaissant  $\lambda, \phi_i, \theta_j, \sigma_\varepsilon^2$  c'est à dire  $\mathbf{Max} \mathbf{L} \lambda, \phi_i, \theta_j, \sigma_\varepsilon^2$

On estime  $\sigma_\varepsilon^2$  par

$$\frac{\partial \mathbf{L}}{\partial \sigma_\varepsilon^2} = -\frac{n}{2\sigma_\varepsilon^2} + \frac{s}{2\sigma_\varepsilon^2} = 0 \quad \Rightarrow \quad \hat{\sigma}_\varepsilon^2 = \frac{s}{n} \quad (1.79)$$

On remplace  $\sigma_\varepsilon^2$  par  $\hat{\sigma}_\varepsilon^2$  et on remarque que maximiser  $\mathbf{L}$  revient à minimiser :

$$\left[ \log \frac{s}{J^{\frac{2}{n}}} + \frac{1}{n} \log[\det(\dot{Z}Z)] \right] \quad (1.80)$$

### 1.3.12 Prévision à l'aide d'un modèle ARMA(p,q)

On supposera que l'on s'est ramené à un processus centré  $\{X_t; t \in \mathbb{Z}\}$  satisfaisant

$$\Phi(B)X_t = \Theta(B)\varepsilon_t \quad (1.81)$$

Sous la forme ARMA, alors

$$X_t = \sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (1.82)$$

donc

$$X_{t+h} = \sum_{j=1}^p \phi_j X_{t+h-j} + \varepsilon_{t+h} + \sum_{i=1}^q \theta_i \varepsilon_{t+h-i} \quad (1.83)$$

On aura donc

$$\hat{X}_{T+h} = E(X_{T+h}/X_T, X_{T-1}, \dots, X_1) \quad (1.84)$$

On peut noter :

$$\hat{X}_{T+h} = \begin{cases} \phi_1 \hat{X}_{T+h-1} + \dots + \phi_h X_T + \dots + \phi_p X_{T+h-p} & h \leq p \\ \phi_1 \hat{X}_{T+h-1} + \dots + \phi_p \hat{X}_{T+h-p} & h > p \end{cases}$$

### 1.3.13 Méthodologie de Box-Jenkins

La méthodologie de Box-Jenkins est une méthode de prévision adaptée au traitement et modélisation des séries chronologiques univariées au moyen des processus ARMA, développée en 1976. Elle est une approche itérative comportant des étapes essentielles afin d'identifier un modèle représentatif du phénomène étudié.

Cette méthode comporte cinq (05) étapes qui sont :

1. **Étape 01** : Transformation des données afin de stabiliser la variance (log, sqrt, ...) et différenciation des données pour les stationariser.
2. **Étape 02** : Visualiser les ACFs et les PACFs empiriques pour identifier les paramètres p et q appropriés.
3. **Étape 03** : Estimation des paramètres du modèle sélectionné.
4. **Étape 04** : Diagnostique et tests d'adéquation du modèle.
5. **Étape 05** : Prévision et c'est la dernière étape, elle consiste à la prévision des valeurs futures à travers du modèle retenu.

## 1.4 Processus aléatoires non-stationnaires

Les chroniques économiques sont rarement des réalisations de processus aléatoires stationnaires déjà étudiés dans la section précédente. Toutefois, la série observée peut montrer un comportement de non stationnarité. Dans ce cas, il faut identifier le type de cette non-stationnarité qui est fréquemment analysée à partir de deux types de processus :

- le processus TS (Trend Stationary)
- le processus DS (Difference Stationary)

### 1.4.1 Processus TS (Trend Stationary)

Il s'agit d'un processus stochastique non stationnaire de nature déterministe. Il s'écrit de la manière suivante :

$$X_t = f_t + \varepsilon_t$$

composé de deux fonctions :

- $f_t$  : une fonction polynômiale du temps.
- $\varepsilon_t$  un bruit blanc stationnaire.

le processus TS le plus simple est représenté par une fonction polynômiale de degré 1. Ce processus s'écrit de la manière suivante :

$$X_t = a_0 + a_1 t + \varepsilon_t$$

si  $\varepsilon_t$  est un bruit blanc, les caractéristiques de ce processus sont alors :

1.  $E(X_t) = a_0 + a_1 t + E(\varepsilon_t) = a_0 + a_1 t$ .
2.  $V(X_t) = 0 + V(\varepsilon_t) = \sigma_\varepsilon^2$
3.  $cov(X_t, X_{t'}) = 0$  pour  $t \neq t'$

**Remarque 1.10.** *Pour stationnariser un processus TS, il suffit d'estimer les paramètres  $a_0$  et  $a_1$  de la tendance en utilisant la méthode des moindres carrés ordinaires pour avoir  $\hat{a}_0$  et  $\hat{a}_1$  ce qui nous permettra de stationnariser le processus en retranchant de la valeur de  $X_t$  en  $t$  la valeur estimée  $\hat{a}_0 + \hat{a}_1 t$ .*

### 1.4.2 Processus DS (Difference Stationary)

Le processus DS peut être stationnarisé en utilisant un filtre aux différences tel que :

$$(1 - B)^d X_t = \beta + \varepsilon_t$$

avec

- $\varepsilon_t$  : un processus stationnaire de type ARMA ou encore un Bruit Blanc.
- $\beta$  : une constante réelle .

- $d$  : ordre du filtre aux différences.

Le processus est dit processus du premier ordre quand il est représenté en utilisant un filtre aux différences premières ( $d=1$ ) et il s'écrit :

$$\begin{cases} (1 - B)X_t = \beta + \varepsilon_t \\ X_t = X_{t-1} + \beta + \varepsilon_t \end{cases}$$

avec :

$\varepsilon_t$  : un processus stationnaire Bruit Blanc (gaussien ou non).

En introduisant la constante  $\beta$  dans le processus DS, on peut définir deux différents processus :

- **cas de  $\beta = 0$  :**

on dit processus DS sans dérive qui s'écrit :

$$X_t = X_{t-1} + \varepsilon_t$$

avec :

1.  $E(X_t) = X_0$ .
2.  $V(X_t) = t\sigma_\varepsilon^2$
3.  $cov(X_t, X_{t'}) = \sigma_\varepsilon^2 \min(t, t')$  si  $t \neq t'$

- **cas de  $\beta \neq 0$  :**

on dit processus DS avec dérive et il s'écrit :

$$X_t = X_{t-1} + \beta + \varepsilon_t$$

avec :

1.  $E(X_t) = X_0 + \beta t$ .
2.  $V(X_t) = t\sigma_\varepsilon^2$
3.  $cov(X_t, X_{t'}) = \sigma_\varepsilon^2 \min(t, t')$  si  $t \neq t'$

## 1.5 Tests de racines unitaires

La stationnarité est une notion qui joue un rôle cruciale dans l'étude des séries chronologiques, elle nous permet d'adopter le choix adéquat de la modélisation de cette dernière. Parmi les outils de détection de la stationnarité on a les tests. Les plus utilisés sont ceux

de racine unitaire "Unit Root Test". Ces tests permettent non seulement de détecter l'existence d'une non-stationnarité mais aussi de déterminer le type de cette non-stationnarité si il s'agit d'un processus DS ou TS donc c'est la bonne méthode pour stationnariser la série.

### 1.5.1 Test de Dickey-Fuller

On distingue deux (02) tests, le test de Dickey-Fuller simple (1979) et le test de Dickey-Fuller augmentés (1981) qui sont à l'origine de l'étude qui consiste à supposer deux (02) hypothèses :

$H_0$  concerne l'étude d'une racine unitaire contre l'hypothèse  $H_1$  de stationnarité du processus.

#### Test de Dickey-Fuller simple

Les tests de Dickey-Fuller (DF) permettent de détecter la présence de la racine unitaire dans un processus autorégressif du premier ordre AR(1). La construction de ce test est basée sur trois modèles autorégressifs du premier ordre dont les erreurs sont identiquement et indépendamment distribuées (**iid**) : le modèle sans constante, le modèle avec constante et le modèle avec constante et avec tendance. Ces derniers seront donc estimés par la méthode des moindres carrés.

$$(1) \quad X_t = \phi_1 X_{t-1} + \varepsilon_t \quad \text{Modèle autorégressif d'ordre 1.}$$

$$(2) \quad X_t = \phi_1 X_{t-1} + \beta + \varepsilon_t \quad \text{Modèle autorégressif avec constante.}$$

$$(3) \quad X_t = \phi_1 X_{t-1} + bt + c + \varepsilon_t \quad \text{Modèle autorégressif avec tendance.}$$

- Sous l'hypothèse  $H_0$ , les règles d'inférence statistique ne peuvent pas être appliquées pour tester cette hypothèse notamment la distribution de Student du paramètre  $\phi_1$ . Pour cela Dickey et Fuller ont étudié la distribution asymptotique de l'estimateur  $\phi_1$  et ils ont tabulé les valeurs critiques pour des échantillons de tailles différentes.

- Pour les trois modèles, on estime le paramètre  $\phi_1$  noté  $\hat{\phi}_1$  par les moindres carrés. L'estimation par cette méthode fournit  $t_{\hat{\phi}_1}$  qui est analogue à la statistique de Student.

- Si l'hypothèse  $H_0 : \phi_1 = 1$  est vérifiée, la chronique est dite non stationnaire.

- Si l'hypothèse  $H_1 : \phi_1 < 1$  est vérifiée, la chronique est dite stationnaire.

L'application du test de Dickey-Fuller simple se fait comme suite :

**Étape 1** : Dans cette étape on va estimer le modèle (3) telque :

$$(3) \quad X_t = \phi_1 X_{t-1} + bt + c + \varepsilon_t$$

On teste l'hypothèse  $H_0 : b = 0$  contre l'hypothèse  $H_1 : b \neq 0$  avec le test de Student.

✓ Si  $t_{calcul} > t_{tab}$  alors on accepte  $H_1$  et le coefficient  $b$  est significativement différent de 0.

Et on teste aussi l'hypothèse  $H_0 : \phi_1 = 1$  contre l'hypothèse  $H_1 : \phi_1 < 1$ .

✓ Si  $t_{\hat{\phi}_1} > t_{tabul}$  alors on accepte  $H_0$  et on dit que le processus est un TS qu'on peut

stationnariser en calculant les résidus par rapport à la tendance estimée par les moindres carrés (MCO).

✓ Si  $t_{calculé} < t_{tab}$  alors on accepte  $H_1$  et le coefficient  $\beta$  n'est pas significativement différent de 0 alors le processus est engendré par un processus DS donc on passe à l'étape(2).

**Étape 2 :** Dans cette étape on va estimer le modèle (2) telque :

$$(2)X_t = \phi_1 X_{t-1} + \beta + \varepsilon_t$$

On teste l'hypothèse  $H_0 : \beta = 0$  contre l'hypothèse  $H_1 : \beta \neq 0$  avec le test de Student.

✓ Si  $t_{calculé} > t_{tab}$  alors on accepte  $H_1$  et le coefficient  $\beta$  est significativement différent de 0. Donc on passe au test de racine unitaire et on teste l'hypothèse  $H_0 : \phi_1 = 1$  contre l'hypothèse  $H_1 : \phi_1 < 1$ .

✓ Si  $t_{\hat{\phi}_1} > t_{tabule}$  alors on accepte  $H_0$  et on dit que le processus n'est pas stationnaire et on va le stationnariser par la différenciation.

✓ Si  $t_{\hat{\phi}_1} < t_{tabul}$  alors on accepte  $H_1$  et on dit que le processus est stationnaire.

✓ Si  $t_{calculé} < t_{tab}$  alors on accepte  $H_0$  et le coefficient  $\beta$  n'est pas significativement différent de 0 alors le processus est engendré par un processus DS donc on passe à l'étape(3).

**Étape 3 :** Dans cette étape on va estimer le modèle (1) telque :

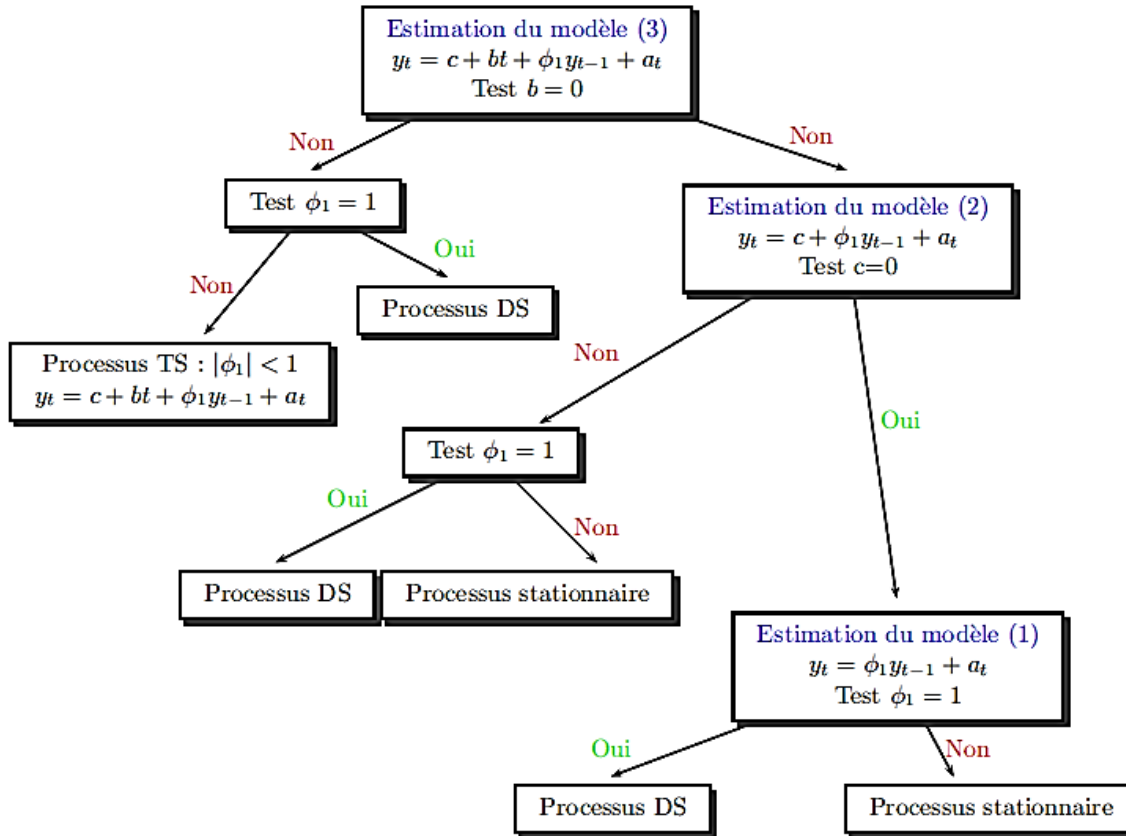
$$(1)X_t = \phi_1 X_{t-1} + \varepsilon_t$$

On teste l'hypothèse  $H_0 : \phi_1 = 1$  contre l'hypothèse  $H_1 : \phi_1 < 1$ .

✓ Si  $t_{\hat{\phi}_1} > t_{tabule}$  alors on accepte  $H_0$  et on dit que le processus n'est pas stationnaire et on va le stationnariser par la différenciation.

✓ Si  $t_{\hat{\phi}_1} < t_{tabul}$  alors on accepte  $H_1$  et on dit que le processus est stationnaire.

Le schéma suivant résume le test de Dickey-Fuller .



**Test de Dickey-Fuller augmenté** Dans le test précédent de Dickey-Fuller, le processus  $\varepsilon_t$  est par hypothèse un bruit blanc. Or il n'y a aucune raison, à priori, l'erreur soit non corrélée. Donc Dickey et Fuller ont conçu un autre test qui est test de Dickey-Fuller augmenté ( ADF ), ce dernier est basé sur l'hypothèse  $H_1 : |\phi_1| < 1$  et l'estimation des modèles par les MCO avec :

$$(4) \Delta X_t = \rho X_{t-1} - \sum_{j=2}^p \phi_j \Delta X_{t-j+1} + \varepsilon_t$$

$$(5) \Delta X_t = \rho X_{t-1} - \sum_{j=2}^p \phi_j \Delta X_{t-j+1} + c + \varepsilon_t$$

$$(6) \Delta X_t = \rho X_{t-1} - \sum_{j=2}^p \phi_j \Delta X_{t-j+1} + bt + c + \varepsilon_t$$

avec  $\varepsilon_t$  *i.i.d*

Ce test déroule de la même manière que le DF, seule différence est celle des tables statistiques.

utilisées. La valeur de  $p$  peut être déterminée en partant d'une valeur suffisamment importante de  $p$ , on estime un modèle à  $p-1$  retard puis à  $p-2$  retard, jusqu'à ce que le coefficient du  $p^{me}$  retard soit significatif.

## 1.5.2 Test de Phillips et Perron (1988)

Ce test est construit sur une correction non-paramétrique des statistiques de Dickey-Fuller pour prendre en compte des erreurs hétéroscédastiques et/ou autocorrélées. Il se déroule en quatre (04) étapes :

✓ Estimer par la méthode des moindres carrés ordinaires (MCO) les trois modèles de base des tests de Dickey-Fuller et calcul des statistiques associées. Soit  $e_t$  le résidu estimé.

✓ Estimer la variance dite du court terme des résidus  $\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n e_t^2$

✓ Estimation du facteur correctif  $S_t^2$  appelé la variance à long terme donné par

$$S_t^2 = \frac{1}{n} \sum_{t=1}^n e_t^2 + 2 \sum_{i=1}^l \left(1 - \frac{i}{l+1}\right) \frac{1}{n} \sum_{t=i+1}^n e_t e_{t-i}$$

avec  $l \approx 4 \left(\frac{n}{100}\right)^{\frac{2}{9}}$

✓ Calcul de la statistique de Phillips et Perron (PP) donnée par

$$t_{\hat{\phi}_1}^* = \sqrt{k} \times \frac{(\hat{\phi}_1 - 1)}{\hat{\sigma}_{\hat{\phi}_1}} + \frac{n(k-1)\hat{\sigma}_{\hat{\phi}_1}}{\sqrt{k}}$$

avec  $k = \frac{\hat{\sigma}^2}{S_t^2}$

Phillips et Perron (1988) montrent que cette correction non-paramétrique apportée à  $t_{\hat{\phi}_1}$  ne modifie pas la distribution asymptotique de la statistique qui reste identique à celle qui est observée dans le cas du test de Dickey-Fuller augmenté. En conséquence, les valeurs critiques tabulées par Dickey et Fuller demeurent également valables pour le test de Phillips-Perron.

# Chapitre 2

## Modélisation VAR et ARDL

### 2.1 Introduction

Nous avons déjà vu dans le chapitre précédent que l'étude des séries chronologiques a pour intérêt l'analyse, l'explication et la prévision d'un phénomène évoluant dans le temps. La prévision est considérée comme un outil d'aide à la prise de décisions car elle permet d'avoir une idée sur le futur en utilisant les séries déjà observées. Cette étude nous permet aussi de définir la relation existante entre les variables constituant la série et aussi d'évaluer l'influence entre ces dernières, pour cela nous allons utiliser une approche économétrique connue par son efficacité et sa contribution dans l'amélioration de la prévision, il s'agit donc de l'approche ARDL développée par Pesaran *et al* (2001), qui sera expliquée dans ce chapitre en donnant en premier lieu un aperçu général sur le côté théorique de cette dernière.

Aussi nous allons voir quelques concepts liés à la modélisation ARDL, il s'agit du concept de cointégration et la modélisation VAR.

### 2.2 Le modèle ARDL

Le modèle ARDL "AutoRegressive Distributed Lag", ou modèle autorégressif à retards échelonnés ou distribués, est un modèle dynamique combinant les caractéristiques de deux autres modèles dynamiques AR et DL.

- **Les modèles autorégressifs (AR)**

C'est des modèles dynamiques où l'on trouve parmi les variables explicatives (exogènes) ( $X_t$ ), la variable dépendante (endogène) décalée.

Sa forme explicite est :

$$Y_t = f(X_t, Y_{t-p})$$

- **Les modèles à retards échelonnés ou distributed lag(DL)**

C'est des modèles dynamiques qui ont pour variables explicatives  $X_t$  et ses valeurs passées

ou décalées.

Sa forme explicite est :

$$Y_t = f(X_t, X_{t-q})$$

Dans le ARDL, on y trouve parmi les variables explicatives ( $X_t$ ), la variable dépendante décalée ( $Y_{t-p}$ ) et les valeurs passées de la variable indépendante ( $X_{t-q}$ ).

Sa forme explicite est :

$$Y_t = f(X_t, Y_{t-p}, X_{t-q})$$

Aussi le modèle ARDL peut s'écrire de la manière suivante :

$$Y_t = \phi + a_1 Y_{t-1} + \dots + a_p Y_{t-p} + b_0 X_t + \dots + b_q X_{t-q} + \varepsilon_t \quad (2.1)$$

ou

$$Y_t = \phi + \sum_{i=1}^p a_i Y_{t-i} + \sum_{j=0}^q b_j X_{t-j} + \varepsilon_t \quad (2.2)$$

avec :

- ✓  $\varepsilon_t \rightsquigarrow iid(0, \sigma)$  : terme d'erreur.
- ✓  $b_0$  traduit l'effet à court terme de  $X_t$  sur  $Y_t$ .
- ✓  $\phi$  traduit l'effet à long terme de  $X_t$  sur  $Y_t$

**Remarque 2.1. Variables endogènes** : ce sont des variables dont la valeur est déterminée par le modèle.

**Variables exogènes** : ce sont les variables du modèle qu'on considère déterminées à l'extérieur du modèle.

## 2.2.1 Caractéristiques du modèle ARDL

- Le modèle ARDL permet d'identifier et d'analyser la relation de long-terme et de court-terme entre les variables explicatives et la variable à expliquer sur des séries qui ne sont pas intégrées de même ordre et d'autre part d'obtenir des meilleures estimations sur des échantillons de petite taille.
- L'avantage du modèle ARDL se manifeste dans sa flexibilité, car ce dernier peut être appliqué même sur les variables étudiées qui ne sont pas intégrées de même ordre, il suffit juste de s'assurer qu'aucune de ces variables n'est intégrée d'ordre deux (2) et plus. Il met en relation la variable à expliquer et les variables explicatives.
- Ce modèle nécessite que les séries chronologiques soient stationnaires au niveau  $I(0)$  ou bien stationnaires après la première différence  $I(1)$ . Pour vérifier la stationnarité on effectue le test ADF déjà vu dans le chapitre précédent.
- Le modèle ARDL est composé de deux parties :
  - ✓ la première partie : une combinaison linéaire de variables en niveau décalées qui représente la dynamique de long terme.

✓ la deuxième partie : une combinaison linéaire de variables différenciées retards qui représente la dynamique de court terme.

• Le fondement théorique du modèle ARDL repose partiellement sur le modèle VAR (Vecteur Autorégressif). Le modèle VAR d'ordre (p) est défini comme suit :

$$Y_t = \phi + \sum_{i=1}^p \beta_i Y_{t-i} + \varepsilon_t \quad (2.3)$$

avec :  $Y_t$  vecteur qui comporte la variable dépendante et la variable endogène à la fois .

## 2.2.2 La méthodologie du modèle ARDL

La procédure ARDL passe par les étapes présentées comme suit :

**Étape 01** : Il s'agit de déterminer d'abord le nombre de retards (retard optimal), de ce fait on se servira des critères d'informations d'Akaike (AIC) et de Schwarz (SIC). La valeur minimale d'un de ces critères offerte par le modèle estimé est appelée "décalage optimal". Leurs valeurs sont calculées comme suit :

$$AIC(p) = \text{Log}|\hat{\Sigma}| + \frac{2}{T}n^2p$$

$$SIC(p) = \text{Log}|\hat{\Sigma}| + \frac{\text{Log}T}{T}n^2p$$

avec :  $|\hat{\Sigma}|$  le déterminant de la matrice de variance-covariance des résidus estimés.

T : nombre d'observations.

p : décalage ou lag du modèle estimé.

n : nombre de régresseurs.

Le SIC permet de sélectionner un nombre plus réduit de retards alors que l'AIC permet de sélectionner le nombre maximum de retards.

**Étape 02** : Il s'agit de vérifier la stationnarité de chaque série car avant d'utiliser l'approche du Bound-Test développé par Pesaran *et al* (2001), il faut d'abord s'assurer préalablement qu'aucune des séries n'est intégrée d'ordre 2 ou plus car les valeurs critiques fournies par Pesaran *et al* (2001) concernent uniquement les niveaux d'intégration 0 et 1. Pour la stationnarité, le test de racine unitaire Dickey-Fuller (ADF) déjà vu dans le chapitre précédent est utilisé. Dans ce cas on va utiliser le test Dickey-Fuller augmenté car dans ce test  $\varepsilon_t$  n'est pas pris pour un bruit blanc car il n'y a aucune raison pour que l'erreur soit non corrélée. La procédure du test est déjà présentée dans le chapitre précédent.

**Étape 03 : Tester la cointégration (Bounds-test)** C'est un test basé sur l'estimation des modèles vectoriels autorégressifs à retards échelonnés. Il est utilisé dans le cas de

présence de au moins de deux (02) séries non stationnaires ce qui nous conduit à rechercher l'existence d'une relation de cointégration (relation d'équilibre à long terme) entre les variables du modèle.

Ce test est applicable lorsque les variables sont de I(0) ou I(1). Cette caractéristique réduit le problème lié à l'incertitude des résultats du test de racine unitaire. L'équation de base du test de cointégration s'écrit comme suit :

$$\Delta Y_t = \gamma_0 + \sum_{i=1}^p \delta_{1i} \Delta Y_{t-1} + \sum_{i=0}^q \pi_{1i} \Delta X_{t-i} + \phi_1 Y_{t-1} + \phi_2 X_{t-1} + \mu_t \quad (2.4)$$

Où :  $\Delta$  l'opérateur de différence première.

p et q les retards déterminés en minimisant le critère d'Akaike (AIC).

Pesaran *et al* (2001) ont simulé deux ensembles de valeurs critiques pour la statistique de test, avec plusieurs cas et différents seuils :

- **Borne inférieure** : C'est l'ensemble de toutes les variables explicatives qui sont I(0).
- **Borne supérieure** : C'est l'ensemble de toutes les variables explicatives qui sont I(1).

Après avoir calculé la **F-stat** (Fisher calculé) on teste comme suit :

- ✓ Si **F-stat** > Borne supérieure : Existence de cointégration.
- ✓ Si **F-stat** < Borne inférieure : Non-existence de cointégration.
- ✓ Si Borne inférieure < **F-stat** < Borne supérieure : On ne peut rien conclure.

#### Étape 04 : Estimer les paramètres du modèle

Il s'agit de l'estimation de long et de court terme.

#### Étape 05 : Validation du modèle

Il s'agit d'effectuer différents tests statistiques de spécification afin de vérifier si le modèle est retenu. Donc on va tester la stabilité du modèle, l'autocorrélation des erreurs et la normalité des résidus.

- **Test de stabilité du modèle** : Il s'agit d'appliquer le test de CUSUM, c'est un test fondé sur la somme des résidus, il nous permet de tester en particulier la stabilité de l'équation estimée.
- **Test d'autocorrélation des erreurs** : Il est très important pour la validité du modèle de vérifier l'hypothèse de bruit blanc. Il risque d'avoir une mauvaise stationnarisation des données ou aussi un mauvais choix de modèle dans le cas où les résidus ne forment pas un bruit blanc, ce qui fait l'importance de ce test.
- **Test de normalité des résidus** : On va effectuer le test de Jarque-Bera en calculant la statistique liée à ce test. On a les coefficients  $\beta_1$  et  $\beta_2$  telque :

$$\beta_1 = \frac{\mu^3}{\sigma^3}$$

et

$$\beta_2 = \frac{\mu^4}{\sigma^4}$$

On propose les estimateurs suivants :

$$b_1 = \frac{\frac{1}{n} \sum_i (x_i - \bar{x})^3}{\frac{1}{n} \sum_i (x_i - \bar{x})^{\frac{3}{2}}} \quad (2.5)$$

$$b_2 = \frac{\frac{1}{n} \sum_i (x_i - \bar{x})^4}{\frac{1}{n} \sum_i (x_i - \bar{x})^2} \quad (2.6)$$

La statistique de Jarque-Bera s'écrit sous la forme suivante :

$$T = n \left( \frac{b_1^2}{6} + \frac{(b_2 - 3)^2}{24} \right) \quad (2.7)$$

Elle est distribuée asymptotiquement selon une loi du  $\chi^2$  à 2 degrés de liberté.

Si La probabilité associée à la statistique de Jarque-Bera  $\mathbf{T} >$  au seuil par exemple on prend  $\alpha = 0.05$  donc la normalité est vérifiée et les résidus d'estimations du modèle sont stationnaires.

## 2.3 Modélisation VAR

Bien que possédant d'excellentes qualités prévisionnelles, le modèle ARMA souffre d'une lacune majeure, il s'agit du fait qu'il est difficile de traiter simultanément plus d'une variable.

Un modèle autoregressif intervient donc afin de permettre de traiter le cas de la présence de plusieurs variables. C'est le modèle Vector AutoRegressive (VAR), dans ces modèles, toutes les variables sont endogènes.

### 2.3.1 Représentation générale d'un modèle VAR

**definition 2.1.** *Un processus vectoriel  $\{X_t; t \in \mathbb{Z}\}$  de dimension  $(n,1)$  admet une représentation VAR d'ordre  $p$ , notée  $VAR(p)$  si :*

$$X_t = \mu - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} + \varepsilon_t \quad (2.8)$$

ou de façon équivalente :

$$\Phi(B)X_t = \mu + \varepsilon_t \quad (2.9)$$

où

- $\mu$  est un vecteur constant de dimension  $(n,1)$
- $\{\varepsilon_t; t \in \mathbb{Z}\}$  est **iid**  $(0_n, \Omega)$  satisfaisant les propriétés suivantes :

$$E(\varepsilon_t) = 0$$

et

$$E(\varepsilon_t \varepsilon_{t-j}') = \begin{cases} \Omega & j = 0 \\ 0 & j \neq 0 \end{cases}$$

où  $\Omega$  : matrice  $(n,n)$  symétrique définie positive.

De la même façon que pour un  $AR(1)$ , on peut exprimer le polynôme matriciel  $\Phi(B)$ , de dimension  $(n;n)$  de la manière suivante :

$$\Phi(B) = \sum_{i=0}^{\infty} \Phi_i B^i = I_n - \phi_1 B^1 - \phi_2 B^2 - \dots - \phi_p B^p \quad (2.10)$$

où :  $\Phi_i$  matrice des de dimension  $(n,n)$  satisfont  $\phi_0 = I_n$  et  $\phi_p = 0_n$   
 $I_n$  est une matrice identité  $(n,n)$ .

On pose donc les définitions suivantes :

$$X_t(n, 1) = \begin{pmatrix} X_{1,t} \\ X_{2,t} \\ \vdots \\ X_{n,t} \end{pmatrix}, \varepsilon_t(n, 1) = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \vdots \\ \varepsilon_{n,t} \end{pmatrix} \text{ et } \phi_i(n, n) = \begin{pmatrix} \phi_{1,1}^i & \phi_{1,2}^i & \dots & \phi_{1,n}^i \\ \phi_{2,1}^i & \phi_{2,2}^i & \dots & \phi_{2,n}^i \\ \vdots & & & \\ \vdots & & & \\ \phi_{n,1}^i & \phi_{n,2}^i & \dots & \phi_{n,n}^i \end{pmatrix}, \forall i \in [1,p]$$

### Processus VAR(1)

On va examiner un processus  $X_t$  qui est un VAR(1) admettant l'écriture suivante :

$$X_t = \mu + \phi_1 X_{t-1} + \varepsilon_t \quad (2.11)$$

Puis on procédera d'une manière récurrente :

- Pour  $t=1$  :

$$X_t = \mu + \phi_1 X_0 + \varepsilon_t$$

- Pour  $t=2$  :

$$X_2 = \mu + \phi_2 X_1 + \varepsilon_2 = \mu + \phi_1(\mu + \phi_1 X_0 + \varepsilon_1) + \varepsilon_2 = (I_n + \phi_1)\mu + \phi_1 \varepsilon_1 + \varepsilon_2 + \phi_1^2 X_0$$

Vers la fin de la procédure le processus  $X_t$  s'écrit sous la forme suivante :

$$X_t = (I_n + \phi_1 + \dots + \phi_1^{t-1})\mu + \phi_1^t X_0 + \sum_{i=0}^{t-1} \phi_1^i \varepsilon_{t-i}$$

**Propriété :** Un processus VAR(1) est stable si les valeurs propres de  $\phi_i$  sont inférieure à 1 en module. C'est à dire on doit avoir

$$\det(I_n - \phi_1 z) \neq 0, \forall |z| \leq 1$$

On distingue donc deux cas :

- Si  $\det(\phi_1) \leq 0$  alors  $\lambda \leq 0$  n'est pas solution de l'équation  $\det(\lambda I_n - \phi_1) = 0$
- Si  $\det(\phi_1) = 0$  alors  $\lambda = 0$  est solution de l'équation  $\det(\lambda I_n - \phi_1) = 0$

D'une autre part les valeurs propres de  $\phi_1$  satisfont l'équation :

$$\det(\lambda I_n - \phi_1) = \lambda^n \det(I_n - \lambda^{-1} \phi_1) = 0$$

et donc on déduit que les solutions satisfont  $|\lambda^{-1}| > 1$  ou encore que les valeurs propres doivent être strictement inférieures à 1 comme on l'a déjà précisé.

### Représentation VAR(1) d'un VAR(p)

On peut aller au processus plus général VAR(p) en appliquant les résultats obtenus du cas particulier VAR(1). Alors, on peut écrire processus VAR(p) comme suite :

$$X_t^* = \mu^* + \phi^* X_{t-1}^* + \varepsilon_t^*$$

où

$$X_t^* = \begin{pmatrix} X_t \\ X_{t-1} \\ \cdot \\ \cdot \\ X_{t-p+1} \end{pmatrix}, \varepsilon_t^* = \begin{pmatrix} \varepsilon_t \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \mu^* = \begin{pmatrix} c \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \phi^* = \begin{pmatrix} \phi_1 & \phi_2 \dots & \phi_{p-1} & \phi_p \\ I_n & 0 \dots & 0 & 0 \\ 0 & I_n \dots & 0 & 0 \\ 0 & \dots & I_n & 0 \end{pmatrix}$$

avec  $X_t^*$ ,  $\varepsilon_t^*$  et  $\mu^*$  sont de dimension  $(np) \times 1$  et  $\phi^*$  est une matrice de dimension  $(np) \times (np)$ .

### 2.3.2 La fonction d'autocovariance d'un VAR(p)

En considérant l'équation suivante

$$X_t^* = \mu^* + \phi^* X_{t-1}^* + \varepsilon_t^*$$

La fonction d'autocovariance d'un VAR(p) peut s'écrire de cette manière

$$\Gamma(h) = E \{ (X_t - \mu)(X_{t-h} - \mu)' \} = E \left\{ \left( \sum_{i=0}^{h-1} \phi_i^* \varepsilon_{t-i} + \sum_{i=0}^{\infty} \phi_{h+i}^* \varepsilon_{t-h-i} \right) \left( \sum_{i=0}^{\infty} \phi_i^* \varepsilon_{t-h-i} \right)' \right\} = \sum_{i=0}^{\infty} \phi_{h+i}^* \Sigma_{\varepsilon} \phi_i^{*'}$$

ou encore sous cette forme

$$\begin{cases} \Gamma(0) = \sum_{i=1}^p \phi_i \Gamma'(i) + \Sigma_\varepsilon & h = 0 \\ \Gamma(h) = \sum_{i=1}^p \phi_i \Gamma(h-i) & h > 0 \end{cases}$$

### 2.3.3 Estimation d'un modèle VAR

Dans le cas d'un processus VAR, l'estimation du modèle peut se faire par la méthode des moindres carrés (MCO) ou celle du maximum de vraisemblance (EMV). On va utiliser donc la méthode du maximum de vraisemblance qu'on a déjà présenté brièvement dans les sections précédentes.

#### Estimation par la méthode du maximum de vraisemblance

On considère le processus VAR(p) suivant :

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t$$

avec :  $\varepsilon_t$  un bruit blanc.

$\Sigma_\varepsilon$  : la matrice des variances covariances.

La vraisemblance conditionnellement à toutes les valeurs passées du processus s'écrit donc comme :

$$\mathbf{L}(X_1, \dots, X_T) = \prod_{t=1}^T \mathbf{L}(X_t | \underline{X}_{t-1})$$

$\underline{X}_{t-1}$  est la mémoire du processus  $X_t$  jusqu'à la date t-1 compris.

Donc la vraisemblance du processus VAR(p) va s'écrire comme :

$$\mathbf{L}(X_1, \dots, X_T) = \prod_{t=1}^T \frac{1}{(\sqrt{2\pi})^N \sqrt{(\det \Sigma_\varepsilon)}} \times \exp\left[-\frac{1}{2} \sum_{t=1}^T (X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p}) \Sigma_\varepsilon^{-1} (X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p})\right]$$

Quant à la log-vraisemblance du processus VAR(p), elle s'écrit :

$$\log \mathbf{L}(X_1, \dots, X_T) = -\frac{TN}{2} \log 2\pi - \frac{T}{2} \log \Sigma_\varepsilon - \frac{1}{2} \sum_{t=1}^T \varepsilon_t' \Sigma_\varepsilon^{-1} \varepsilon_t$$

La maximisation de cette log-vraisemblance permet d'obtenir les estimations des paramètres  $\phi_1, \phi_2, \dots, \phi_p$  et de la matrice de variance-covariance  $\Sigma_\varepsilon$ .

### 2.3.4 Détermination du nombre de retards

Dans le but de déterminer le nombre de retards d'un modèle à retards échelonnés, on va s'appuyer sur le critère de Akaike et Schwarz que nous avons déjà abordé. Ces critères peuvent être utilisés pour déterminer l'ordre  $p$  du modèle dans le cas d'une représentation VAR. La procédure de sélection de l'ordre de la représentation consiste à estimer tous les modèles VAR pour un ordre allant de 0 à  $h$  ( $h$  étant le retard maximum admissible par la théorie économique ou par les données disponibles).

Les critères  $AIC(p)$  et  $SIC(p)$  sont calculés de cette façon :

$$AIC(p) = \ln[\det |\Sigma_\varepsilon|] + \frac{2k^2p}{n}$$

$$SIC(p) = \ln[\det |\Sigma_\varepsilon|] + \frac{k^2p \ln(n)}{n}$$

avec  $k$  = nombre de variables du système.

$n$  : nombre d'observations.

$p$  : nombre de retards.

$\Sigma_\varepsilon$  : matrice des variances covariances des résidus du modèle.

Le retard  $p$  qui minimise les critères  $AIC$  et  $SIC$  est retenu.

### 2.3.5 Prévision dans un VAR stationnaire

#### Cas d'un VAR(1)

Soit le modèle VAR(1) centré suivant :

$$X_t = \phi_0 + \phi_1 X_{t-1} + \varepsilon_t \quad \forall t \in \mathbb{Z}$$

Supposons que l'on dispose de  $T$  réalisations  $(X_1, X_2, \dots, X_T)$ , d'un estimateur convergent  $\hat{\phi}_1^*$  de  $\phi_1$  et d'un estimateur convergent  $\hat{\phi}_0^*$  et de  $\phi_0$ , la prévision de la réalisation à la date  $T+1$  du processus  $X_t$  est donc donnée par :

$$\hat{X}_{T+1} = E(X_{T+1} | \underline{X}_T, X_{T-1}, \dots, X_1) = \hat{\phi}_0 + \hat{\phi}_1 X_t$$

A l'horizon  $T+2$  :

$$\hat{X}_{T+2} = E(X_{T+2} | \underline{X}_T, X_{T-1}, \dots, X_1) = \hat{\phi}_0 + \hat{\phi}_1 \hat{X}_{T+1} = (I + \hat{\phi}_1) \hat{\phi}_0 + \hat{\phi}_1^2 X_T$$

De façon générale, à l'horizon  $h$  donc, on aura

$$\hat{X}_{T+h} = E(X_{T+h} | \underline{X}_T, X_{T-1}, \dots, X_1) = (I + \hat{\phi}_1 + \hat{\phi}_1^2 + \hat{\phi}_1^{h-1}) \hat{\phi}_0 \hat{\phi}_1^h X_t$$

## Cas d'un VAR(p)

A la date T, on souhaite effectuer une prévision de  $X_{T+h}$ .  
Soit le modèle suivant :

$$\phi_B X_T = \mu + \varepsilon_t$$

avec  $\det \phi(Z)$  ayant toutes ses racines de module strictement supérieur à 1 et on s'intéresse à la solution stationnaire ( $X_t$ ).  
D'une manière itérative, on va obtenir :

$$\hat{X}_{T+1} = \mu + \phi_1 X_T + \dots + \phi_p X_{T+1+p} + \underbrace{EB(\varepsilon_{T+1} | \mathbf{1}, X_T)}_0$$

$$X_{T+2} = \mu + \phi_1 X_{T+1} + \dots + \phi_p X_{T+2+p} + \varepsilon_{T+2}$$

$$\hat{X}_{T+2} = \mu + \phi_1 \hat{X}_{T+1} + \dots + \hat{\phi}_p X_{T+2+p} + \underbrace{EB(\varepsilon_{T+2} | \mathbf{1}, X_T)}_0$$

·  
·  
·

$$X_{T+h} = \mu + \sum_p^1 \phi_k X_{T+h-k} + \varepsilon_{T+h}$$

$$\hat{X}_{T+h} = \mu + \sum_p^1 \phi_k \hat{X}_{T+h-k} + 0$$

avec  $\hat{X}_{T+h-k} = X_{T+h-k}$  si  $k \geq h$

### 2.3.6 Causalité

Une bonne connaissance du concept de causalité est importante pour une correcte formulation de la politique économique. En effet, connaître le sens de la causalité est aussi important que de mettre en évidence une liaison entre les variables économiques. Il existe plusieurs définitions de la causalité, on illustre :

- Causalité au sens de Granger(1969).
- Causalité au sens de Sims (1980).

## Causalité au sens de Granger

En 1969 Granger a proposé les concepts de causalité et exogénéité : la variable  $X_{2t}$  est la cause de  $X_{1t}$ , si la prédictibilité de  $X_{1t}$  est améliorée lorsque l'information relative à  $X_{2t}$  est incorporée dans l'analyse. Soit le modèle VAR(p) pour lequel les variables  $X_{1t}$  et  $X_{2t}$  sont stationnaires :

$$\begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} a_1^1 & b_1^1 \\ a_1^2 & b_1^2 \end{pmatrix} \begin{pmatrix} X_{1t-1} \\ X_{2t-1} \end{pmatrix} + \begin{pmatrix} a_2^1 & b_2^1 \\ a_2^2 & b_2^2 \end{pmatrix} \begin{pmatrix} X_{1t-2} \\ X_{2t-2} \end{pmatrix} + \dots + \begin{pmatrix} a_p^1 & b_p^1 \\ a_p^2 & b_p^2 \end{pmatrix} \begin{pmatrix} X_{1t-p} \\ X_{2t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

Le bloc de variables  $(X_{2t-1}, X_{2t-2}, \dots, X_{2t-p})$  est jugé comme exogène par rapport au bloc de variables  $(X_{1t-1}, X_{1t-2}, \dots, X_{1t-p})$ , si le fait de rajouter  $X_{2t}$  n'améliore pas significativement la détermination des variables  $X_{1t}$ , ce qui signifie qu'un test de restrictions sur les coefficients des variables  $X_{2t}$  de la représentation VAR noté RVAR (Restricted VAR) sera effectué. Aussi la détermination du retard  $p$  est fait par biais des critères AIC et SIC déjà vu.

On effectue un test de Fisher classique et on teste :

- $H_0 : b_1^1 = b_2^1 = \dots = b_p^1 = 0$ , si  $H_0$  est acceptée donc  $X_{2t}$  ne cause pas  $X_{1t}$ .
- $H_0 : a_1^2 = a_2^2 = \dots = a_p^2 = 0$ , si  $H_0$  est acceptée donc  $X_{1t}$  ne cause pas  $X_{2t}$ .

**Remarque 2.2.** *Le terme prédictibilité semble préférable au terme de causalité, en effet lorsque  $Y_t$  cause  $X_t$ , signifie qu'il est préférable de prédire  $X_t$  en connaissant  $Y_t$  que sans le connaître.*

## Causalité au sens de Sims

En 1980 Sims a présenté une spécification de test légèrement différente, en considérant que si les valeurs futures de  $X_{1t}$  permettent d'expliquer les valeurs présentes de  $X_{2t}$ , alors  $X_{2t}$  est la cause de  $X_{1t}$ . Ceci se traduit par la représentation suivante :

$$X_{1t} = a_1^0 + \sum_{i=1}^p a_{1i}^1 X_{1t-i} + \sum_{i=1}^p a_{1i}^2 X_{2t-i} + \sum_{i=1}^p b_i^2 X_{2t+i} + \varepsilon_{1t}$$

$$X_{2t} = a_2^0 + \sum_{i=1}^p a_{2i}^1 X_{1t-i} + \sum_{i=1}^p a_{2i}^2 X_{2t-i} + \sum_{i=1}^p b_i^1 X_{1t+i} + \varepsilon_{2t}$$

On va effectuer un test de Fisher classique :

- ✓  $H_0 : b_1^2 = b_2^2 = \dots = b_p^2 = 0$  Si cette hypothèse est acceptée donc  $X_{1t}$  ne cause pas  $X_{2t}$ .
- ✓  $H_0 : b_1^1 = b_2^1 = \dots = b_p^1 = 0$  Si cette hypothèse est acceptée donc  $X_{2t}$  ne cause pas  $X_{1t}$ .

## 2.4 Le concept de cointégration

L'analyse de la cointégration permet d'identifier clairement la relation véritable entre deux variables en recherchant l'existence d'un vecteur de cointégration et en éliminant son effet, le cas échéant.

### 2.4.1 Les propriétés de l'ordre d'intégration d'une série chronologique

Une série est intégrée d'ordre  $d$  (notée  $X_t \rightarrow I(d)$ ), s'il convient de la différencier  $d$  fois afin de la stationnariser.

Soit la série  $X_{1t}$  stationnaire et la série  $X_{2t}$  intégrée d'ordre 1 :

$$\begin{aligned} X_{1t} &\rightarrow I(0) \\ X_{2t} &\rightarrow I(1) \\ \implies X_{1t} + X_{2t} &\rightarrow I(0) \end{aligned}$$

La série  $Y_t = X_{1t} + X_{2t}$  n'est pas stationnaire puisque l'on somme deux séries dont une est affectée d'une tendance et l'autre est stationnaire.

Soient deux séries  $X_{1t}$  et  $X_{2t}$  intégrées d'ordre  $p$  :

$$\begin{aligned} X_{1t} &\rightarrow I(d) \\ X_{2t} &\rightarrow I(d) \\ \implies X_{1t} + X_{2t} &\rightarrow I(?) \end{aligned}$$

Soit la combinaison linéaire

$$\alpha X_{1t} + \beta X_{2t} \rightarrow I(?)$$

Donc le résultat dépend des signes des coefficients  $\alpha$ ,  $\beta$  et de l'existence d'une dynamique non stationnaire commune.

Prenons un autre cas dont  $d \neq d'$  :

$$\begin{aligned} X_{1t} &\rightarrow I(d) \\ X_{2t} &\rightarrow I(d') \\ \implies X_{1t} + X_{2t} &\rightarrow I(?) \end{aligned}$$

Il est impossible de conclure car on somme deux séries d'ordre d'intégration différent.

## 2.4.2 Représentation du modèle à correction d'erreur(VECM)

Le modèle VECM permet de modéliser conjointement les dynamiques de court terme (représenter par les variables en différence première) et de long terme (représentation par les variables en niveau). En effet, les avantages principaux du VECM sont :

- La capacité de capture d'une structure et des interactions dynamiques riches.
- la possibilité de tester et d'estimer une relation de cointégration.
- Il permet de faire l'inférence sur la relation de cointégration (paramètre à long terme) sur la dynamique du système (paramètre à court terme), ainsi que sur la vitesse d'ajustement.

Le modèle à correction d'erreur (VECM) s'obtient à partir de la dynamique de court terme. Par conséquent, nous pouvons distinguer deux dynamiques selon lesquelles sera estimé notre modèle.

✓ La dynamique de court terme suivante :

$$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-1} + a_3 Y_{t-1} + \mu_t.$$

✓ La dynamique de long terme suivante :

$$Y_t = aX_t + \beta + \varepsilon_t.$$

A partir de ces deux dynamiques on aura l'équation du modèle VECM suivante :

$$\Delta Y_t = \gamma \Delta X_t + \delta(y_{t-1} - aX_{t-1}) + \varepsilon_t.$$

# Chapitre 3

## La modélisation des dépenses sanitaires en Algérie

### 3.1 Introduction

L'objectif de ce chapitre est la modélisation des dépenses de santé. Nous allons donc appliquer la méthode d'estimation ARDL pour modéliser l'équilibre de long terme et la dynamique de court terme de l'évolution des dépenses de santé. La modélisation va nous permettre de savoir quels sont les facteurs qui expliquent la croissance des dépenses de santé en Algérie.

Cette partie du travail est composée de deux sections : dans une première section, nous allons présenter les variables à utiliser dans notre étude. Dans une deuxième section, nous allons nous intéresser à l'étude empirique (tests, estimation et validation du modèle choisi).

### 3.2 Choix des variables et données utilisées

Dans notre travail empirique qui porte sur la modélisation des dépenses de santé en Algérie, nous avons choisi comme variables endogènes (variable à expliquer) les dépenses sanitaires (DS), et le niveau de vie (DM), la densité médicale et la densité lits d'hôpitaux (DLH), la prise en charge publique (PCP), la croissance et le vieillissement de la population (POP) et la proportion des 60 ans et plus dans la population totale (POP60) comme des variables exogènes (variables explicatives). Cette section comprend les étapes suivantes : la spécification du modèle, l'identification des variables choisies.

#### Dépenses de santé (DS)

Les dépenses de santé par habitant ont connu des fluctuations sur la période étudiée, notamment, entre 1990 et 2000. A partir des années 2000, les dépenses ont connu une augmentation, passant de 6 695,23 DA en 2000 à un pic en 2013 de près de 17 000 DA

par habitant, avec un taux de croissance de 153,9 Cette hausse est, sans doute, dépend des facteurs exogènes. Dès lors, il importe de se demander quels sont les facteurs qui sont à l'origine de cette croissance ?

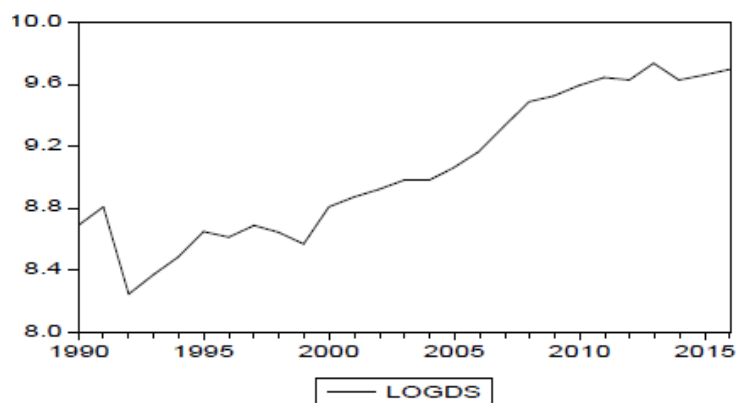


FIG. 3.1 – Evolution des dépenses de santé per capita (en log)

### Le niveau de vie

Le PIB par habitant est utilisé ici comme approximation du niveau de vie. En Algérie, le PIB a connu une croissance spectaculaire, en particulier à partir des années 2000, correspondant à la période de l'aisance financière, en raison de l'augmentation des prix des hydrocarbures. Avec à peine 143 000 DA en 1990, il atteint plus de 390000 DA en 2016, en dinars constant, soit une augmentation de plus de 172%.

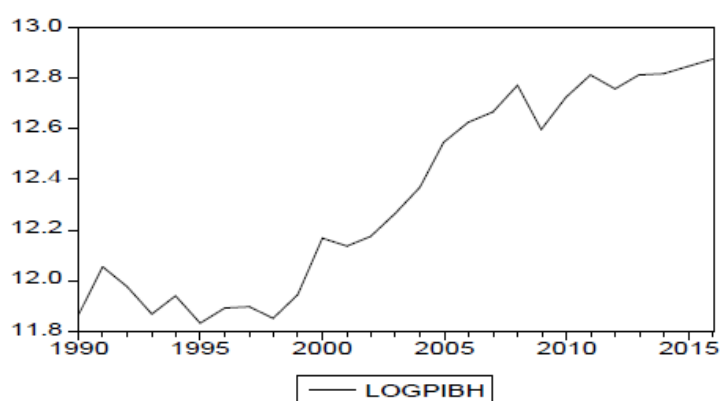


FIG. 3.2 – Evolution du PIB par habitant

### La densité médicale et la densité lits d'hôpitaux (DLH)

En Algérie, de 1990 à 2016, le nombre de médecins tant public que privé, a augmenté de 160%. Cette augmentation est largement supérieure à la croissance de la population. Cette observation a conduit à formuler l'hypothèse selon laquelle l'augmentation de l'offre conduit à un accroissement de la demande (demande induite par l'offre). Contrairement à toutes les variables introduites dans notre modèle, la densité de lits d'hôpital a connu une diminution, malgré le nombre de lits en croissance. Depuis 1990, la densité de lits passe de 2,5 lits pour 1000 habitants à 1,5 lits pour 1000 habitants en 2016, ceci est dû, sans doute, à une forte croissance de la population. Il est à noter, qu'il s'agit, exclusivement, de lits publics, car nous ne disposons pas de données concernant le secteur privé pour la période étudiée.

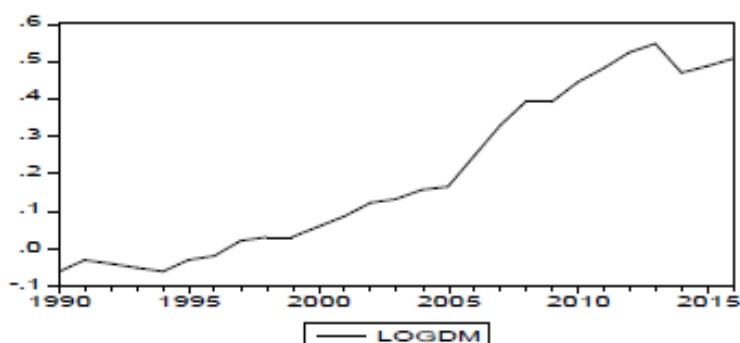


FIG. 3.3 – Evolution de la densité médicale

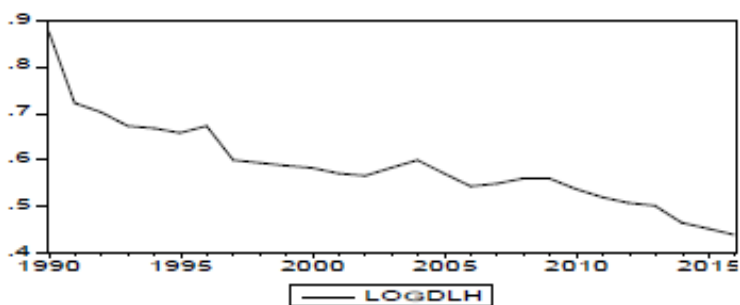


FIG. 3.4 – Evolution de la densité de lits d'hôpitaux

### La prise en charge publique (PCP)

Le système de santé algérien est financé en grande partie par l'Etat et la sécurité sociale. Leurs contributions à ce financement sont variables dans le temps. Elles dépendent, bien évidemment, de la situation économique du pays. En effet, avec une participation de moins de 75% en moyenne entre 1990 à 2000, elle dépasse 77% à partir des années 2000, grâce à la conjoncture économique favorable qu'a connu le pays.

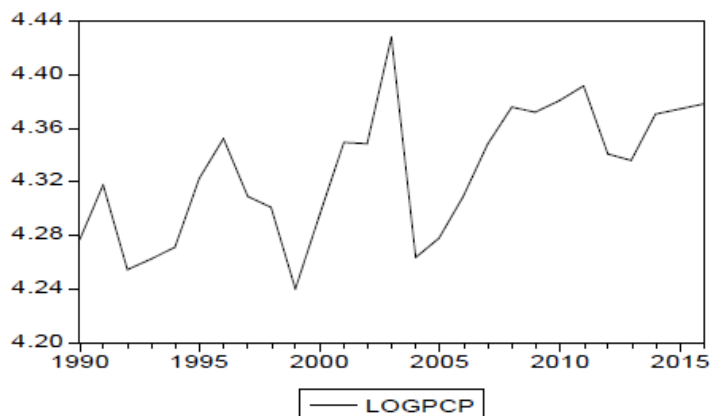


FIG. 3.5 – Evolution de la prise en charge publique

### La croissance et le vieillissement de la population (POP)

L'augmentation de la population totale est traditionnellement évoquée comme facteur important dans la variation des dépenses de santé. Il est à souligner que lorsqu'on évoque l'effet de la structure démographique, cela sous entend à la fois l'augmentation de la population totale et le vieillissement de celle-ci. La consommation des soins et l'âge sont supposés intimement liés. La part de la population de 60 ans et plus ne cesse de croître en Algérie.

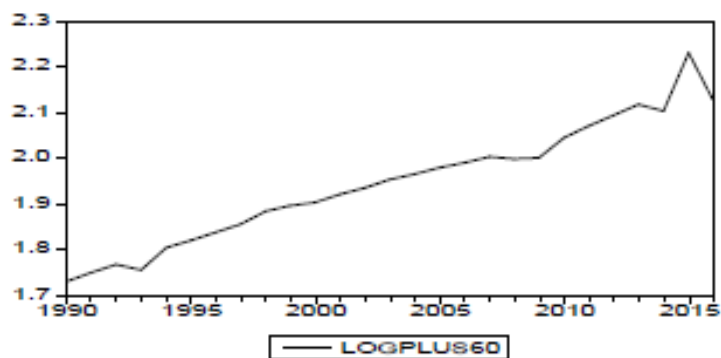


FIG. 3.6 – Evolution de la part des 60 ans et plus dans la population totale

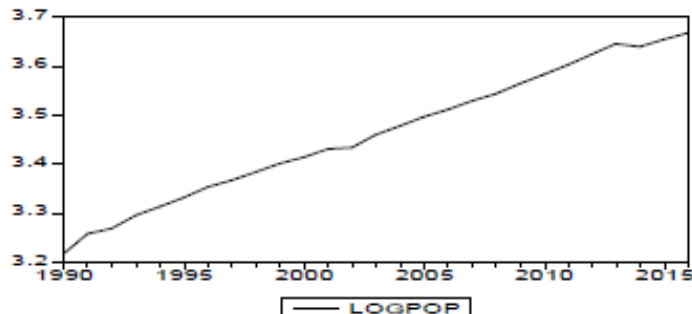


FIG. 3.7 – Evolution de la population totale

### 3.2.1 Spécification du modèle

L'objectif de cette étape est d'examiner l'influence des variables explicatives retenues sur l'évolution des dépenses de santé en Algérie, durant la période considérée, à travers un modèle de cointégration. Pour éviter le problème d'hétéroscédasticité, la transformation des séries d'origine en logarithme est indispensable. On peut alors écrire le modèle sous forme d'équation suivante :

$$\log DS_t = \alpha + \beta_1 \log PIBH_t + \beta_2 \log PCP_t + \beta_3 \log POP_t + \beta_4 \log DM_t + \beta_5 \log DLH_t + \beta_6 \log POP60_t + \varepsilon_t \quad (3.1)$$

avec  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$  sont respectivement les multiplicateurs des variables PIBH, PCP, POP, DM, DLH et POP60 et  $\varepsilon_t$  le terme de l'erreur.

En se basant sur l'approche ARDL, l'équation (3.1) peut s'écrire de la manière suivante :

$$\begin{aligned} \Delta \log DS_t = & \alpha_0 + \sum_{i=1}^p \delta_{1i} \Delta \log DS_{t-i} + \sum_{i=0}^p \delta_{2i} \Delta \log PIBH_{t-i} + \sum_{i=0}^p \delta_{3i} \Delta \log PCP_{t-i} + \sum_{i=0}^p \delta_{4i} \Delta \log \\ & POP_{t-i} + \sum_{i=0}^p \delta_{5i} \Delta \log DM_{t-i} + \sum_{i=0}^p \delta_{6i} \Delta \log DLH_{t-i} + \sum_{i=0}^p \delta_{7i} \Delta \log POP60_{t-i} \\ & + \beta_1 \log DS_{t-1} + \beta_2 \log PIBH_{t-1} + \beta_3 \log PCP_{t-1} + \beta_4 \log POP_{t-1} + \\ & \beta_5 \log DM_{t-1} + \beta_6 \log DLH_{t-1} + \beta_7 \log POP60_{t-1} + \varepsilon_t \end{aligned} \quad (3.2)$$

où  $\Delta$  est la différence première des variables .

$\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7$  désignent la dynamique de court terme du modèle.

$\beta_1$  à  $\beta_7$  représentent la relation de long terme.

Le bounds test est utilisé pour tester la relation de cointégration entre les variables. Cette procédure est basée sur la F-statistique. L'hypothèse nulle est l'absence de la relation de cointégration  $H_0$  contre l'hypothèse alternative  $H_1$ .

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$$

contre

$$H_1 : \beta_1 \neq \beta_2 \neq \beta_3 \neq \beta_4 \neq \beta_5 \neq \beta_6 \neq \beta_7 \neq 0$$

Une fois que les résultats des tests rejettent l'hypothèse nulle de l'absence de la cointégration, alors il est possible de passer à la deuxième étape de la procédure d'estimation du modèle de long terme ARDL ( $p_1, p_2, p_3, p_4, p_5, p_6, p_7$ ) comme suit :

$$\begin{aligned} \log DS_t = & \alpha_0 + \sum_{i=1}^{p_1} \beta_1 \log DS_{t-i} + \sum_{i=0}^{p_2} \beta_2 \log PIBH_{t-i} + \sum_{i=0}^{p_3} \beta_3 \log \\ & POP_{t-i} + \sum_{i=0}^{p_4} \beta_4 \log PCP_{t-i} + \sum_{i=0}^{p_5} \beta_5 \log DM_{t-i} + \sum_{i=0}^{p_6} \beta_6 \log \\ & DLH_{t-i} + \sum_{i=0}^{p_7} \beta_7 \log POP60_{t-i} + \varepsilon_t \end{aligned} \quad (3.3)$$

Enfin, on estime, par le modèle à correction d'erreur (ECM), les paramètres de la dynamique de court terme associés aux paramètres de long terme. La spécification de cette estimation peut être présentée comme suit :

$$\begin{aligned} \Delta \log DS_t = & \sum_{i=1}^{p-1} \delta_{1i} \log DS_{t-i} + \sum_{i=0}^{p-1} \delta_{2i} \log PIBH_{t-i} + \sum_{i=0}^{p-1} \delta_{3i} \log POP_{t-i} + \\ & \sum_{i=0}^{p-1} \delta_{4i} \log PCP_{t-i} + \sum_{i=0}^{p-1} \delta_{5i} \log DM_{t-i} + \sum_{i=0}^{p-1} \delta_{6i} \log DLH_{t-i} + \\ & \sum_{i=0}^{p-1} \delta_{7i} \log POP60_{t-i} + \lambda \approx ECM_{t-1} + \varepsilon_t \end{aligned} \quad (3.4)$$

où  $\delta$  sont les coefficients de la dynamique de court terme (de la convergence du modèle vers l'équilibre).

$ECM_{t-1}$  sont les résidus retardés obtenus de la relation de long terme ARDL de l'équation (3.2) estimée.

$\lambda$  représente la vitesse d'ajustement,  $\lambda > 0$ .

### 3.3 Analyse Empirique

Cette section comprend les étapes suivantes : Les tests de racine unitaire, et les tests de cointégration, estimation et validation du modèle (tests d'autocorrélation des erreurs, test de normalité des résidus, test de CUSUM).

### 3.3.1 Etude de la stationnarité

Pour toute analyse d'étude statistique, on doit vérifier les caractéristiques aléatoires des séries. Il s'agit d'étudier leurs stationnarité. pour cela, nous utiliserons le test de stationnarité de Dickey-Fuller (ADF), le principe de ce test est bien expliqué dans le chapitre 1, il permet de vérifier l'existence ou non de racine unitaire.

Les résultats du test Dickey-Fuller effectué sur les séries sont donnés à partir d'EvIEWS 10 comme suit :

#### Etude de la série log (DS)

<u>Modèle avec tendance et constante (3)</u>		
Null Hypothesis: LOGDS has a unit root		
Exogenous: Constant, Linear Trend		
Lag Length: 7 (Automatic based on SIC, MAXLAG=8)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.408848	0.3637
Test critical values:	1% level	-4.532598
	5% level	-3.673616
	10% level	-3.277364
*MacKinnon (1996) one-sided p-values.		
Warning: Probabilities and critical values calculated for 20 observations and may not be accurate for a sample size of 19		
<u>Modèle sans tendance et avec constante (2)</u>		
Null Hypothesis: LOGDS has a unit root		
Exogenous: Constant		
Lag Length: 0 (Automatic based on SIC, MAXLAG=8)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.358798	0.9025
Test critical values:	1% level	-3.711457
	5% level	-2.981038
	10% level	-2.629906
*MacKinnon (1996) one-sided p-values.		
<u>Modèle sans constante (1)</u>		
Null Hypothesis: LOGDS has a unit root		
Exogenous: None		
Lag Length: 0 (Automatic based on SIC, MAXLAG=8)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	1.320675	0.9487
Test critical values:	1% level	-2.656915
	5% level	-1.954414
	10% level	-1.609329
*MacKinnon (1996) one-sided p-values.		

TAB. 3.1 – Application du test ADF sur la série log (DS)

Les probabilités critiques sont toutes supérieures à 0.05, nous ne rejetons pas l'hypothèse  $H_0$ , nous pouvons donc conclure que le processus LOGDS possède une racine unitaire et n'est donc pas stationnaire. On applique maintenant le test ADF à la série en différence première.

Null Hypothesis: D(LOGDS) has a unit root		
Exogenous: Constant, Linear Trend		
Lag Length: 0 (Automatic based on SIC, MAXLAG=8)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.902318	0.0003
Test critical values:	1% level	-4.374307
	5% level	-3.603202
	10% level	-3.238054

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(LOGDS,2)				
Method: Least Squares				
Date: 01/10/17 Time: 19:12				
Sample(adjusted): 1992 2016				
Included observations: 25 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LOGDS(-1))	-1.214107	0.205700	-5.902318	0.0000
C	-0.016394	0.064872	-0.252712	0.8028
@TREND(1990)	0.004287	0.004164	1.029602	0.3144
R-squared	0.613131	Mean dependent var	-0.003308	
Adjusted R-squared	0.577961	S.D. dependent var	0.228619	
S.E. of regression	0.148521	Akaike info criterion	-0.864010	
Sum squared resid	0.485289	Schwarz criterion	-0.717745	
Log likelihood	13.80012	F-statistic	17.43338	
Durbin-Watson stat	1.170762	Prob(F-statistic)	0.000029	

TAB. 3.2 – Application du test ADF sur la série  $\Delta \log$  (DS)

On remarque qu'on a  $-5.90 < -3.60$  donc la série  $\Delta(\text{LOGDS})$  est stationnaire. On conclue qu'après avoir différencié la série LOGDS une fois, on obtient une série stationnaire.

## Etude de la série log (DM)

### Modèle avec tendance et constante (3)

Null Hypothesis: LOGDM has a unit root

Exogenous: Constant, Linear Trend

Lag Length: 8 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.092964	0.5150
Test critical values: 1% level	-4.571559	
5% level	-3.690814	
10% level	-3.286909	

\*MacKinnon (1996) one-sided p-values.

Warning: Probabilities and critical values calculated for 20 observations and may not be accurate for a sample size of 18

### Modèle avec constante (2)

Null Hypothesis: LOGDM has a unit root

Exogenous: Constant

Lag Length: 0 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	0.051392	0.9552
Test critical values: 1% level	-3.711457	
5% level	-2.981038	
10% level	-2.629906	

\*MacKinnon (1996) one-sided p-values.

### Modèle sans constante (1)

Null Hypothesis: LOGDM has a unit root

Exogenous: None

Lag Length: 0 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	2.063560	0.9884
Test critical values: 1% level	-2.656915	
5% level	-1.954414	
10% level	-1.609329	

\*MacKinnon (1996) one-sided p-values.

TAB. 3.3 – Application du test ADF sur la série log (DM)

Les probabilités critiques sont toutes supérieures à 0.05, nous ne rejetons pas l'hypothèse  $H_0$ , nous pouvons donc conclure que le processus LOG(DM) possède une racine unitaire et n'est donc pas stationnaire.

Etude de la série différenciée.

Null Hypothesis: D(LOGDM) has a unit root

Exogenous: Constant, Linear Trend

Lag Length: 7 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.754533	0.0447
Test critical values: 1% level	-4.571559	
5% level	-3.690814	
10% level	-3.286909	

\*MacKinnon (1996) one-sided p-values.

Warning: Probabilities and critical values calculated for 20 observations and may not be accurate for a sample size of 18

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(LOGDM,2)

Method: Least Squares

Date: 01/10/17 Time: 19:21

Sample(adjusted): 1999 2016

Included observations: 18 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LOGDM(-1))	-5.450793	1.451790	-3.754533	0.0056
D(LOGDM(-1),2)	4.280413	1.382010	3.097237	0.0147
D(LOGDM(-2),2)	4.664599	1.442096	3.234596	0.0120
D(LOGDM(-3),2)	4.022962	1.277593	3.148861	0.0136
D(LOGDM(-4),2)	3.666288	1.137700	3.222545	0.0122
D(LOGDM(-5),2)	3.269407	0.848512	3.853107	0.0049
D(LOGDM(-6),2)	2.574314	0.660197	3.899310	0.0045
D(LOGDM(-7),2)	1.664331	0.497807	3.343325	0.0102
C	-0.070958	0.038510	-1.842601	0.1026
@TREND(1990)	0.012431	0.004122	3.015849	0.0167
R-squared	0.841375	Mean dependent var		0.000471
Adjusted R-squared	0.662921	S.D. dependent var		0.045361
S.E. of regression	0.026336	Akaike info criterion		-4.135573
Sum squared resid	0.005549	Schwarz criterion		-3.640922
Log likelihood	47.22016	F-statistic		4.714812
Durbin-Watson stat	2.196861	Prob(F-statistic)		0.019872

TAB. 3.4 – Application du test ADF sur la série  $\Delta \log (DM)$

On remarque qu'on a  $-3.75 < -3.69$  donc la série  $\Delta(\text{LOGDM})$  est stationnaire. On conclue qu'après avoir différencié la série LOGDM une fois, on obtient une série stationnaire.

## Etude de la série log (POP)

Modèle avec tendance et constante (3)

Null Hypothesis: LOGPOP has a unit root

Exogenous: Constant, Linear Trend

Lag Length: 0 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.730986	0.0380
Test critical values: 1% level	-4.356068	
5% level	-3.595026	
10% level	-3.233456	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(LOGPOP)

Method: Least Squares

Date: 01/10/17 Time: 18:08

Sample(adjusted): 1991 2016

Included observations: 26 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOGPOP(-1)	-0.572950	0.153565	-3.730986	0.0011
C	1.869210	0.495132	3.775176	0.0010
@TREND(1990)	0.009416	0.002620	3.593641	0.0015
R-squared	0.448155	Mean dependent var		0.017280
Adjusted R-squared	0.400168	S.D. dependent var		0.007749
S.E. of regression	0.006002	Akaike info criterion		-7.285434
Sum squared resid	0.000828	Schwarz criterion		-7.140269
Log likelihood	97.71064	F-statistic		9.339172
Durbin-Watson stat	1.787083	Prob(F-statistic)		0.001074

TAB. 3.5 – Application du test ADF sur la série log (POP)

La série log(POP) est stationnaire car on a  $-3.73 < -3.59$ .

## Etude de la série log (PCP)

Null Hypothesis: LOGPCP has a unit root  
 Exogenous: Constant, Linear Trend  
 Lag Length: 3 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.608464	0.0008
Test critical values: 1% level	-4.416345	
5% level	-3.622033	
10% level	-3.248592	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(LOGPCP)  
 Method: Least Squares  
 Date: 01/10/17 Time: 19:48  
 Sample(adjusted): 1994 2016  
 Included observations: 23 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOGPCP(-1)	-2.368834	0.422368	-5.608464	0.0000
D(LOGPCP(-1))	1.199327	0.299545	4.003827	0.0009
D(LOGPCP(-2))	0.973947	0.250152	3.893416	0.0012
D(LOGPCP(-3))	0.519259	0.200019	2.596048	0.0188
C	10.11482	1.800995	5.616238	0.0000
@TREND(1990)	0.009340	0.001993	4.687143	0.0002
R-squared	0.711885	Mean dependent var		0.005004
Adjusted R-squared	0.627145	S.D. dependent var		0.050232
S.E. of regression	0.030672	Akaike info criterion		-3.911448
Sum squared resid	0.015994	Schwarz criterion		-3.615233
Log likelihood	50.98166	F-statistic		8.400849
Durbin-Watson stat	2.061686	Prob(F-statistic)		0.000373

TAB. 3.6 – Application du test ADF sur la série log (PCP)

La série LOG(PCP) est stationnaire, car on a  $-5.60 < -3.62$ .

## Etude de la série log (PIBH)

### Modèle 3.

Null Hypothesis: LOGPIBH has a unit root

Exogenous: Constant, Linear Trend

Lag Length: 0 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.044925	0.5507
Test critical values: 1% level	-4.356068	
5% level	-3.595026	
10% level	-3.233456	

\*MacKinnon (1996) one-sided p-values.

### Modèle 2

Null Hypothesis: LOGPIBH has a unit root

Exogenous: Constant

Lag Length: 0 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.413771	0.8927
Test critical values: 1% level	-3.711457	
5% level	-2.981038	
10% level	-2.629906	

\*MacKinnon (1996) one-sided p-values.

### Modèle 1.

Null Hypothesis: LOGPIBH has a unit root

Exogenous: None

Lag Length: 0 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	2.046003	0.9879
Test critical values: 1% level	-2.656915	
5% level	-1.954414	
10% level	-1.609329	

\*MacKinnon (1996) one-sided p-values.

TAB. 3.7 – Application du test ADF sur la série log (PIBH)

Les probabilités critiques sont toutes supérieures à 0.05, nous ne rejetons pas l'hypothèse  $H_0$ , nous pouvons donc conclure que le processus LOGPIBH possède une racine unitaire

et n'est donc pas stationnaire.

### Etude de la série LOG(PIBH) différenciée

#### Modèle 3.

Null Hypothesis: D(LOGPIBH) has a unit root  
 Exogenous: Constant, Linear Trend  
 Lag Length: 1 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.860433	0.0305
Test critical values: 1% level	-4.394309	
5% level	-3.612199	
10% level	-3.243079	

\*MacKinnon (1996) one-sided p-values.

TAB. 3.8 – Application du test ADF sur la série  $\Delta \log(\text{PIBH})$

La série LOG(PIBH) est stationnaire, car on a  $-3.86 < -3.61$ .

### Etude de la série log (POP60)

#### Modèle 3.

Null Hypothesis: LOGPOPS60 has a unit root  
 Exogenous: Constant, Linear Trend  
 Lag Length: 1 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.345292	0.8520
Test critical values: 1% level	-4.374307	
5% level	-3.603202	
10% level	-3.238054	

\*MacKinnon (1996) one-sided p-values.

#### Modèle 2.

Null Hypothesis: LOGPOP60 has a unit root  
 Exogenous: Constant  
 Lag Length: 1 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	0.703899	0.9897
Test critical values: 1% level	-3.724070	
5% level	-2.986225	
10% level	-2.632604	

\*MacKinnon (1996) one-sided p-values.

Modèle (1).

Null Hypothesis: LOGPOP60 has a unit root

Exogenous: None

Lag Length: 1 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	6.018207	1.0000
Test critical values: 1% level	-2.660720	
5% level	-1.955020	
10% level	-1.609070	

\*MacKinnon (1996) one-sided p-values.

TAB. 3.9 – Application du test ADF sur la série log (POP60)

Les probabilités critiques sont toutes supérieures à 0.05, nous ne rejetons pas l'hypothèse  $H_0$ , nous pouvons donc conclure que le processus LOG(PLUS60) possède une racine unitaire et n'est donc pas stationnaire.

**Test de Dickey-Fuller appliqué à la série différenciée**

Null Hypothesis: D(LOGPOP60) has a unit root

Exogenous: Constant, Linear Trend

Lag Length: 0 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-10.79330	0.0000
Test critical values: 1% level	-4.374307	
5% level	-3.603202	
10% level	-3.238054	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(LOGPOP60,2)

Method: Least Squares

Date: 01/10/17 Time: 20:42

Sample(adjusted): 1992 2016

Included observations: 25 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LOGPOP60(-1))	-2.070218	0.191806	-10.79330	0.0000
C	0.028379	0.010615	2.673617	0.0139
@TREND(1990)	0.000588	0.000685	0.858621	0.3998
R-squared	0.844380	Mean dependent var		-0.004833
Adjusted R-squared	0.830233	S.D. dependent var		0.058298
S.E. of regression	0.024021	Akaike info criterion		-4.507651
Sum squared resid	0.012694	Schwarz criterion		-4.361386
Log likelihood	59.34563	F-statistic		59.68514
Durbin-Watson stat	2.048091	Prob(F-statistic)		0.000000

TAB. 3.10 – Application du test ADF sur la série log (POP60)

La série LOG(POP60) est stationnaire, car on a  $-10.79 < -3.60$ .

### Etude de la série log (DLH)

#### Modèle 03.

Null Hypothesis: LOGDLH has a unit root  
 Exogenous: Constant, Linear Trend  
 Lag Length: 0 (Automatic based on SIC, MAXLAG=8)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-6.962782	0.0000
Test critical values:		
1% level	-4.356068	
5% level	-3.595026	
10% level	-3.233456	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(LOGDLH)  
 Method: Least Squares  
 Date: 01/10/17 Time: 21:12  
 Sample(adjusted): 1991 2016  
 Included observations: 26 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOGDLH(-1)	-0.678105	0.097390	-6.962782	0.0000
C	0.464412	0.072029	6.447611	0.0000
@TREND(1990)	-0.005875	0.001144	-5.137493	0.0000
R-squared	0.705769	Mean dependent var		-0.016816
Adjusted R-squared	0.680184	S.D. dependent var		0.033728
S.E. of regression	0.019074	Akaike info criterion		-4.972795
Sum squared resid	0.008368	Schwarz criterion		-4.827630
Log likelihood	67.64633	F-statistic		27.58491
Durbin-Watson stat	1.436135	Prob(F-statistic)		0.000001

TAB. 3.11 – Application du test ADF sur la série log (DLH)

La série log(DLH) est stationnaire, car on a  $-6.96 < -3.59$ .

L'application du test de stationnarité ADF nous conduit à accepter l'hypothèse  $H_0$  de non stationnarité pour les variables PIBH, DS, DM et POP60. Par contre les variables PCP, POP et DLH sont stationnaires en niveau. Les résultats montrent qu'il y a des séries qui sont stationnaires en niveau et d'autres non stationnaires. Il convient, tout de même,

de signaler qu'aucune des séries considérées dans notre analyse n'est intégrée d'ordre 2 ou plus, car cette condition est primordiale pour l'application de l'approche ARDL.

Variables	En niveau		Première différence		Ordre d'intégration
	T-statistique	<i>P-value</i>	T-statistique	<i>P-value</i>	
Log DS	-2.408848	0.3637	-5.902318	0.0003*	I (1)
Log PIBH	-3.860433	0.0305	-3.860433	0.0305**	I (1)
Log PCP	-5.608464	0.0008*	-	-	I (0)
Log POP	-3.730986	0.0380**	-	-	I (0)
Log DM	-2.092964	0.5150	-3.754533	0.0447**	I (1)
Log DLH	-6.962782	0.0000*	-	-	I (0)
Log POP60	-1.345292	0.8520	-10.79330	0.0000*	I (1)

\* Significatif à 1%, \*\* Significatif à 5%, \*\*\* Significatif à 10%.

TAB. 3.12 – Résultats du test ADF de racine unitaire

### 3.3.2 Test de cointégration et estimation du modèle ARDL

Dans cette partie, on va s'intéresser au test de cointégration, donc on va tester la présence ou pas d'une relation de cointégration entre les variables étudiées, et cela se fait à partir du test de Fisher, il s'agit de tester l'hypothèse nulle (absence de cointégration) et dans le cas de rejet de cette hypothèse, l'existence d'une relation de cointégration est confirmée.

Le choix de  $p$  est très important dans la mesure où l'introduction d'un nombre insuffisant de retards peut rendre la qualité du test problématique. La valeur  $p$  peut être déterminée par plusieurs méthodes, dans notre étude, on va utiliser le critère de Schwarz (SIC), le critère d'information d'Akaike (AIC).

#### Détermination du nombre de retard $p$

À partir du graphe ci-dessous et selon le critère d'information Schwarz, on constate que le modèle ARDL(1.0.0.1.1.1.1) est le meilleur modèle car la valeur du SIC est la minimale.

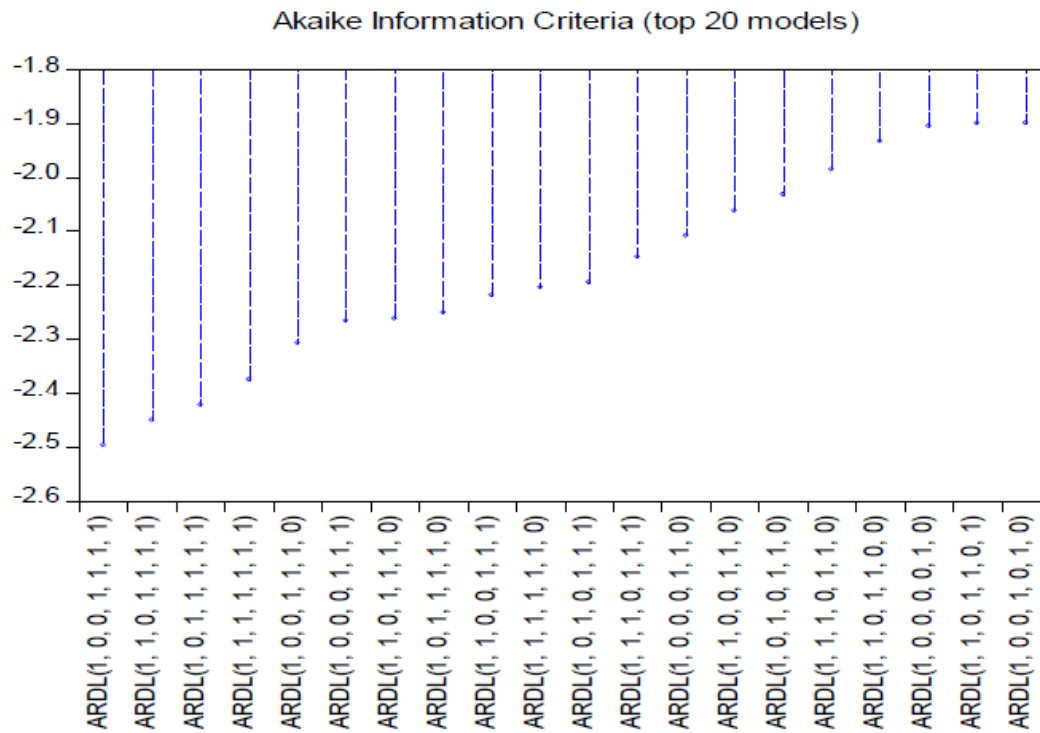


FIG. 3.8 – Le graphique du critère d'information Schwarz (SIC)

### Estimation du modèle ARDL (1.0.0.1.1.1.1)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LogDS(-1)	-0.180469	0.144191	-1.251599	0.2312
LogPIBH	0.091069	0.126369	0.720660	0.4830
Log POP	0.463246	2.714696	0.170644	0.8669
LogPOP(-1)	3.646645	2.407973	1.514404	0.1522
LogPCP	1.466814	0.296597	4.945472	0.0002
LogPCP(-1)	-0.497102	0.376948	-1.318753	0.2084
LogDM	1.887734	0.521934	3.616803	0.0028
LogDM(-1)	-1.356290	0.423954	-3.199149	0.0064
LogDLH	0.347516	0.748444	0.464319	0.6496
LogDLH(-1)	2.679450	0.519924	5.153544	0.0001
LogPOP60	0.471727	1.323073	0.356539	0.7268
C	-11.67765	3.371449	-3.463690	0.0038
$R^2 = 0.9933$ , $R^2$ ajusté = 0.98 AIC = -2.793159, DW = 2.67				

TAB. 3.13 – Estimation du modèle ARDL(1.0.0.1.1.1.1)

D'après les résultats d'estimation et selon la règle de Granger telque  $R^2 = 0.98 < DW = 2.67$  le modèle nous confirme que les variables utilisées sont bien stationnaires

## Test de cointégration

Selon l'approche de Pesaran et *al*(2001), le test de cointégration dans les modèles ARDL consiste à tester la nullité conjointe des coefficients des variables en niveau et retardées du modèle. Les résultats du Bounds-test sont présentés dans le tableau ci-dessous :

Signification	I0 Bound	I1 Bound
10%	1.99	2.94
5%	2.27	3.28
2.5%	2.55	3.61
1%	2.88	3.99
<b>F-statistique= 15.68834</b>		

TAB. 3.14 – Bounds-test

On constate que la statistique de Fisher (15.68834) est supérieure à la borne supérieure pour les différents seuils de significativité. Donc on rejette l'hypothèse nulle et on conclue à l'existence d'une relation de cointégration entre les dépenses de santé et leurs déterminants considérés dans cette étude.

### 3.3.3 L'estimations de long et de court terme selon le modèle ARDL

Il s'agit des estimations d'équilibre de long terme et de la dynamique de court terme entre les dépenses de santé et leurs déterminants.

#### Estimation de la relation à long terme selon le modèle ARDL

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LogPIBH	0.077147	0.104044	0.741479	0.4707
LogPOP	3.481573**	1.202350	2.871756	0.0123
LogPCP	0.821463**	0.364300	2.254912	0.0407
LogDM	0.450197***	0.251933	1.786968	0.0956
LogDLH	2.564205*	0.787388	3.256598	0.0057
LogPOP60	0.399610	1.117284	0.357662	0.7259
C	-9.892380	2.814999	-3.514169	0.0034
* Significatif à 1%, ** Significatif à 5%, *** Significatif à 10%.				
Log (DS) = -9,89 + 0,077 log(PIBH) + 3,48 log (POP) + 0,82 log (PCP) + 0,45 log (DM) + 2,56 log (DLH) + 0,39 log (POP60).				

TAB. 3.15 – Résultats de l'estimation de la relation à long terme

On a :

$$\log(DS) = -9.89 + 0.077 \log(PIBH) + 3.48 \log(POP) + 0.82 \log(PCP) + 0.45 \log(DM) + 2.5 \log(DLH) + 0.39 \log(POP60)$$

Contrairement à toute attente, les estimations des coefficients de la relation de long terme à partir du modèle ARDL montrent que le PIB par habitant (PIBH) n'est pas significatif. De plus, l'élasticité-revenue des dépenses de santé est très faible, soit 0.077. Ces résultats suggèrent que le niveau de vie, mesuré par le PIB par habitant n'a pas d'impact significatif sur l'évolution des dépenses de santé, à long terme.

L'impact significatif de l'augmentation de la population totale sur les dépenses de santé est confirmé par nos résultats. Le tableau montre une élasticité élevée de 3,48 au seuil de 5%.

Avec une élasticité de 0,82 au seuil de significativité de 5%, les résultats de nos estimations montrent l'impact de long terme de la prise en charge publique (PCP). La densité médicale (DM) présente une élasticité relativement faible de 0,45 mais significative au seuil de 10%. Quant à la densité de lits d'hôpitaux, cette variable est très significative au seuil de 1% avec une élasticité positive très élevée de 2,56.

Tout comme le PIBH, la proportion de la population de 60 ans et plus dans la population totale (POP60) n'a pas d'impact sur les dépenses de santé, même si cette variable présente une élasticité positive, mais statistiquement non significative.

#### Estimation de la relation à court terme selon le modèle ARDL

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLogPIBH	0.118441	0.095367	1.241960	0.2347
DLogPOP	1.200037	1.081891	1.109203	0.2860
DLogPCP	1.432822*	0.179079	8.001083	0.0000
DLogDM	1.897642*	0.299124	6.343989	0.0000
DLogDLH	-0.376886	0.282375	1.334700	0.2033
DLogPOP60	-0.448091	0.944693	-0.474324	0.6426
ECM <sub>t-1</sub>	-1.160224*	0.096995	-11.961726	0.0000

\* Significatif à 1%, \*\* Significatif à 5%, \*\*\* Significatif à 10 %.

TAB. 3.16 – Résultats de l'estimation de la relation à court terme

On désigne par D la différence première des variables considérées. En outre, le terme  $ECM_{t-1}$  correspond au résidu retardé d'une période issu de l'équation d'équilibre de long terme. Son coefficient estimé est négatif et largement significatif au seuil de 1% , confirmant ainsi l'existence d'un mécanisme à correction d'erreur. Ce coefficient, qui exprime le degré

avec lequel la variable DS sera rappelée vers la cible de long terme, est estimé de (-1,160) pour notre modèle ARDL.

Les résultats des estimations de court terme ne présentent pas de différences quant à l'impact du PIBH sur les dépenses de santé. En effet, avec une élasticité toujours faible, cette variable n'est pas significative même à court terme. Contrairement à ce qu'elle présente à long terme, la variable POP ne semble pas avoir la même importance à court terme. En effet, cette variable n'est pas, statistiquement, significative à court terme, même en présence d'un coefficient relativement élevé et positif.

La prise en charge publique (PCP) demeure un facteur important même à court terme. Elle présente une élasticité très élevée (1,43) et très significative au seuil de 1%. Même chose, pour l'impact de la densité médicale, cette variable est significative, elle présente une élasticité de 1,89 beaucoup plus élevée à celle trouvée dans les estimations de long terme.

Pour ce qui est de la variable de la densité de lits, les estimations de court terme présentent des différences notables par rapport à celles de long terme. Étant un facteur déterminant des dépenses de santé à long terme, cette variable, en plus de présenter une élasticité négative, ne semble pas être significative à court terme. La variable de la proportion de 60 ans et plus dans la population totale demeure non significative à court terme et présente une élasticité négative.

### **3.3.4 Validation du modèle**

La validation du modèle se réfère à divers tests statistiques de spécification qui permettent de vérifier si le modèle est congru et pertinent, c'est à dire, qu'il ne peut être mis à défaut.

#### **Test d'autocorrélation des erreurs**

L'une des hypothèses qui doit être vérifiées de manière rigoureuse pour la validité d'un modèle ajusté est celle de bruit blanc. En effet, si les résidus ne forment pas un bruit blanc on peut penser à une mauvaise stationnarisation des données ou même à un mauvais choix du modèle. Si les résidus obéissent à un bruit blanc, il ne doit pas exister d'autocorrélation dans la série. Les résultats du tableau ci-dessous montrent une indépendance des erreurs, car la plus value est de 0.22, supérieure à 0.05.

F-statistic		2.223872	Prob. F(2,12)		0.1508
Obs*R-squared		10.030834	Prob. Chi-Square(2)		0.2297
Variable	Coefficient	Std Error	t-Statistic	Prob.	
LogDS(-1)	0,073588	0.158640	0.463870	0.6510	
LogPIBH	-0.037391	0.133034	-0.281061	0.7835	
LogPOP	-0.304673	2.511847	-0.121295	0.9055	
LogPOP(-1)	0.536858	2.243770	0.239266	0.8149	
LogPCP	0.085530	0.288617	0.296343	0.7720	
LogPCP(-1)	-0.087514	0.409303	-0.213813	0.8343	
LogDM	-0.137111	0.497996	-0.275325	0.7878	
LogDM(-1)	0.000620	0.427841	0.001449	0.9989	
LogDLH	0.219658	0.715904	0.306826	0.7642	
LogDLH(-1)	-0.008254	0.485927	-0.016986	0.9867	
LogPOP60	-0.053150	1.236883	-0.042971	0.9664	
C	-0.981706	3.332496	-0.294586	0.7733	
RESID(-1)	-0.614232	0.403449	-1.235167	0.0858	
RESID(-2)	-0.32895300	0.318016	-1.034389	0.3214	
R-squared		0.270417	Mean dependent var		-3.21E-15
Adjusted R-squared		-0.519965	S.D. dependent var		0.038487
S.E. of regression		0.047449	Akaike info criterion		-2.954594
Sum squared resid		0.027017	Schwarz criterion		-2.277158
Log likelihood		52.40973	Hannan-Quinn criter.		-2.759517
F-statistic		0.342134	Durbin-Watson stat		1.985458
Prob(F-statistic)		0.966812			

TAB. 3.17 – Test d'autocorrélation des erreurs

### Test de normalité des résidus

Pour vérifier l'hypothèse de normalité des résidus des séries, on fait recours au test de normalité de Jarque Bera.

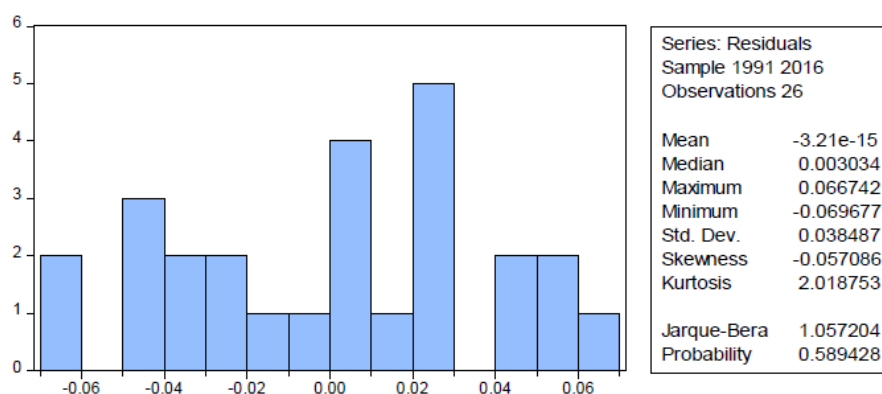


FIG. 3.9 – Test de normalité des résidus

La probabilité associée à la statistique de Jarque-Bera 0.58 est supérieure à 0.05. L'hypothèse de normalité des résidus est donc vérifiée.

### Etude de la stabilité du modèle

Afin de quantifier la qualité de notre modèle, en particulier, la stabilité de l'équation estimée, nous avons appliqué le test de CUSUM. Ce test est fondé sur la somme des résidus. Il représente la courbe de la somme cumulée des résidus ensemble avec 5% des lignes critiques.

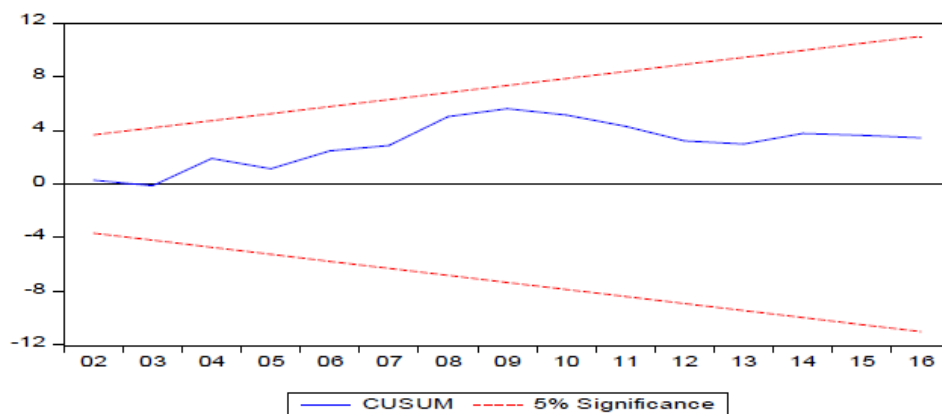


FIG. 3.10 – Courbe de la somme cumulée des résidus (CUSUM)

Nous observons sur le graphique que le CUSUM est à l'intérieure du corridor. On peut donc dire que la relation est stable.

# Conclusion

Dans le premier chapitre, notre intérêt s'est accentué sur des concepts de base sur les séries chronologiques, un passage très important afin de bien comprendre le fondement de la méthode utilisée dans notre étude.

Dans le deuxième chapitre, nous nous sommes basés sur l'approche ARDL, modèle VAR et modèle VECM, des outils statistiques très importants pour l'étude empirique que nous allons voir dans le dernier chapitre.

Dans le dernier chapitre, nous avons exploité la théorie vue dans les chapitres précédents, afin de mener une application, nous avons modélisé les dépenses sanitaires en Algérie à l'aide de l'approche ARDL, cette dernière nous a permis de tester l'existence d'une relation de cointégration entre la variable DS et les autres variables (déterminants) puis de tester la relation à long terme et celle à court terme. Les calculs ainsi que les procédures statistiques ont été réalisés à base du logiciel Eviews10, très connu et utilisé dans la construction des modèles économétriques.

Le bien fondé des résultats théorique a été mis en évidence sur données réelles relatives aux dépenses de santé en Algérie. Ces données sont fournies par l'ONS.

Comme perspectives futures, il serait intéressant d'actualiser cette étude en la complétant par des données récentes. Aussi, effectuer une étude comparative avec des études existantes faites avec d'autres types de modèles et ressortir la supériorité du modèle ARDL.

# Bibliographie

- [1]. Régis BOURBOUNNAIS, Michel TERRAZA, Analyse des séries temporelles, 2<sup>ème</sup> édition DUNOD.
- [2]. Régis BOURBOUNNAIS, Econométrie cours et exercices corrigés, 9<sup>ème</sup> édition DUNOD.
- [3]. Séries chronologiques, Agnès Lagnoux, Université de toulouse, Le MIRAIL, France (pdf).
- [4]. Steven Fortier, Les modèles MA, AR et ARMA.
- [5]. Renforcement statistique, Séries chronologiques, Agnès Lagnoux, Université de toulouse, Le MIRAIL, France (pdf).
- [6]. Jonas KIBALA KUMA, Modélisation ARDL, test de cointégration aux bornes et approche de Toda Yammoto, HAL.
- [7]. Mohamed BOUTAHAR, Séries temporelles, Estimation paramétrique et non paramétrique avec le logiciel R, Université de Marseille, France.
- [8]. Séries chronologiques à une et plusieurs variables : synthèse des méthodes classiques et modèles à base de copules, Mémoire présenté par YAN GIRARD. Université Trois rivières, Quebec, Mai 2011.
- [9]. M-C.Viano, A-Philippe, Maîtrise d'économétrie, cours de séries temporelles (1999-2004), Université de Lille, France.
- [10]. Validation des modèles de séries chronologiques Spatio-temporelles multivariées, Mémoire présenté par Robinson Saint Frard, Juin 2011, Université de Montréal.
- [11]. Introduction aux séries temporelles, Master Mathématiques Appliquées, 2017/2018, Université Paris-Dauphine.

- [12].Richard BEN Jelil, Procédures du test de la racine unitaire.
- [13].Introduction aux séries chronologiques, O.Roustant, Novembre 2008, Ecole Nationale Supérieure des Mines, Saint-Etienne.
- [14].Christophe-Hurlin, Econométrie Appliquée, Séries temporelles, Chapitre 2, Chapitre 3 (Pdf).
- [15].Séries temporelles : Méthodologie de BOX-JENKINS, Mohamed Essaied Hamitra, Décembre 2013, Institut Supérieur de Mathématiques Appliquées et Informatique, KAI-ROUAN.
- [16].Agricultural growth, lical transformation of natural resources and industrialization in TogO KPEMOUA, Pola Kiyém, November 2016, Munich.
- [17].BOUZNIT Mohammed, Rendement du capital humain et dynamique de la croissance au sein des pays sous développés, Thèse de doctorat, ENSSEA, 2016.
- [18].Les processus VAR, cours de séries temporelles multivariées, Gilbert Gilletaz, Janvier 2019.
- [19].C-Chouquet, Modèles linéaires, Université Paul Sabatier, Toulouse, France.
- [20].Peter J.Brockwell-Richard, A.Davis, Times Series : Theory and Methods Springer.